Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

AN ITERATIVE ALGORITHM FOR η-(ANTI)-HERMITIAN LEAST-SQUARES SOLUTIONS OF QUATERNION MATRIX EQUATIONS*

FATEMEH PANJEH ALI BEIK† AND SALMAN AHMADI-ASL†.

Abstract. Recently, some research has been devoted to finding the explicit forms of the η-Hermitian and η-anti-Hermitian solutions of several kinds of quaternion matrix equations and their associated least-squares problems in the literature. Although exploiting iterative algorithms is superior than utilizing the explicit forms in application, hitherto, an iterative approach has not been offered for finding η-(anti)-Hermitian solutions of quaternion matrix equations. The current paper deals with applying an efficient iterative manner for determining η-Hermitian and η-anti-Hermitian least-squares solutions corresponding to the quaternion matrix equation $AXB + CYD = E$. More precisely, first, this paper establishes some properties of the η-Hermitian and η-anti-Hermitian matrices. These properties allow for the demonstration of how the well-known conjugate gradient least-squares (CGLS) method can be developed for solving the mentioned problem over the η-Hermitian and η-anti-Hermitian matrices. In addition, the convergence properties of the proposed algorithm are discussed with details. In the circumstance that the coefficient matrices are ill-conditioned, it is suggested to use a preconditioner for accelerating the convergence behavior of the algorithm. Numerical experiments are reported to reveal the validity of the elaborated results and feasibility of the proposed iterative algorithm and its preconditioned version.

Key words. Quaternion matrix equation, η- Hermitian matrix, η-Anti-Hermitian matrix, Iterative algorithm, Convergence, Preconditioner.

AMS subject classifications. 15A24, 65F10.

1. Introduction.

Let us first present some notations and symbols used throughout this paper. We use $\mathbb{Q}^{m\times n}$ to refer the set of all $m\times n$ matrices over the quaternion ring

$$\mathbb{Q} = \{a_1 + a_2 i + a_3 j + a_4 k \mid i^2 = j^2 = k^2 = ijk = -1,\ a_1, a_2, a_3, a_4 \in \mathbb{R}\},$$

and $\mathbb{R}^{m\times n}$ stands for the set of all $m\times n$ real matrices. For a given $m\times n$ matrix A, the symbols A^T, $\bar{A}$, A^H and $\mathrm{tr}(A)$ are respectively utilized to represent the transpose, the conjugate, the conjugate transpose and the trace of A. The real part of a quaternion a is denoted by $\mathrm{Re}(a)$, i.e., if $a = a_1 + a_2 i + a_3 j + a_4 k$ then $\mathrm{Re}(a) = a_1$. For given $a = a_1 + a_2 i + a_3 j + a_4 k \in \mathbb{Q}$, the *conjugate* of a is defined by $\bar{a} = a_1 - a_2 i - a_3 j - a_4 k$.

*Received by the editors on January 15, 2015. Accepted for publication on June 15, 2015. Handling Editor: James G. Nagy.

†Department of Mathematics, Vali-e-Asr University of Rafsanjan, PO Box 518, Rafsanjan, Iran (f.beik@vru.ac.ir, beik.fatemeh@gmail.com; s.ahmadiasl@stu.vru.ac.ir).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

The *inner product over* $\mathbb{Q}^{m\times n}$ is defined as follows:

$$\langle A, B\rangle = \mathrm{Re}(\mathrm{tr}(B^H A)) \quad \text{for } A, B \in \mathbb{Q}^{m\times n},$$

and the *induced matrix* norm is specified by

$$\|A\| = \sqrt{\langle A, A\rangle} = \sqrt{\mathrm{Re}(\mathrm{tr}(A^H A))}.$$

As a natural way, the *inner product over* $\mathbb{Q}^{m\times n} \times \mathbb{Q}^{m\times n}$ can be elucidated by

$$\langle [A_1, A_2], [B_1, B_2]\rangle = \langle A_1, B_1\rangle + \langle A_2, B_2\rangle,$$

where $[A_1, A_2]$ and $[B_1, B_2]$ belong to $\mathbb{Q}^{m\times n}\times\mathbb{Q}^{m\times n}$. The set of all $n\times n$ real symmetric and anti-symmetric matrices are respectively indicated by $\mathbb{SR}^{n\times n}$ and $\mathbb{ASR}^{n\times n}$. A matrix $A \in \mathbb{Q}^{n\times n}$ is η-*Hermitian* if $A^{\eta H} = A$ and it is called η-*anti-Hermitian* if $A^{\eta H} = -A$ where $A^{\eta H} = -\eta A^H \eta$ and $\eta \in \{i, j, k\}$. The set of all η-Hermitian and η-anti-Hermitian matrices are respectively signified by $\eta\mathbb{HQ}^{n\times n}$ and $\eta\mathbb{AQ}^{n\times n}$.

The linear matrix equations have a crucial role in many branches of applied and pure mathematics such as control and system theory, stability theory, perturbation analysis, etc; for further details one may refer to [1, 2, 9, 18, 42, 43, 44] and the references therein. So far, several types of iterative algorithms have been proposed for solving various kinds of (coupled) matrix equations in the literature. For instance, Beik and Salkuyeh [3] have presented the Global Full Orthogonalization Method (Gl-FOM) and Global Generalized Minimum Residual (Gl-GMRES) method to determine the unique solution group $(X_1, X_2, \ldots, X_p)$ of the subsequent coupled linear matrix equations

$$\sum_{j=1}^{p} A_{ij} X_j B_{ij} = C_i,$$

where the nonsingular matrices $A_{ij} \in \mathbb{R}^{n\times n}$ and $B_{ij} \in \mathbb{R}^{n\times n}$ and the right hand side $C_i \in \mathbb{R}^{n\times n}$ are known matrices for $i, j = 1, 2, \ldots, p$. In [4], an efficient algorithm has been presented to solve the next coupled Sylvester-transpose matrix equations over the generalized centro-symmetric matrices

$$\sum_{j=1}^{q} A_{ij} X_j B_{ij} + C_{ij} X_j^T D_{ij} = C_i, \quad i = 1, \ldots, p,$$

in which $A_{ij}, C_{ij} \in \mathbb{R}^{r_i\times n_j}$, $B_{ij}, D_{ij} \in \mathbb{R}^{n_j\times s_i}$ and $C_i \in \mathbb{R}^{r_i\times s_i}$ are given matrices and the matrices $X_j \in \mathbb{R}^{n_j\times n_j}$ are unknown for $j = 1, 2, \ldots, q$. By making use of convex optimization theory, Cai and Chen [8] have proposed an iterative algorithm for finding the least-squares bisymmetric solutions of the coupled matrix equations

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

$(A_1XB_1, A_2XB_2) = (C_1, C_2)$. Ding and Chen [15, 16, 17] have developed some stationary iterative methods named gradient-based iterative methods to solve some kinds of matrix equations. In [11, 13], Dehghan and Hajarian have developed the well-known conjugate gradient (CG) method to solve the ensuring coupled matrix equations

$$\begin{cases} \sum\limits_{i=1}^{l} A_iXB_i + \sum\limits_{i=1}^{l} C_iYD_i = M \\ \sum\limits_{i=1}^{l} E_iXF_i + \sum\limits_{i=1}^{l} G_iYH_i = N \end{cases}$$

and

$$\begin{cases} A_{11}X_1B_{11} + A_{12}X_2B_{12} + \cdots + A_{1l}X_lB_{1l} = C_1 \\ A_{21}X_1B_{21} + A_{22}X_2B_{22} + \cdots + A_{2l}X_lB_{2l} = C_2 \\ \qquad\vdots \\ A_{l1}X_1B_{l1} + A_{l2}X_2B_{l2} + \cdots + A_{ll}X_lB_{ll} = C_l \end{cases},$$

respectively. In [12], the extended version of the CG method has been examined to solve the succeeding system of matrix equations

$$\sum_{j=1}^{p} A_{ij}X_jB_{ij} = C_i, \qquad i = 1, 2, \ldots, p,$$

over the generalized bisymmetric matrices. In [14], the authors have exploited the idea of the CG method to construct an iterative algorithm to find the generalized reflexive and anti-reflexive solutions of the following system of linear operator equations

$$\begin{cases} \mathcal{F}_1(X) = A_1 \\ \mathcal{F}_2(X) = A_2 \\ \quad\vdots \\ \mathcal{F}_n(X) = A_n \end{cases}.$$

The outline of this paper is organized as follows. Before ending the current section, we momentarily describe our inspiration for presenting this paper and express two main problems of concern. Section 2 is devoted to proving some properties of the η-Hermitian and η-anti-Hermitian matrices. In Section 3, we propose an iterative algorithm to solve our main mentioned problems and analyze the convergence of the proposed algorithm. The preconditioned version of the algorithm is proposed to accelerate the convergence rate of the method in Section 4. In Section 5, some numerical examples are solved to demonstrate the validity of the established results and effectiveness of the proposed algorithm. Finally the paper is ended with a brief conclusion and some suggestions for further works in Section 6.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

1.1. Motivations and highlight points. Recently, there is a growing interest to find the solutions of quaternion matrix equations. A linear quaternion matrix equation is a linear matrix equation over the quaternion ring; see [20, 21, 22, 35, 40, 41] for more details. We would like to comment here that the quaternion matrices and quaternion matrix equations arise in the fields of quantum physics, signal and color image processing, computer science and so on; for further details see [5, 6, 23, 25, 33] and the references therein. Lately, some research has focused on developing iterative methods to solve such problems, for instance one may refer to [22, 28, 35].

Obtaining the explicit forms of the η-Hermitian and η-anti-Hermitian solutions of linear quaternion matrix equations have been studied in [19, 39, 40]. It is well-known that the explicit forms are not computationally practical even in the case that the size of coefficient matrices are moderate. More precisely, Yuan et al. [40] have derived the explicit forms of η-Hermitian and η-anti-Hermitian solutions of $AXB + CYD = E$ and also the explicit forms of the solution pair $[X, Y]$ with least-norm theoretically. Nevertheless there exist some drawbacks for utilizing these forms such as:

- The complicated forms of the obtained explicit solutions and the requirement to compute the Moore-Penrose pseudoinverse which would be expensive for large scale quaternion least-squares problems.

- For deriving solutions by means of obtained explicit forms, Kronecker product of matrices is required to be explicitly formed. Therefore the size of original problem increases and causes high computational cost.

These disadvantages in using explicit forms inspire us to develop an iterative algorithm for computing the η-Hermitian and η-anti-Hermitian solutions of $AXB + CYD = E$. We would like to emphasize that, to the best of our knowledge, a neat and feasible iterative technique to solve such solutions has not been examined so far. In the literature, the well-known CGLS method has been successfully implemented to solve different types of matrix equations; the reader may refer to [26, 27, 29, 30, 31] and the references therein. Hence, our main goal is to extend the CGLS method for resolving the least-squares problem corresponding to the quaternion matrix equation $AXB + CYD = E$ over the η- Hermitian and η-anti-Hermitian matrix pair $[X, Y]$. To this end, we first need to establish some properties of η- Hermitian and η-anti-Hermitian matrices. For improving the convergence speed of the proposed algorithm, the application of a preconditioner is examined.

1.2. Problem reformulation. For simplicity, we consider the linear operator $\mathcal{M} : \mathbb{Q}^{n\times n} \times \mathbb{Q}^{n\times n} \to \mathbb{Q}^{n\times n}$ such that

$$\mathcal{M}(X, Y) = \mathcal{M}_1(X) + \mathcal{M}_2(Y),$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

where $\mathcal{M}_i : \mathbb{Q}^{n\times n} \to \mathbb{Q}^{n\times n}$ for $i = 1, 2$. The linear operators $\mathcal{M}_1$ and $\mathcal{M}_2$ are specified by

$$\mathcal{M}_1(X) = AXB \quad \text{and} \quad \mathcal{M}_2(Y) = CYD. \tag{1.1}$$

Using (1.1), the matrix equation $AXB + CYD = E$ can be reformulated in the following form

$$\mathcal{M}(X, Y) = E. \tag{1.2}$$

Definition 1.1. Let $\mathcal{L}$ be a linear operator from $\mathbb{Q}^{m\times n}$ onto $\mathbb{Q}^{m\times n}$, then the *adjoint of* $\mathcal{L}$ is denoted by $\mathcal{L}^*$ and satisfies

$$\langle \mathcal{L}(X), Z\rangle = \langle X, \mathcal{L}^*(Z)\rangle .$$

By straightforward computations and using the fact that $\operatorname{tr}(XY) = \operatorname{tr}(YX)$, we may conclude that the adjoint operators corresponding to the linear operators $\mathcal{M}_1$ and $\mathcal{M}_2$ have the next forms

$$\mathcal{M}_1^*(Z) = A^H Z B^H \quad \text{and} \quad \mathcal{M}_2^*(Z) = C^H Z D^H,$$

which implies that $\mathcal{M}^* = (\mathcal{M}_1^*, \mathcal{M}_2^*)$.

In this paper, we focus on the solutions of the subsequent two problems.

Problem 1.2. Presume that the $n \times n$ quaternion matrices A, B, C, D and E are given. Find the matrices X and Y such that

$$\|AXB + CYD - E\| \text{ is minimized,} \tag{1.3}$$

and $X \in \eta\mathbb{H}\mathbb{Q}^{n\times n}$ and $Y \in \eta\mathbb{A}\mathbb{Q}^{n\times n}$.

Problem 1.3. Presume that the $n \times n$ quaternion matrices A, B, C, D and E are given. Suppose that S_E stands for the solution set of Problem 1.2. Given $\tilde{X} \in \eta\mathbb{H}\mathbb{Q}^{n\times n}$ and $\tilde{Y} \in \eta\mathbb{A}\mathbb{Q}^{n\times n}$, find $X^* \in \eta\mathbb{H}\mathbb{Q}^{n\times n}$ and $Y^* \in \eta\mathbb{A}\mathbb{Q}^{n\times n}$ such that

$$\left\|\tilde{X} - X^*\right\| + \left\|\tilde{Y} - Y^*\right\| = \min_{[X,Y]\in S_E} \left\|\tilde{X} - X\right\| + \left\|\tilde{Y} - Y\right\|.$$

2. On the η-Hermitian and η-anti-Hermitian matrices. In this section, we scrutinize some properties of the η-Hermitian and η-anti-Hermitian matrices which have an essential role for constructing our algorithm to solve the mentioned problems. Exploiting the derived results, two new linear operators are also expounded.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

For an arbitrary given $A \in \mathbb{R}^{n\times n}$, suppose that the operators $\mathcal{H}(A)$ and $\mathcal{S}(A)$ give the Hermitian and skew-Hermitian parts of matrix A, respectively. That is,

$$\begin{aligned} \mathcal{H} : \mathbb{R}^{n\times n} &\to \mathbb{SR}^{n\times n} \\ A &\mapsto \mathcal{H}(A) = \frac{1}{2}(A + A^T), \end{aligned}$$

and,

$$\begin{aligned} \mathcal{S} : \mathbb{R}^{n\times n} &\to \mathbb{ASR}^{n\times n} \\ A &\mapsto \mathcal{S}(A) = \frac{1}{2}(A - A^T). \end{aligned}$$

The following theorem reveals that the set of all $n \times n$ quaternion matrices can be written as a direct sum of $\eta\mathbb{HQ}^{n\times n}$ and $\eta\mathbb{AQ}^{n\times n}$. The theorem supplies a cardinal tool for extending the CGLS algorithm to solve Problems 1.2 and 1.3.

THEOREM 2.1. *Presume that $\eta\mathbb{HQ}^{n\times n}$ and $\eta\mathbb{AQ}^{n\times n}$ stand for the set of all $n \times n$ η-Hermitian and η-anti-Hermitian matrices, respectively. Then,*

$$\mathbb{Q}^{n\times n} = \eta\mathbb{HQ}^{n\times n} \oplus \eta\mathbb{AQ}^{n\times n}.$$

Proof. We only demonstrate the validity of the assertion for $\eta = i$, the strategy of the proof for $\eta = j$ $(\eta = k)$ is similar. For clarification, the proof is divided into three steps as follows:

Step 1. In this step, we show that for an arbitrary given matrix $A \in \mathbb{Q}^{n\times n}$, there exist matrices $B \in i\mathbb{HQ}^{n\times n}$ and $C \in i\mathbb{AQ}^{n\times n}$ such that

$$A = B + C. \tag{2.1}$$

To do so, it is required to characterize the structure of the members of $i\mathbb{HQ}^{n\times n}$ and $i\mathbb{AQ}^{n\times n}$. Let $B = B_1 + B_2 i + B_3 j + B_4 k$ be an i-Hermitian matrix, i.e, $-iB^H i = B$. Straightforward computations reveal that

$$B_1 + B_2 i + B_3 j + B_4 k = B_1^T - B_2^T i + B_3^T j + B_4^T k. \tag{2.2}$$

In view of (2.2), it can be deduced that

$$B_1^T = B_1, \quad B_2^T = -B_2, \quad B_3^T = B_3 \quad \text{and} \quad B_4^T = B_4.$$

On the other hand if $B_1^T = B_1$, $B_2^T = -B_2$, $B_3^T = B_3$ and $B_4^T = B_4$, then it is not difficult to verify that $B = B_1 + B_2 i + B_3 j + B_4 k$ is an i-Hermitian matrix. With a similar manner, it can be seen that $C \in i\mathbb{AQ}^{n\times n}$ (i.e., $iC^H i = C$) if and only

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

if $C_1^T = -C_1$, $C_2^T = C_2$, $C_3^T = -C_3$, and $C_4^T = -C_4$. Consequently, we may set $A = B + C$ in which

$$\begin{aligned} B &= \mathcal{H}(A_1) + \mathcal{S}(A_2)\, i + \mathcal{H}(A_3)\, j + \mathcal{H}(A_4)\, k, \\ C &= \mathcal{S}(A_1) + \mathcal{H}(A_2)\, i + \mathcal{S}(A_3)\, j + \mathcal{S}(A_4)\, k, \end{aligned}$$

where the operators $\mathcal{H}$ and $\mathcal{S}$ are defined as before. Evidently, $B \in i\mathbb{HQ}^{n\times n}$ and $C \in i\mathbb{AQ}^{n\times n}$.

Step 2. Here it is demonstrated that if $B \in i\mathbb{HQ}^{n\times n}$ and $C \in i\mathbb{AQ}^{n\times n}$ then $\langle B, C\rangle = 0$.

Suppose that $B = B_1 + B_2 i + B_3 j + B_4 k$ and $C = C_1 + C_2 i + C_3 j + C_4 k$, then

$$\begin{aligned} C^H B &= \left(C_1^T - C_2^T i - C_3^T j - C_4^T k\right)(B_1 + B_2 i + B_3 j + B_4 k) \\ &= \left(C_1^T B_1 + C_2^T B_2 + C_3^T B_3 + C_4^T B_4\right) + \left(C_1^T B_2 - C_2^T B_1 - C_3^T B_4 + C_4^T B_3\right) i \\ &\quad + \left(C_1^T B_3 + C_2^T B_4 - C_3^T B_1 - C_4^T B_2\right) j + \left(C_1^T B_4 - C_2^T B_3 + C_3^T B_2 - C_4^T B_1\right) k. \end{aligned}$$

Thence, we have

$$\begin{aligned} \langle B, C\rangle &= \mathrm{Re}\left(\mathrm{tr}\left(C^H B\right)\right) \\ &= \mathrm{tr}\left(C_1^T B_1 + C_2^T B_2 + C_3^T B_3 + C_4^T B_4\right). \end{aligned}$$

From the earlier discussions in the first step, it is not onerous to see that

$$\mathrm{Re}\left(\mathrm{tr}\left(C_i^T B_i\right)\right) = 0 \quad \text{for } i = 1, 2, 3, 4.$$

Step 3. In this part, it is shown that the splitting (2.1) is unique. In fact, if $A = B + C = B' + C'$ such that $B, B' \in i\mathbb{HQ}^{n\times n}$ and $C, C' \in i\mathbb{AQ}^{n\times n}$ then $B - B' = C - C'$. On the other hand, $B - B' \in i\mathbb{HQ}^{n\times n}$ and $C - C' \in i\mathbb{AQ}^{n\times n}$. Now Step 2 implies that $B = B'$ and $C = C'$. The result can be concluded immediately from Steps 1, 2 and 3. ∎

The following corollary is a direct conclusion of Theorem 2.1.

Corollary 2.2. *For each $A \in \mathbb{Q}^{n\times n}$, there exist unique matrices $U_1 \in \eta\mathbb{HQ}^{n\times n}$ and $U_2 \in \eta\mathbb{AQ}^{n\times n}$ such that $A = U_1^\eta + U_2^\eta$ and $\langle U_1^\eta, U_2^\eta\rangle = 0$.*

Based on the above lemma, we introduce two useful linear operators as follows:

$$\begin{aligned} \mathcal{L}_1^\eta : \mathbb{Q}^{n\times n} &\to \eta\mathbb{HQ}^{n\times n} \\ U &\mapsto \mathcal{L}_1^\eta(U) = U_1^\eta, \end{aligned}$$

and

$$\begin{aligned} \mathcal{L}_2^\eta : H^{n\times n} &\to \eta\mathbb{AQ}^{n\times n} \\ U &\mapsto \mathcal{L}_2^\eta(U) = U_2^\eta. \end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

We finish the current section by presenting the next two useful remarks.

Remark 2.3. In view of Theorem 2.1, we conclude that for given arbitrary matrices $U \in \mathbb{Q}^{n\times n}$, $Q \in \eta\mathbb{HQ}^{n\times n}$ and $W \in \eta\mathbb{AQ}^{n\times n}$, we have

$$\langle U, Q\rangle = \langle \mathcal{L}_1^\eta(U), Q\rangle \quad \text{and} \quad \langle U, W\rangle = \langle \mathcal{L}_2^\eta(U), W\rangle . \tag{2.3}$$

Remark 2.4. Note that Theorem 2.1 helps us to derive the explicit forms of the operators $\mathcal{L}_1^\eta(U)$ and $\mathcal{L}_2^\eta(U)$. For instance, consider the case that $\eta = i$. Then,

$$\begin{aligned} \mathcal{L}_1^\eta(U) &= \mathcal{H}(U_1) + \mathcal{S}(U_2)\, i + \mathcal{H}(U_3)\, j + \mathcal{H}(U_4)\, k, \\ \mathcal{L}_2^\eta(U) &= \mathcal{S}(U_1) + \mathcal{H}(U_2)\, i + \mathcal{S}(U_3)\, j + \mathcal{S}(U_4)\, k, \end{aligned}$$

where the operators $\mathcal{H}$ and $\mathcal{S}$ are defined as before and $U = U_1 + U_2 i + U_3 j + U_4 k$.

3. Proposed iterative scheme. The well-known CGLS method for resolving the following least-squares problem

$$\|Ax - b\|_2 ,$$

is obtained by applying the conjugate gradient (CG) algorithm on the normal equations $A^T Ax = A^T b$ and the derived algorithm is called CGLS; see [7, 36] for more details. We would like to comment here that the CGLS method is also known as the CGNR method; for more details see [32, Chapter 8].

As a matter of fact, the CGLS method is an oblique projection technique onto

$$\mathcal{K}_n = \left\{ r_0, \left(A^T A\right) r_0, \ldots, \left(A^T A\right)^{n-1} r_0 \right\},$$

and orthogonal to $\mathcal{L}_n = A\mathcal{K}_n$ where $r_0 = A^T b - A^T Ax_0$ and x_0 is a given initial guess. It can be theoretically shown that in the absent of round-off errors, the CGLS method converges to the exact solution of the normal equations $A^T Ax = A^T b$ within finite number of steps [32].

The main objective of the present section is to exploit the idea of the CGLS approach for constructing an iterative manner for resolving Problems 1.2 and 1.3 and study its convergence properties. To this end, we mainly utilize the results obtained in the previous section. In the rest of this section, first, an iterative algorithm is offered for determining the solution set of Problem 1.2. Afterward an approach is proposed to solve the second problem.

3.1. An iterative algorithm and its convergence properties. Developing the idea of the CGLS method and using the linear operators introduced in the previous

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

section, we may propose an iterative algorithm to determine the least–squares the η-Hermitian and η-anti-Hermitian solution pair of the linear matrix equation $AXB + CYD = E$. The proposed algorithm is given by Algorithm 3.1.

ALGORITHM 3.1. *The developed CGLS method for* Problem 1.2.
Data: *Input* $A \in \mathbb{Q}^{n\times n}$, $B \in \mathbb{Q}^{n\times n}$, $C \in \mathbb{Q}^{n\times n}$, $D \in \mathbb{Q}^{n\times n}$, $E \in \mathbb{Q}^{n\times n}$, $X(0) \in \eta\mathbb{H}\mathbb{Q}^{n\times n}$, $Y(0) \in \eta\mathbb{A}\mathbb{Q}^{n\times n}$ *and choose tolerance* ϵ.
Initialization:

- $k = 0$;
- $R(0) = E - AX(0)B - CY(0)D$;
- $P_{0,x} = \mathcal{L}_1^{\eta}\left(\mathcal{M}_1^*\left(R(0)\right)\right)$;
- $P_{0,y} = \mathcal{L}_2^{\eta}\left(\mathcal{M}_2^*\left(R(0)\right)\right)$;
- $Q_{0,x} = P_{0,x}$;
- $Q_{0,y} = P_{0,y}$;

While $\|R^{(0)}\| > \epsilon$ *or* $\|P_{0,x}\|^2 + \|P_{0,y}\|^2 > \epsilon$ *Do:*

- $\alpha_k = \frac{\|P_{k,x}\|^2+\|P_{k,y}\|^2}{\|\mathcal{M}(Q_{k,x},Q_{k,y})\|^2}$;
- $X(k+1) = X(k) + \alpha_k Q_{k,x}$;
- $Y(k+1) = Y(k) + \alpha_k Q_{k,y}$;
- $R(k+1) = R(k) - \alpha_k \mathcal{M}\left(Q_{k,x}, Q_{k,y}\right)$;
- $P_{k+1,x} = \mathcal{L}_1^{\eta}\left(\mathcal{M}_1^*\left(R(k+1)\right)\right)$;
- $P_{k+1,y} = \mathcal{L}_2^{\eta}\left(\mathcal{M}_2^*\left(R(k+1)\right)\right)$;
- $\beta_k = \frac{\|P_{k+1,x}\|^2+\|P_{k+1,y}\|^2}{\|P_{k,x}\|^2+\|P_{k,y}\|^2}$;
- $Q_{k+1,x} = P_{k+1,x} + \beta_k Q_{k,x}$;
- $Q_{k+1,y} = P_{k+1,y} + \beta_k Q_{k,y}$;
- $k = k + 1$;

EndDo

We will inspect the convergence behavior of Algorithm 3.1. First we need to recall the next lemma which is called "Projection Theorem" and its proof can be found in [37].

LEMMA 3.2. *Let* $\mathbb{X}$ *be a finite-dimensional inner product space,* $\mathbb{M}$ *be a subspace of* $\mathbb{X}$*, and* $\mathbb{M}^{\perp}$ *be the orthogonal complement of* $\mathbb{M}$*. For a given* $x \in \mathbb{X}$*, always, there exists an* $m_0 \in \mathbb{M}$ *such that* $\|x - m_0\| \leq \|x - m\|$ *for all* $m \in \mathbb{M}$ *where* $\|.\|$ *is the norm associated with the inner product defined in* $\mathbb{X}$*. Moreover,* $m_0 \in \mathbb{M}$ *is the unique minimization vector in* $\mathbb{M}$ *if and only if* $(x - m_0) \perp \mathbb{M}$ *which is equivalent to say that* $(x - m_0) \in \mathbb{M}^{\perp}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

THEOREM 3.3. *Suppose that* $\tilde{R} = E - A\tilde{X}B - C\tilde{Y}D$ *where* $[\tilde{X}, \tilde{Y}] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$. *If* $\mathcal{L}_i^\eta(\mathcal{M}_i^*(\tilde{R})) = 0$ *for* $i = 1, 2$, *then* $[\tilde{X}, \tilde{Y}]$ *is a solution pair of* Problem 1.2.

Proof. Consider the following linear subspace

$$\mathbb{V} = \left\{ V \;\middle|\; V = AXB + CYD,\ X \in \eta\mathbb{HQ}^{n\times n} \text{ and } Y \in \eta\mathbb{AQ}^{n\times n} \right\}.$$

Presume that $\tilde{X} \in \eta\mathbb{HQ}^{n\times n}$ and $\tilde{Y} \in \eta\mathbb{AQ}^{n\times n}$, then $\tilde{E} = A\tilde{X}B + C\tilde{Y}D \in \mathbb{V}$. Lemma 3.2 implies that $[\tilde{X}, \tilde{Y}]$ is a least-squares solution of $AXB + CYD = E$ if and only if

$$\left\langle E - \tilde{E}, AXB + CYD \right\rangle = \left\langle \tilde{R}, AXB + CYD \right\rangle = 0, \tag{3.1}$$

for all $X \in \eta\mathbb{HQ}^{n\times n}$ and $Y \in \eta\mathbb{AQ}^{n\times n}$. On the other hand, invoking (2.3), we have

$$\begin{aligned} \left\langle \tilde{R}, AXB + CYD \right\rangle &= \left\langle A^H\tilde{R}B^H, X \right\rangle + \left\langle C^H\tilde{R}D^H, Y \right\rangle \\ &= \left\langle \mathcal{L}_1^\eta\left(A^H\tilde{R}B^H\right), X \right\rangle + \left\langle \mathcal{L}_2^\eta\left(C^H\tilde{R}D^H\right), Y \right\rangle \\ &= \left\langle \mathcal{L}_1^\eta(\mathcal{M}_1^*(\tilde{R})), X \right\rangle + \left\langle \mathcal{L}_2^\eta(\mathcal{M}_2^*(\tilde{R})), Y \right\rangle. \end{aligned}$$

Consequently, from (3.1), the assumptions $\mathcal{L}_1^\eta(\mathcal{M}_1^*(\tilde{R})) = 0$ and $\mathcal{L}_2^\eta(\mathcal{M}_2^*(\tilde{R})) = 0$ ensure that $[\tilde{X}, \tilde{Y}]$ is the least-squares solution pair of $AXB + CYD = E$ which completes the proof. $\square$

LEMMA 3.4. [24] *Let* $f(Z)$ *be a continuous, differentiable and convex function on subspace* $\mathbb{Y}$, *then there exists* $Z^* \in \mathbb{Y}$ *such that* $f(Z^*) = \min_{Z\in\mathbb{Y}} f(Z)$ *if and only if the projection of the gradient* $\nabla f(Z^*)$ *onto* $\mathbb{Y}$ *equals to zero.*

REMARK 3.5. Using Lemma 3.4, we can present an alternative proof for Theorem 3.3. To this end, we define

$$\begin{aligned} f &: \mathbb{Q}^{n\times n} \times \mathbb{Q}^{n\times n} \to \mathbb{R} \\ &[X, Y] \mapsto f(X, Y) = \|AXB + CYD - E\|^2. \end{aligned} \tag{3.2}$$

It is not difficult to establish that $f(X, Y)$ is a continuous, differentiable and convex function. With a similar approach used in [26], the gradient of $f(X, Y)$ can be obtained. To do so, consider the next auxiliary function

$$\begin{aligned} g &: \mathbb{R} \to \mathbb{R} \\ &t \mapsto g(t) = f(X + tP_1, Y + tP_2) \end{aligned} \tag{3.3}$$

in which P_1 and P_2 are arbitrary quaternion matrices. It can be seen that

$$g'(0) = \langle \nabla_X f(X, Y), P_1 \rangle + \langle \nabla_Y f(X, Y), P_2 \rangle. \tag{3.4}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

From (3.2), (3.3) and some straightforward computations, we get

$$
\begin{aligned}
g'(0) &= 2\left\langle AXB + CYD - E, AP_1B\right\rangle \\
&\quad +2\left\langle AXB + CYD - E, CP_2D\right\rangle \\
&= 2\left\langle A^H\left(AXB + CYD - E\right)B^H, P_1\right\rangle \\
&\quad +2\left\langle C^H\left(AXB + CYD - E\right)D^H, P_2\right\rangle .
\end{aligned} \tag{3.5}
$$

In view of (3.4) and (3.5), it turns out that

$$\nabla_X f(X,Y) = 2A^H\left(AXB + CYD - E\right)B^H = -2A^HRB^H,$$

$$\nabla_Y f(X,Y) = 2C^H\left(AXB + CYD - E\right)D^H = -2C^HRD^H.$$

Now from the discussion of this remark, it is not difficult to conclude that if

$$\mathcal{L}_1^\eta\left(A^HRB^H\right) = 0 \quad \text{and} \quad \mathcal{L}_2^\eta\left(C^HRD^H\right) = 0,$$

where $R = E - AXB - CYD$, then $[X,Y]$ is the minimizer of (3.2) over the η-Hermitian and η-anti-Hermitian matrices.

The next theorem can be proved by mathematical induction and its proof is elaborated in Appendix A.

THEOREM 3.6. *Suppose that k steps of* Algorithm 3.1 *have been performed, i.e., $\|P_{l,x}\|^2 + \|P_{l,y}\|^2 \neq 0$ and $\mathcal{M}(Q_{l,x}, Q_{l,y}) \neq 0$ for $l = 0, 1, \ldots, k$. The sequences $P_{l,x}, P_{l,y}, Q_{l,x}$ and $Q_{l,y}$ $(l = 0, 1, \ldots, k)$ produced by* Algorithm 3.1 *satisfy the following statements*

$$\langle P_{i,x}, P_{j,x}\rangle + \langle P_{i,y}, P_{j,y}\rangle = 0, \tag{3.6}$$

$$\langle \mathcal{M}(Q_{i,x}, Q_{i,y}), \mathcal{M}(Q_{j,x}, Q_{j,y})\rangle = 0, \tag{3.7}$$

$$\langle P_{i,x}, Q_{j,x}\rangle + \langle P_{i,y}, Q_{j,y}\rangle = 0, \tag{3.8}$$

for $i, j = 0, 1, 2, \ldots, k$ $(i \neq j)$.

REMARK 3.7. Note that if in a specific step of Algorithm 3.1, say l-th step, we face to the situation that $\|P_{l,x}\|^2 + \|P_{l,y}\|^2 = 0$. Then,

$$P_{l,x} = \mathcal{L}_1^\eta\left(\mathcal{M}_1^*\left(R(l)\right)\right) = 0 \quad \text{and} \quad P_{l,y} = \mathcal{L}_2^\eta\left(\mathcal{M}_2^*\left(R(l)\right)\right) = 0,$$

which, in view of Theorem 3.3, implies that $[X(l), Y(l)]$ is a solution pair of Problem 1.2. On the other hand, the subsequent computations reveal that if $\mathcal{M}(Q_{l,x}, Q_{l,y}) = 0$, then $\|P_{l,x}\|^2 + \|P_{l,y}\|^2 = 0$,

$$
\begin{aligned}
\left\langle \mathcal{M}\left(Q_{l,x}, Q_{l,y}\right), R(l)\right\rangle &= \left\langle \mathcal{M}_1\left(Q_{l,x}\right) + \mathcal{M}_2\left(Q_{l,y}\right), R(l)\right\rangle \\
&= \left\langle Q_{l,x}, \mathcal{L}_1^\eta\left(\mathcal{M}_1^*\left(R(l)\right)\right)\right\rangle + \left\langle Q_{l,y}, \mathcal{L}_2^\eta\left(\mathcal{M}_2^*\left(R(l)\right)\right)\right\rangle \\
&= \left\langle Q_{l,x}, P_{l,x}\right\rangle + \left\langle Q_{l,y}, P_{l,y}\right\rangle \\
&= \|P_{l,x}\|^2 + \|P_{l,y}\|^2.
\end{aligned}
$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

The following theorem can be deduced from Theorem 3.6 immediately. The proof of the theorem is straightforward and is omitted.

THEOREM 3.8. *For any initial matrix pair* $[X(0), Y(0)] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$, Algorithm 3.1 *converges to the exact solution of* Problem 1.2 *within finite number of steps in the absence of round-off errors.*

In the sequent part of this subsection, our goal is to demonstrate that the least-norm solution of Problem 1.2 can be obtained by choosing appropriate initial matrices. To this end, let us define the next linear subspace over $\eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$,

$$\mathbb{W} = \left\{ [W_1, W_2] \;\middle|\; W_1 = \mathcal{L}_1^\eta\left(A^H Z B^H\right) \text{ and } W_2 = \mathcal{L}_2^\eta\left(C^H Z D^H\right) \right\}. \tag{3.9}$$

where $Z \in \mathbb{Q}^{n\times n}$ is an arbitrary given matrix.

LEMMA 3.9. *Let* $[\tilde{X}, \tilde{Y}]$ *be a solution pair of* Problem 1.2. *Suppose that* $[N_1, N_2] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$. *Then,* $[\tilde{X} + N_1, \tilde{Y} + N_2]$ *is a solution pair of* Problem 1.2 *if and only if* $AN_1B + CN_2D = 0$.

Proof. If $[\tilde{X} + N_1, \tilde{Y} + N_2] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$ and $AN_1B + CN_2D = 0$, then

$$\left\| A\left(\tilde{X} + N_1\right) B + C\left(\tilde{Y} + N_2\right) D - E \right\|^2 = \left\| A\tilde{X}B + C\tilde{Y}D - E \right\|^2,$$

which ensures that $[\tilde{X} + N_1, \tilde{Y} + N_2]$ is a solution pair of Problem 1.2. Conversely, suppose that $[\tilde{X} + N_1, \tilde{Y} + N_2] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$ is a solution pair of Problem 1.2. By the hypophysis, $[\tilde{X}, \tilde{Y}]$ is a solution pair of Problem 1.2, therefore

$$\left\| A\tilde{X}B + C\tilde{Y}D - E \right\|^2 = \left\| A\left(\tilde{X} + N_1\right) B + C\left(\tilde{Y} + N_2\right) D - E \right\|^2. \tag{3.10}$$

As $[\tilde{X}, \tilde{Y}]$ is a least-squares solution of $AXB + CYD = E$, Lemma 3.2 implies that

$$\left\langle A\tilde{X}B + C\tilde{Y}D - E, AXB + CYD \right\rangle = 0.$$

In view of the above relation and the next computations, we have

$$\begin{aligned} \left\| A\left(\tilde{X} + N_1\right) B + C\left(\tilde{Y} + N_2\right) D - E \right\|^2 &= \left\| A\tilde{X}B + C\tilde{Y}D - E \right\|^2 + \|AN_1B + CN_2D\|^2 \\ &\quad - 2\left\langle A\tilde{X}B + C\tilde{Y}D - E, AN_1B + CN_2D \right\rangle \\ &= \left\| A\tilde{X}B + C\tilde{Y}D - E \right\|^2 + \|AN_1B + CN_2D\|^2. \end{aligned}$$

Now the result follows from (3.10) immediately. $\square$

THEOREM 3.10. *If the initial matrix pair* $[X(0), Y(0)]$ *belongs to* $\mathbb{W}$, *then* Algorithm 3.1 *converges to the least-norm solution pair of* Problem 1.2.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

Proof. Suppose that the initial matrices are chosen such that $[X(0), Y(0)] \in \mathbb{W}$. Presume that ℓ steps of Algorithm 3.1 have been performed. In this case, for the sequence of approximate solutions $\{X(j)\}_{j=1}^{\ell}$ and $\{Y(j)\}_{j=1}^{\ell}$ produced by Algorithm 3.1, we have $[X(j), Y(j)] \in \mathbb{W}$ for some $Z \in \mathbb{Q}^{n\times n}$. Suppose that $X(j) \to X^*$ and $Y(j) \to Y^*$ as $j \to \infty$. From Lemma 3.9, it reveals that any arbitrary solution pair $[X, Y]$ of Problem 1.2 can be expressed as $[X^* + N_1, Y^* + N_2]$ such that $[N_1, N_2] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$ and $AN_1B + CN_2D = 0$. On the other hand,

$$\begin{aligned}
\langle X^*, N_1\rangle + \langle Y^*, N_2\rangle &= \left\langle \mathcal{L}_1^{\eta}\left(A^H ZB^H\right), N_1\right\rangle + \left\langle \mathcal{L}_2^{\eta}\left(C^H ZD^H\right), N_2\right\rangle \\
&= \left\langle A^H ZB^H, N_1\right\rangle + \left\langle C^T ZD^T, N_2\right\rangle \\
&= \langle Z, AN_1B\rangle + \langle Z, CN_2D\rangle \\
&= \langle Z, AN_1B + CN_2D\rangle . \\
&= 0.
\end{aligned}$$

Hence, we get

$$\begin{aligned}
\|X^* + N_1\|^2 + \|Y^* + N_2\|^2 &= \|X^*\|^2 + \|Y^*\|^2 + \|N_1\|^2 + \|N_2\|^2 \\
&\geq \|X^*\|^2 + \|Y^*\|^2.
\end{aligned}$$

This implies that the solution pair $[X^*, Y^*]$ is the least-norm solution pair of Problem 1.2. □

The next theorem illustrates that the sequence of approximate solutions generated via Algorithm 3.1 satisfies a minimization property.

THEOREM 3.11. *Let $[X(0), Y(0)] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$ be an arbitrary initial matrix pair, then the generated matrix solution pair $[X(k), Y(k)]$ at the k-th iteration step is the minimizer of the following optimization problem*

$$\min_{[X,Y]\in U_k} \|AXB + CYD - E\|^2,$$

where the affine subspace U_k is expounded as follows:

$$U_k = [X(0), Y(0)] + span\{[Q_{0,x}, Q_{0,y}], [Q_{1,x}, Q_{1,y}], \ldots, [Q_{k-1,x}, Q_{k-1,y}]\}. \tag{3.11}$$

Proof. Consider an arbitrary matrix pair $[X, Y] \in U_k$. From (3.11), there exist real numbers $\beta_0, \beta_2, \ldots, \beta_{k-1}$ such that

$$[X, Y] = [X(0), Y(0)] + \sum_{i=0}^{k-1} \beta_i [Q_{i,x}, Q_{i,y}].$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

We define the next function with respect to $\beta_0, \beta_1, \ldots, \beta_{k-1}$,

$$\begin{aligned} f(\beta_0,\beta_1,\ldots,\beta_{k-1}) &= \|AXB + CYD - E\|^2 \\ &= \left\| A\left(X(0) + \sum_{i=0}^{k-1} \beta_i Q_{i,x}\right) B + C\left(Y(0) + \sum_{i=0}^{k-1} \beta_i Q_{i,y}\right) D - E\right\|^2. \end{aligned} \tag{3.12}$$

It is well-known that the function (3.12) is continuous and differentiable with respect to the variables $\beta_0, \beta_1, \ldots, \beta_{k-1}$. Straightforward computations reveal that

$$\begin{aligned} f(\beta_0,\beta_2,\ldots,\beta_{k-1}) &= \left\| A\left(X(0) + \sum_{i=0}^{k-1} \beta_i Q_{i,x}\right) B + C\left(Y(0) + \sum_{i=0}^{k-1} \beta_i Q_{i,y}\right) D - E\right\|^2 \\ &= \left\| AX(0)B + CY(0)D - E + \sum_{i=0}^{k-1} \beta_i AQ_{i,x}B + \sum_{i=0}^{k-1} \beta_i CQ_{i,y}D \right\|^2 \\ &= \|R(0)\|^2 + \sum_{i=0}^{k-1} \beta_i^2 \|AQ_{i,x}B + CQ_{i,y}D\|^2 \\ &\quad -2\sum_{i=0}^{k-1} \beta_i \left\langle R(0), AQ_{i,x}B + CQ_{i,y}D \right\rangle, \end{aligned}$$

where $R(0) = E - AX(0)B - CY(0)D$.

From Algorithm 3.1, it is seen that $R(0)$ can be written as

$$\begin{aligned} R(0) = R\,(i) &+ \gamma_{i-1}\left(AQ_{i-1,x}B + CQ_{i-1,y}D\right) \\ &+ \cdots + \gamma_0\left(AQ_{0,x}B + CQ_{0,y}D\right), \end{aligned} \tag{3.13}$$

for some $\gamma_0, \gamma_1, \ldots, \gamma_{i-1}$. Note that Lemma 3.4 demonstrates that the necessary and sufficient conditions for minimizing $f(\beta_0, \beta_1, \ldots, \beta_{k-1})$ are

$$\frac{\partial f(\beta_0,\beta_1,\ldots,\beta_{k-1})}{\partial \beta_i} = 0, \qquad i = 0, 1, \ldots, k-1.$$

Hence, the optimal solution is obtained as soon as

$$\beta_i = \frac{\left\langle R(0), AQ_{i,x}B + CQ_{i,y}D\right\rangle}{\|AQ_{i,x}B + CQ_{i,y}D\|^2}, \qquad i = 0, 1, \ldots, k-1. \tag{3.14}$$

Using Eqs. (3.7), (3.8), (3.13) and some computations, we can simplify (3.14) as follows:

$$\begin{aligned} \beta_i &= \frac{\left\langle R(0), AQ_{i,x}B + CQ_{i,y}D\right\rangle}{\|AQ_{i,x}B + CQ_{i,y}D\|^2} = \frac{\left\langle R(i), AQ_{i,x}B + CQ_{i,y}D\right\rangle}{\|AQ_{i,x}B + CQ_{i,y}D\|^2} \\ &= \frac{\left\langle \mathcal{L}_1^\eta\left(A^H R(i) B^H\right), Q_{i,x}\right\rangle + \left\langle \mathcal{L}_2^\eta\left(C^H R(i) D^H\right), Q_{i,y}\right\rangle}{\|AQ_{i,x}B + CQ_{i,y}D\|^2} \\ &= \frac{\langle P_{i,x}, Q_{i,x}\rangle + \langle P_{i,y}, Q_{i,y}\rangle}{\|AQ_{i,x}B + CQ_{i,y}D\|^2} = \frac{\|P_{i,x}\|^2 + \|P_{i,y}\|^2}{\|AQ_{i,x}B + CQ_{i,y}D\|^2} = \alpha_i, \qquad i = 0, 1, \ldots, k-1. \end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

Now the result follows from the following fact immediately,

$$\min_{\beta_i} f(\beta_0, \beta_1, \ldots, \beta_{k-1}) = \min_{[X,Y] \in U_k} \|AXB + CYD - E\|^2. \quad \square$$

REMARK 3.12. Let us define the sets U_{k-1} and U_k as follows:

$$U_{k-1} = [X(0), Y(0)] + span\{[Q_{0,x}, Q_{0,y}], [Q_{1,x}, Q_{1,y}], \ldots, [Q_{k-1,x}, Q_{k-1,y}]\},$$
$$U_k = [X(0), Y(0)] + span\{[Q_{0,x}, Q_{0,y}], [Q_{1,x}, Q_{1,y}], \ldots, [Q_{k,x}, Q_{k,y}].$$

It is obvious that $U_{k-1} \subseteq U_k$ which illustrates that

$$\|R(k)\|^2 = \min_{U_k} \|AXB + CYD - E\|^2 \leq \min_{U_{k-1}} \|AXB + CYD - E\|^2 = \|R(k-1)\|^2.$$

Evidently, the above inequality demonstrates that the sequence of the norm of residuals produced by Algorithm 3.1 is decreasing which shows that the algorithm is convergent.

3.2. The solution of Problem 2. In this subsection, we briefly express how the second problem can be solved using Algorithm 3.1. Suppose that a matrix pair $[\hat{X}, \hat{Y}] \in \eta\mathbb{HQ}^{n\times n} \times \eta\mathbb{AQ}^{n\times n}$ is given. Evidently, we may find the solution pair of Problem 1.3 by finding the least-norm solution pair $[Z, W]$ of the following least-squares problem

$$\min \|AZB + CWD - F\|, \tag{3.15}$$

where $Z = X - \hat{X}$, $W = Y - \hat{Y}$ and $F = E - A\hat{X}B - C\hat{Y}D$.

To this end, we exploit Algorithm 3.1 with the initial matrices $[Z(0), W(0)] \in \mathbb{W}$, e.g. $[Z(0), W(0)] = [0, 0]$, for determining the least-norm solution pair $[Z^*, W^*]$. Then the solution pair of Problem 1.3 can be obtained by setting $[X^*, Y^*] = [\hat{X} + Z^*, \hat{Y} + W^*]$.

4. Application of a preconditioner. In this section, we propose a preconditioned form of Algorithm 3.1 for solving Problem 1.2 to accelerate the convergence of the algorithm. Exploiting the offered preconditioner can be profitable when the quaternion coefficient matrices are extremely ill-conditioned. First, we need to recall some theorems and definitions.

THEOREM 4.1. [38] *Let $A \in \mathbb{Q}^{m\times n}$ with rank$(A) = r$. Then there exist unitary quaternion matrices $U \in \mathbb{Q}^{m\times n}$ and $V \in \mathbb{Q}^{n\times n}$ such that*

$$U^H AV = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

where $\Sigma_r = diag\,(\sigma_1, \sigma_2, \ldots, \sigma_r)\,,\, \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,$ *and* $\sigma_1, \sigma_2, \ldots, \sigma_r$ *are the nonzero singular values of quaternion matrix* A.

Let $A = A_1 + iA_2 + jA_3 + kA_4 \in \mathbb{Q}^{m\times n}$ be an arbitrary quaternion matrix, then the real representation of A is specified as follows:

$$A^R = \begin{pmatrix} A_1 & -A_2 & -A_3 & -A_4 \\ A_2 & A_1 & -A_4 & A_3 \\ A_3 & A_4 & A_1 & -A_2 \\ A_4 & -A_3 & A_2 & A_1 \end{pmatrix}. \tag{4.2}$$

From (4.1) and using the properties of the real representation operator (4.2), we get

$$\left(U^R\right)^T A^R V^R = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}^R,$$

which is the SVD decomposition of matrix A^R. Consequently, the singular values of the quaternion matrix A and its real representation form A^R are the same.

Definition 4.2. The *spectral condition number of a quaternion matrix* $A \in \mathbb{Q}^{m\times n}$ is defined by

$$\mathrm{cond}(A) = \|A\|_2 \left\|A^+\right\|_2 = \frac{\sigma_1}{\sigma_r},$$

where A^+, stands for the Moore-Penrose pseudoinverse of the quaternion matrix A.

By straightforward computations, it can be proved that $\mathrm{cond}(A) = \mathrm{cond}(A^R)$. Developing the idea utilized in [28], the preconditioned form of Algorithm 3.1 to solve (1.2) can be presented. Consider the ensuing four nonsingular preconditioners

$$Q_1 = diag\,(1/\left\|A(:,1)\right\|, 1/\left\|A(:,2)\right\|, \ldots, 1/\left\|A(:,n)\right\|)\,,$$

$$Q_2 = diag\,(1/\left\|B(1,:)\right\|, 1/\left\|B(2,:)\right\|, \ldots, 1/\left\|B(n,:)\right\|)\,,$$

$$Q_3 = diag\,(1/\left\|C(:,1)\right\|, 1/\left\|C(:,2)\right\|, \ldots, 1/\left\|C(:,n)\right\|)\,,$$

$$Q_4 = diag\,(1/\left\|D(1,:)\right\|, 1/\left\|D(2,:)\right\|, \ldots, 1/\left\|D(n,:)\right\|)\,,$$

where for arbitrary given matrix Z, $\|Z(:,i)\|$ and $\|Z(i,:)\|$ signify the Frobenius norm of i-th column and i-th row of matrix Z, respectively. As discussed in [28], the idea behind these choices of preconditionres comes from an established theorem in [34, Theorem 3.5]. As a matter of fact after applying the preconditioners of the above forms, the condition numbers of the coefficient matrices of the obtained preconditioned problem become smaller; for more details see [28, 34].

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

Consider the equivalent form of (1.3) as follows:

$$\min \left\| AQ_1Q_1^{-1}XQ_2^{-1}Q_2B + CQ_3Q_3^{-1}YQ_4^{-1}Q_4D - E \right\|. \tag{4.3}$$

Assume that

$$\hat{A} = AQ_1,\ \hat{B} = Q_2B,\ \hat{C} = CQ_3,\ \hat{D} = Q_4D,\ \hat{X} = Q_1^{-1}XQ_2^{-1},\text{ and } \hat{Y} = Q_3^{-1}YQ_4^{-1}.$$

Then (4.3) can be rewritten in the next form

$$\min \left\| \hat{A}\hat{X}\hat{B} + \hat{C}\hat{Y}\hat{D} - E \right\|. \tag{4.4}$$

By employing Algorithm 3.1 to solve (4.4), we can obtain $\hat{X}$ and $\hat{Y}$. Afterward, the solutions of (4.3) can be obtained by setting $X = Q_1\hat{X}Q_2$ and $Y = Q_3\hat{Y}Q_4$.

5. Numerical experiments. In this section, we examine some test examples to numerically confirm the validity of the results elaborated in the previous sections and to demonstrate that Algorithm 3.1 is robust and effective for solving Problems 1.2 and 1.3. We comment here that all of the numerical experiments were performed on a Pentium 4 PC with a 2.67 GHz CPU and 4.00GB of RAM using some Matlab code in MATLAB 8.1.0.604.

Example 5.1. Consider Problem 1.2 where

$$A = \begin{pmatrix} 1+4i+7j+2k & 1+2i-4j+2k \\ 3+i+4j+3k & 1+2i+3j+4k \end{pmatrix},$$
$$B = \begin{pmatrix} -6+i+5j+8k & 3+2i+j+4k \\ 7+9i+3j+2k & -2-3i+4j-5k \end{pmatrix},$$

$$C = \begin{pmatrix} -3+4i+j+5k & 0 \\ 5+i+3j+8k & 0 \end{pmatrix},\ D = \begin{pmatrix} -7+i+2j-k & -1+2i+9j-k \\ -3-3i+2j+k & 1+2i+3j+k \end{pmatrix},$$

and

$$E = \begin{pmatrix} -95.33-75.67i-110j-34.83k & -75-17.67i-72.33j-19.17k \\ -156.7-92.33i-90j-64.67k & -148.5-0.666i-30.17j+47.17k \end{pmatrix}.$$

We aim to apply Algorithm 3.1 to solve Problem 1.2. In addition, for two given matrices $\hat{X} \in i\mathbb{HQ}^{2\times 2}$ and $\hat{Y} \in i\mathbb{AQ}^{2\times 2}$,

$$\hat{X} = \begin{pmatrix} 3+5j+k & 2-i+2j+2k \\ 2+i+2j+2k & 1+3j+4k \end{pmatrix},$$

and

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

$$\hat{Y} = \begin{pmatrix} 4i & -2-0.5i-2j-0.5k \\ 2-0.5i+2j+0.5k & -2i \end{pmatrix},$$

the solution pair of Problem 1.3 is determined.

It can be verified that the matrix equation $AXB + CYD = E$ is consistent over the i-Hermitian and i-anti-Hermitian matrices as the following i-Hermitian and i-anti-Hermitian matrices satisfy $AXB + CYD = E$,

$$X = \begin{pmatrix} 1+j+k & 0.5+0.5j+0.5k \\ 0.5+0.5j+0.5k & 0.33+0.33j+0.33k \end{pmatrix},$$

and,

$$Y = \begin{pmatrix} i & 0.5i \\ 0.5i & 10i \end{pmatrix}.$$

In order to resolve Problem 1.2, we set $[X(0), Y(0)] = [0, 0]$ as the initial guess and exploit the subsequent stopping criterion,

$$\|R(k)\| = \|F - AX(k)B - CY(k)D\| < 10^{-10}.$$

Utilizing Algorithm 3.1, the approximate least-norm solution pair of mentioned problem can be derived respectively as follows:

$$X^{(21)} = \begin{pmatrix} 1+j+k & 0.5+0.5j+0.5k \\ 0.5+0.5j+0.5k & 0.33+0.33j+0.33k \end{pmatrix},$$

and

$$Y^{(21)} = \begin{pmatrix} i & 0.5i \\ 0.5i & 0 \end{pmatrix},$$

where the corresponding residual norm is $\|R(21)\| = \|F - AX(21)B - CY(21)D\| = 4.1507 \times 10^{-10}$. It can be seen that

$$\|X\| + \|Y\| = 10.3118 \quad \text{and} \quad \left\|X^{(21)}\right\| + \left\|Y^{(21)}\right\| = 2.5166.$$

Therefore,

$$\|X\| + \|Y\| > \left\|X^{(21)}\right\| + \left\|Y^{(21)}\right\|,$$

which confirms the established fact that performing Algorithm 3.1 with the initial approximate of the form (3.9) gives the least-norm solution pair of Problem 1.2. For

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

more details, the required iteration numbers and CPU time (sec) for the convergence and the Frobenius norm of residuals at different iterates are reported in Table 5.1.

TABLE 5.1
Numerical results for Example 5.1.

Iteration	15	20	25	30
CPU time	0.3327	0.4354	0.5618	0.6543
$\|R(k)\|$	0.0019	$1.0057e-10$	$2.1703e-12$	$3.4083e-13$

For solving Problem 1.3, Algorithm 3.1 is used to determine the solution pair of (3.15) in which

$$F = E - A\hat{X}B - C\hat{Y}D =$$

$$\begin{pmatrix} 143.2+203.3i-33.5j+143.2k & 147+35.83i+155.7j+185.3k \\ 374.8+278.7i+65.5j+163.3k & 146.5-56.83i+176.8j-158.3k \end{pmatrix}.$$

By the following choice of the initial guess $[Z(0), W(0)] = [0, 0]$, we derive

$$\|R(21)\| = \|F - AZ(21)B - CW(21)D\| = 5.0558 \times 10^{-11}.$$

The computed approximate solution pair for $[Z^*, W^*]$ is given by

$$Z^* \simeq Z^{(21)} = \begin{pmatrix} -2-4j & -1.5+i-1.5j-1.5k \\ -1.5-i-1.5j-1.5k & -0.6667-2.667j-3.667k \end{pmatrix},$$

and

$$W^* \simeq W^{(21)} = \begin{pmatrix} -3i & 2+i+2j+0.5k \\ -2+i-2j-0.5k & 0 \end{pmatrix}.$$

Using the presented discussion in Subsection 3.2, the solution of Problem 1.3 is calculated by

$$X^* = \hat{X} + Z^* \simeq \begin{pmatrix} 1+j+k & 0.5+0.5j+0.5k \\ 0.5+0.5j+0.5k & 0.33+0.33j+0.33k \end{pmatrix},$$

and,

$$Y^* = \hat{Y} + W^* \simeq \begin{pmatrix} i & 0.5i \\ 0.5i & -2i \end{pmatrix}.$$

For more clarification, we exhibit the convergence curves of the methods in Figures 5.1 where

$$\varepsilon(k) = \log_{10} \|R(k)\|. \tag{5.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

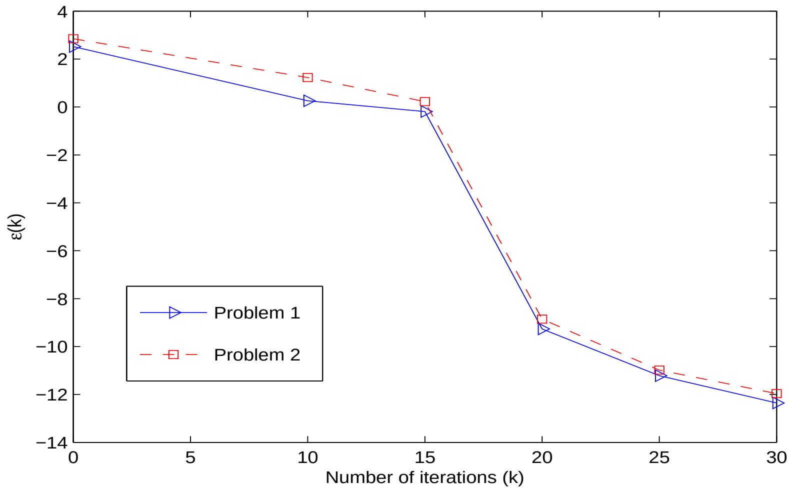

FIG. 5.1. *Convergence history for Problem 1.2.*

EXAMPLE 5.2. In this instance, we focus on Problem 1.2 such that

$$A = A_1 + A_2 i + A_3 j + A_4 k,\ B = B_1 + B_2 i + B_3 j + B_4 k,$$

$$C = C_1 + C_2 i + C_3 j + C_4 k,\ D = D_1 + D_2 i + D_3 j + D_4 k,$$

$$E = E_1 + E_2 i + E_3 j + E_4 k,$$

where

$$A_1 = \mathrm{triu}(\mathrm{hilb}(n)), \quad A_2 = \mathrm{triu}(\mathrm{ones}(n,n)), \ A_3 = \mathrm{eye}(n),\ A_4 = \mathrm{zeros}(n,n),$$

$$B_1 = \mathrm{tridiag}(n,-1,2,-1),\ B_2 = \mathrm{eye}(n),\ B_3 = \mathrm{zeros}(n,n),$$
$$B_4 = \mathrm{tridiag}(n,0.5,6,-0.5),$$

$$C_1 = C_2 = C_3 = C_4 = \mathrm{ones}(n,n),\ D_1 = D_2 = D_3 = D_4 = \mathrm{ones}(n,n),$$

$$E_1 = \mathrm{hankel}(1:n), E_2 = \mathrm{zeros}(n,n),\ E_3 = \mathrm{zeros}(n,n),\ E_4 = \mathrm{zeros}(n,n).$$

Our goal is to find the least-norm solution pair of Problem 1.2.

It can be seen that the matrix equation $AXB+CYD = E$ mentioned in Example 5.2 is not consistent over k-Hermitian and k-anti-Hermitian matrices. From Theorem 3.10, it is known that if the initial matrix group $[X(0), Y(0)]$ are chosen so that $[X(0), Y(0)] \in \mathbb{W}$, then the least-norm solution pair can be obtained. We apply

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

Algorithm 3.1 to solve Problem 1.2 for different values of n with $[X(0), Y(0)] = [0, 0]$. The iterations are stopped as soon as

$$\eta(k) = \log_{10}\sqrt{\frac{\|P_{k,x}\|^2 + \|P_{k,y}\|^2}{\|P_{0,x}\|^2 + \|P_{0,y}\|^2}} < 10^{-5}. \tag{5.2}$$

The details of numerical results including the required iteration numbers and the Frobenius norm of residuals are recorded in Table 5.2 where $R(k) = F - AX(k)B - CY(k)D$. The convergence histories are depicted in Figure 5.2 where $\eta(k)$ defined by (5.2).

TABLE 5.2
Numerical results for Example 5.2.

n	$\|X(k)\| + \|Y(k)\|$	$\min \|R(k)\|$	Iteration
20	13.3815	27.9922	76
40	38.8499	65.7652	178
60	107.3637	71.9070	287

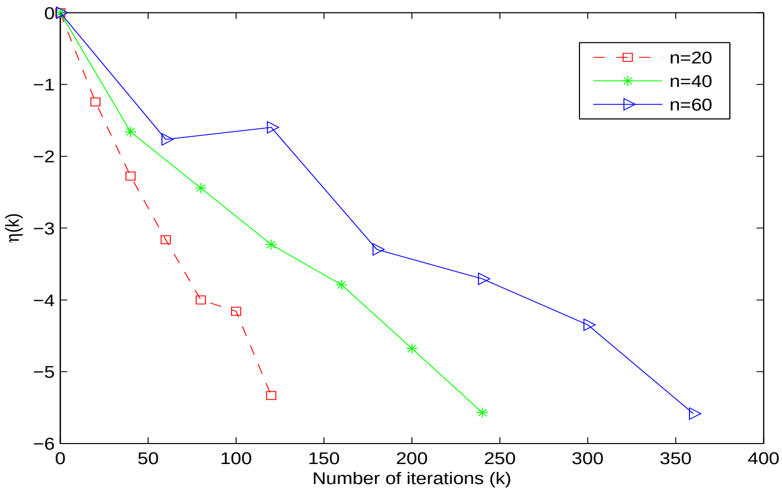

FIG. 5.2. *Convergence history for Example 5.2.*

REMARK 5.3. We would like to comment here that it is not possible to use the presented scheme in [40] for $n = 40, 60$ during the programming process due to

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

the required high computational costs. Whereas Algorithm 3.1 can be successfully handled which reveals the superiority of the proposed algorithm.

EXAMPLE 5.4. In this example, we examine the efficiency of Algorithm 3.1 for solving sparse quaternion least-squares problems. The test matrices were chosen from HB (Harwell-Boeing) group of Tim Davis's collection [10]. The four test matrices utilized in this experiment appear in structural problem ($bcsstm01$) and directed weighted graph problems(gre_115, gre_185 and gre_343). The sparsity structure of mentioned matrices are illustrated in Figure 5.3 where "nz" stands for the number of nonzero entries of the matrix. In addition, the properties of the test matrices are presented in Table 5.3 for more details.

TABLE 5.3
The sparsity structure of test matrices for Example 5.4.

Matrix	number of rows	number of columns	structure	condition number
$bcsstm01$	48	48	*symmetric*	Inf
gre_115	115	115	*unsymmetric*	49.6709
gre_185	185	185	*unsymmetric*	103.015
gre_343	343	343	*unsymmetric*	111.976

The main objective is to find the solution pair of Problem 1.2 associated with the matrix equation $AXB + CYD = E$ such that

$$A = A_1 + A_2 i + A_3 j + A_4 k,\ B = B_1 + B_2 i + B_3 j + B_4 k,$$

$$C = C_1 + C_2 i + C_3 j + C_4 k,\ D = D_1 + D_2 i + D_3 j + D_4 k,$$

$$E = E_1 + E_2 i + E_3 j + E_4 k.$$

Presume that

$$A_1 = B_1 = C_1 = D_1, \quad A_2 = B_2 = C_2 = D_2,$$

$$A_3 = B_3 = C_3 = D_3, \quad A_4 = B_4 = C_4 = D_4.$$

The right hand side E is constructed such that $X = \mathcal{L}_1^j(C)$ and $Y = \mathcal{L}_2^j(D)$ satisfy $AXB + CYD = E$. Hence, the problem is consistent over j-Hermitian and j-anti-Hermitian matrices. The following four instances are examined:

CASE 1:

$$A_1 = tridiag(-1, 2, -1),\ A_2 = bcsstm01,\ A_3 = tridiag(1, 3, 1),\ A_4 = bcsstm01.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

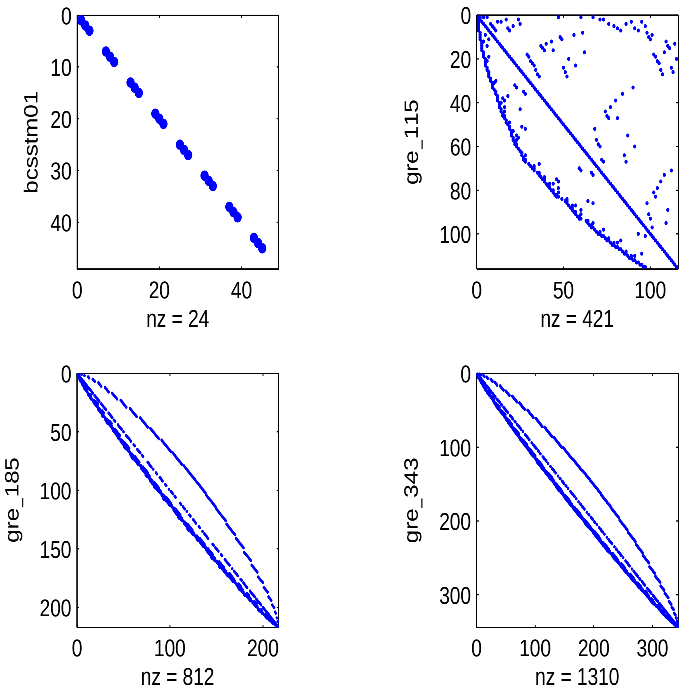

FIG. 5.3. *The sparsity structure of test matrices.*

CASE 2:

$$A_1 = tridiag(-1, 2, -1),\ A_2 = gre\_115,\ A_3 = tridiag(1, 3, 1),\ A_4 = gre\_115.$$

CASE 3:

$$A_1 = tridiag(-1, 2, -1),\ A_2 = gre\_185,\ A_3 = tridiag(1, 3, 1),\ A_4 = gre\_185.$$

CASE 4:

$$A_1 = tridiag(-1, 2, -1),\ A_2 = gre\_343,\ A_3 = tridiag(1, 3, 1),\ A_4 = gre\_343.$$

We employed Algorithm 3.1 and its preconditioned version for Example 5.4, the obtained results are presented in the following figures. For Case 1, Algorithm 3.1 has slow convergence rate than preconditioned version since the matrix $bcsstm01$ is an ill-conditioned matrix. The convergence histories of Algorithm 3.1 and its preconditioned version for this case are depicted in Figure 5.4 where $\varepsilon(k)$ specified by (5.1). However,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

for Cases 2–4, the convergence rate of Algorithm 3.1 and its preconditioned version is almost equal as the coefficients are well conditioned. The convergence history of Algorithm 3.1 for Cases 2–4 is plotted in Figure 5.5.

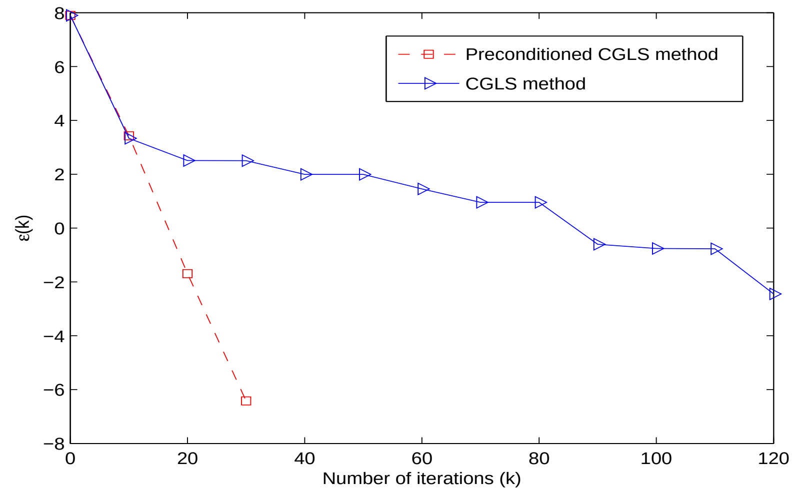

FIG. 5.4. *Convergence history for Case 1 of Example 5.4.*

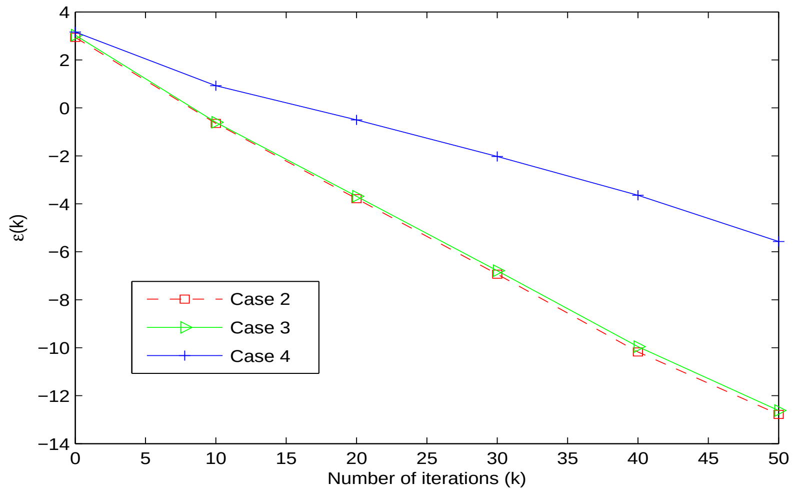

FIG. 5.5. *Convergence history for Cases 2, 3, 4 of Example 5.4.*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

6. Conclusion and further works. The paper has been concerned with solving two main problems referred Problem 1.2 and Problem 1.3. More precisely, Problem 1.2 deals with solving the least-squares problem

$$\|AXB + CYD - E\| = \min,$$

over $X \in \eta\mathbb{HQ}^{n\times n}$ and $Y \in \eta\mathbb{AQ}^{n\times n}$ matrices. The main aim of the second problem was to obtain the optimal approximate solution pair of Problem 1.2 with respect to an arbitrary given matrix pair. After establishing some properties of the η-Hermitian and η-anti-Hermitian matrices, the well-known CGLS iterative algorithm has been developed to construct an efficient and robust algorithm for solving Problems 1.2 and 1.3. The convergence properties of the algorithm have been analyzed and a preconditioned version of the algorithm has been suggested. Numerical results have illustrated the efficiency and feasibility of the presented algorithm. The proposed iterative scheme can be generalized for solving the following generalized quaternion least-squares problem over η-Hermitian and η-anti-Hermitian matrices,

$$\min \left\| \sum_{j=1}^{n} \sum_{i=1}^{n} \left(A_i X_i B_i + C_i X_i^H D_i - E_i \right) \right\|.$$

Future work can investigate applying suitable preconditioners to speed up the convergence of the CGLS algorithm for solving the least-squares problems corresponding to coupled matrix equations.

Acknowledgment. The authors wish to thank anonymous referees for their helpful comments.

REFERENCES

[1] F.A. Aliev and V.B. Larin. *Optimization of Linear Control Systems: Analytical Methods and Computational Algorithms*. Gordon and Breach Science Publishers, Amsterdam, 1998.

[2] U. Baur and P. Benner. Factorized solution of Lyapunov equations based on hierarchical matrix arithmetic. *Computing*, 78:211–234, 2006.

[3] F.P.A. Beik and D.K. Salkuyeh. On the global Krylov subspace methods for solving general coupled matrix equations. *Computers and Mathematics with Applications*, 62:4605–4613, 2011.

[4] F.P.A. Beik and D.K. Salkuyeh. The coupled Sylvester-transpose matrix equations over generalized centrosymmetric matrices. *International Journal of Computer Mathematics*, 90:1546–1566, 2013.

[5] N.L. Bihan and J. Mars. Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing. *Signal Processing*, 84:1177–1199, 2004.

[6] N.L. Bihan and S.J. Sangwine. Color image decomposition using quaternion singular value decomposition. IEEE International Conference on Visual Information Engineering of quaternion (VIE), Guildford, UK, 2003.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

[7] A. Björck. *Numerical Methods for Least Squares Problems.* SIAM, Philadelphia, 1996.

[8] J. Cai and G. Chen. An iterative algorithm for the least squares bisymmetric solutions of the matrix equations $A_1XB_1 = C_1, A_2XB_2 = C_2$. *Mathematical and Computer Modelling*, 50:1237–1244, 2009.

[9] E.B. Castelan and V.G. da Silva. On the solution of a Sylvester equation appearing in descriptor systems control theory. *Systems & Control Letters*, 54:109–117, 2005.

[10] T. A. Davis and Y. Hu. The University of Florida Sparse Matrix Collection. *ACM Transactions on Mathematical Software* 38:1–25, 2011. Available at `http://www.cise.ufl.edu/research/sparse/matrices`.

[11] M. Dehghan and M. Hajarian. An efficient algorithm for solving general coupled matrix equations and its application. *Mathematical and Computer Modelling*, 51:1118–1134, 2010.

[12] M. Dehghan and M. Hajarian. The general coupled matrix equations over generalized bisymmetric matrices. *Linear Algebra and its Applications*, 432:1531–1552, 2010.

[13] M. Dehghan and M. Hajarian. Analysis of an iterative algorithm to solve the generalized coupled Sylvester matrix equations. *Applied Mathematical Modelling*, 35:3285–3300, 2011.

[14] M. Dehghan and M. Hajarian. On the generalized reflexive and anti-reflexive solutions to a system of matrix equations. *Linear Algebra and its Applications*, 437:2793–2812, 2012.

[15] F. Ding and T. Chen. Gradient based iterative algortithms for solving a class of matrix equations. *IEEE Transactions Automatic Control*, 50:1216–1221, 2005.

[16] F. Ding and T. Chen. On iterative solutions of general coupled matrix equations. *SIAM Journal on Control and Optimization*, 44:2269–2284, 2006.

[17] F. Ding, P.X. Liu, and J. Ding. Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle. *Applied Mathematics and Computation*, 197:41–50, 2008.

[18] Z. Gajic and M.T.J. Qureshi. *Lyapunov Matrix Equation in System Stability and Control.* Dover Publications, Inc., Mineola, NY, 2008. (Original work published in 1995).

[19] Z.H. He and Q.W. Wang. A real quaternion matrix equation with applications. *Linear and Multilinear Algebra*, 61:725–740, 2013.

[20] Z.H. He and Q.W. Wang. The η-bihermitian solution to a system of real quaternion matrix equations. *Linear and Multilinear Algebra*, 62:1509–1528, 2014.

[21] L.P. Huang. The matrix equation $AXB - GXD = E$ over the quaternion field. *Linear Algebra and its Applications*, 234:197–208, 1996.

[22] J. Jiang and N. Li. An efficient algorithm for the generalized (P,Q)-reflexive solution to a quaternion matrix equation and its optimal approximation. *Journal of Applied Mathematics and Computing*, 45:297–326, 2014.

[23] T.S. Jiang and M.S. Wei. On a solution of the quaternion matrix equation $X - A\tilde{X}B = C$ and its application. *Acta Mathematica Sinica (English Series)*, 21:483–490, 2005.

[24] C.T. Kelley. *Iterative Methods for Linear and Nonlinear Equations.* SIAM, Philadelphia, PA, 1995.

[25] S.D. Leo and G. Scolarici. Right eigenvalue equation in quaternionic quantum mechanics. *Journal of Physics A: Mathematical and General*, 33:2971–2995, 2000.

[26] J.F. Li, X.Y. Hu, and L. Zhang. Numerical solutions of $AXB = C$ for centrosymmetric matrix X under a specified submatrix constraint. *Numerical Linear Algebra with Applications*, 18:857–873, 2011.

[27] J.F. Li, W. Li, and Z.Y. Peng. A hybrid algorithm for solving minimization problem over (R,S)-symmetric matrices with the matrix inequality constraint. *Linear Multilinear Algebra*, 63:1049-1072, 2015.

[28] S. Ling, M. Wang, and M. Wei. Hermitian tridiagonal solution with the least norm to quaternionic least squares problem. *Computer Physics Communications*, 181:481–488, 2010.

[29] Z. Peng. The reflexive least squares solutions of the matrix equation $A_1X_1B_1 + A_2X_2B_2 + \cdots + A_lXlB_l = C$ with a submatrix constraint. *Numerical Algorithms*, 64:455–480, 2013.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

[30] Z. Peng. The (R, S)-symmetric least squares solutions of the general coupled matrix equations. *Linear and Multilinear Algebra*, 63:1086-1105, 2015.
[31] Z. Peng and H. Xin. The reflexive least squares solutions of the general coupled matrix equations with a submatrix constraint. *Applied Mathematics and Computation*, 225:425–445, 2013.
[32] Y. Saad. *Iterative Methods for Sparse Linear Systems*, second edition. SIAM, 2003.
[33] S.J. Sangwine, T.A. Ell, and C.E. Moxey. Vector phase correlation. *Electronics Letters*, 37:1513–1515, 2001.
[34] A. van der Sluis. Condition numbers and equilibration of matrices. *Numerische Mathematik*, 14:14–23, 1969.
[35] C. Song, G. Chen, and X. Zhang. An iterative solution to coupled quaternion matrix equations. *Filomat*, 26:809–826, 2012.
[36] E. Stiefel. Ausgleichung ohne aufstellung der Gausschen normalgleichungen. *Wissenschaftliche Zeitschrift der Technischen Hochschule Dresden*, 2:441–442, 1953.
[37] R.S. Wang. *Functional Analysis and Optimization Theory.* Beijing University of Aeronautics and Astronautics Press, Beijing 2003.
[38] N.A. Wiegmann. Some theorems on matrices with real quaternion elements. *Canadian Journal of Mathematics*, 7:191–201, 1955.
[39] S.F. Yuan and Q.W. Wang. Two special kinds of least squares solutions for the quaternions matrix equation $AXB + CXD = E$. *Electronic Journal of Linear Algebra*, 23:257–274, 2012.
[40] S.F. Yuan, Q.W. Wang, and X. Zhang. Least-squares problem for the quaternion matrix equation $AXB + CYD = E$ over different constrained matrices. *International Journal of Computer Mathematics*, 90:565–576, 2013.
[41] F. Zhang. Quaternions and matrices of quaternions. *Linear Algebra and its Applications*, 251:21–57, 1997.
[42] G. Zhang, Q. Zhang, T. Chen, and Y. Lin. On Lyapunov theorems for descriptor systems, *Dynamics of Continuous, Discrete & Impulsive Systems. Series B. Applications & Algorithms*, 10:709–725, 2003.
[43] B. Zhou and G.R. Duan. On the generalized Sylvester mapping and matrix equations. *Systems & Control Letters*, 57:200–208, 2008.
[44] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control.* Prentice Hall, New Jersey, 1996.

Appendix A. Proof of Theorem 3.6. Because of commutative property of the inner product $\langle ., . \rangle$, we only need to prove the validity of Eqs. (3.6), (3.7) and (3.8) for $1 \le i < j \le k$. For $k = 1$, we have

$$
\begin{aligned}
&\langle P_{0,x}, P_{1,x} \rangle + \langle P_{0,y}, P_{1,y} \rangle \\
&\quad = \left\langle P_{0,x}, P_{0,x} - \alpha_0 \mathcal{L}_1^\eta \left(A^H \left(\mathcal{M}(Q_{0,x}, Q_{0,y}) \right) B^H \right) \right\rangle \\
&\qquad + \left\langle P_{0,y}, P_{0,y} - \alpha_0 \mathcal{L}_2^\eta \left(C^H \left(\mathcal{M}(Q_{0,x}, Q_{0,y}) \right) D^H \right) \right\rangle \\
&\quad = \|P_{0,x}\|^2 + \|P_{0,y}\|^2 - \alpha_0 \langle AP_{0,x}B, \mathcal{M}(Q_{0,x}, Q_{0,y}) \rangle - \alpha_0 \langle CP_{0,y}D, \mathcal{M}(Q_{0,x}, Q_{0,y}) \rangle \\
&\quad = \|P_{0,x}\|^2 + \|P_{0,y}\|^2 - \alpha_0 \langle \mathcal{M}(Q_{0,x}, Q_{0,y}), \mathcal{M}(Q_{0,x}, Q_{0,y}) \rangle \\
&\quad = 0.
\end{aligned}
$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

$$
\begin{aligned}
&\langle \mathcal{M}\left(Q_{0,x}, Q_{0,y}\right), \mathcal{M}\left(Q_{1,x}, Q_{1,y}\right)\rangle \\
&\quad = \langle \mathcal{M}\left(Q_{0,x}, Q_{0,y}\right), \mathcal{M}\left(P_{1,x} + \beta_0 Q_{0,x}, P_{1,y} + \beta_0 Q_{0,y}\right)\rangle \\
&\quad = \beta_0 \|\mathcal{M}\left(Q_{0,x}, Q_{0,y}\right)\|^2 + \langle \mathcal{M}\left(Q_{0,x}, Q_{0,y}\right), \mathcal{M}\left(P_{1,x}, P_{1,y}\right)\rangle \\
&\quad = \beta_0 \|\mathcal{M}\left(Q_{0,x}, Q_{0,y}\right)\|^2 \\
&\qquad + \frac{1}{\alpha_0}\left(\langle \mathcal{M}_1^*\left(R_0 - R_1\right), P_{1,x}\rangle + \langle \mathcal{M}_2^*\left(R_0 - R_1\right), P_{1,y}\rangle\right) \\
&\quad = \beta_0 \|\mathcal{M}\left(Q_{0,x}, Q_{0,y}\right)\|^2 \\
&\qquad + \frac{1}{\alpha_0}\left(\langle \mathcal{L}_1^\eta\left(\mathcal{M}_1^*\left(R_0 - R_1\right)\right), P_{1,x}\rangle + \langle \mathcal{L}_2^\eta\left(\mathcal{M}_2^*\left(R_0 - R_1\right)\right), P_{1,y}\rangle\right) \\
&\quad = \beta_0 \|\mathcal{M}\left(Q_{0,x}, Q_{0,y}\right)\|^2 \\
&\qquad + \frac{1}{\alpha_0}\left(\langle P_{0,x} - P_{1,x}, P_{1,x}\rangle + \langle P_{0,y} - P_{1,y}, P_{1,y}\rangle\right) \\
&\quad = 0.
\end{aligned}
$$

Evidently, it can be seen that

$$
\langle P_{1,x}, Q_{0,x}\rangle + \langle P_{1,y}, Q_{0,y}\rangle = \langle P_{1,x}, P_{0,x}\rangle + \langle P_{1,y}, P_{0,y}\rangle = 0.
$$

Hence, the equations (3.6)–(3.8) are valid for $k = 1$. Suppose that assertions (3.6)–(3.8) are true for $i \leq s$. To complete the proof, in the sequel, we demonstrate the validity of the conclusions for $k = s + 1$. For $i < s$, it can be deuced that

$$
\begin{aligned}
&\langle P_{i,x}, P_{s+1,x}\rangle + \langle P_{i,y}, P_{s+1,y}\rangle \\
&\quad = \left\langle P_{i,x}, P_{s,x} - \alpha_s \mathcal{L}_1^\eta\left(A^H\left(\mathcal{M}(Q_{s,x}, Q_{s,y})\right) B^H\right)\right\rangle \\
&\qquad + \left\langle P_{i,y}, P_{s,y} - \alpha_s \mathcal{L}_2^\eta\left(C^H\left(\mathcal{M}(Q_{s,x}, Q_{s,y})\right) D^H\right)\right\rangle \\
&\quad = -\alpha_s \langle \mathcal{M}_1(P_{i,x}), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle \\
&\qquad - \alpha_s \langle \mathcal{M}_2(P_{i,y}), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle \\
&\quad = -\alpha_s \langle \mathcal{M}_1(Q_{i,x} - \beta_{i-1} Q_{i-1,x}), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle \\
&\qquad - \alpha_s \langle \mathcal{M}_2(Q_{i,y} - \beta_{i-1} Q_{i-1,y}), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle \\
&\quad = 0,
\end{aligned}
$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

ELA

$$
\begin{aligned}
&\langle \mathcal{M}\left(Q_{i,x}, Q_{i,y}\right), \mathcal{M}\left(Q_{s+1,x}, Q_{s+1,y}\right)\rangle \\
&\quad = \langle \mathcal{M}\left(Q_{i,x}, Q_{i,y}\right),\ \mathcal{M}\left(P_{s+1,x} + \beta_s Q_{s,x}, P_{s+1,y} + \beta_s Q_{s,y}\right)\rangle \\
&\quad = \langle \mathcal{M}\left(Q_{i,x}, Q_{i,y}\right), \mathcal{M}\left(P_{s+1,x}, P_{s+1,y}\right)\rangle \\
&\quad = \frac{1}{\alpha_i}\langle R(i) - R(i+1), \mathcal{M}\left(P_{s+1,x}, P_{s+1,y}\right)\rangle \\
&\quad = \frac{1}{\alpha_i}\langle \mathcal{M}_1^*\left(R(i) - R(i+1)\right), P_{s+1,x}\rangle + \frac{1}{\alpha_i}\langle \mathcal{M}_2^*\left(R(i) - R(i+1)\right), P_{s+1,y}\rangle \\
&\quad = \frac{1}{\alpha_i}\langle \mathcal{M}_1^*\left(R(i)\right) - \mathcal{M}_1^*(R(i+1)), P_{s+1,x}\rangle \\
&\qquad + \frac{1}{\alpha_i}\langle \mathcal{M}_2^*\left(R(i)\right) - \mathcal{M}_2^*\left(R(i+1)\right), P_{s+1,y}\rangle \\
&\quad = \frac{1}{\alpha_i}\langle \mathcal{L}_1^\eta\left(\mathcal{M}_1^*\left(R(i)\right) - \mathcal{M}_1^*\left(R(i+1)\right)\right), P_{s+1,x}\rangle \\
&\qquad + \frac{1}{\alpha_i}\langle \mathcal{L}_2^\eta\left(\mathcal{M}_2^*\left(R(i)\right) - \mathcal{M}_2^*\left(R(i+1)\right)\right), P_{s+1,y}\rangle \\
&\quad = -\frac{1}{\alpha_i}\left(\langle P_{i+1,x}, P_{s+1,x}\rangle + \langle P_{i+1,y}, P_{s+1,y}\rangle\right) = 0,
\end{aligned}
$$

and,

$$
\begin{aligned}
&\langle Q_{i,x}, P_{s+1,x}\rangle + \langle Q_{i,y}, P_{s+1,y}\rangle \\
&\quad = \left\langle Q_{i,x}, P_{s,x} - \alpha_s \mathcal{L}_1^\eta\left(A^H\left(\mathcal{M}(Q_{s,x}, Q_{s,y})\right)B^H\right)\right\rangle \\
&\qquad + \left\langle Q_{i,x}, P_{s,x} - \alpha_s \mathcal{L}_2^\eta\left(C^H\left(\mathcal{M}(Q_{s,x}, Q_{s,y})\right)D^H\right)\right\rangle \\
&\quad = -\alpha_s\langle \mathcal{M}_1\left(Q_{i,x}\right), \mathcal{M}\left(Q_{s,x}, Q_{s,y}\right)\rangle - \alpha_s\langle \mathcal{M}_2\left(Q_{i,y}\right), \mathcal{M}\left(Q_{s,x}, Q_{s,y}\right)\rangle \\
&\quad = -\alpha_s\langle \mathcal{M}\left(Q_{i,x}, Q_{i,y}\right), \mathcal{M}\left(Q_{s,x}, Q_{s,y}\right)\rangle = 0.
\end{aligned}
$$

For $i = s$, we have

$$
\begin{aligned}
&\langle P_{s,x}, P_{s+1,x}\rangle + \langle P_{s,y}, P_{s+1,y}\rangle \\
&\quad = \left\langle P_{s,x}, P_{s,x} - \alpha_s \mathcal{L}_1^\eta\left(A^H\left(\mathcal{M}(Q_{s,x}, Q_{s,y})\right)B^H\right)\right\rangle \\
&\qquad + \left\langle P_{s,y}, P_{s,y} - \alpha_s \mathcal{L}_2^\eta\left(C^H\left(\mathcal{M}(Q_{s,x}, Q_{s,y})\right)D^H\right)\right\rangle \\
&\quad = \|P_{s,x}\|^2 + \|P_{s,y}\|^2 - \alpha_s\langle \mathcal{M}_1\left(P_{s,x}\right), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle \\
&\qquad - \alpha_s\langle \mathcal{M}_2\left(P_{s,y}\right), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle \\
&\quad = \|P_{s,x}\|^2 + \|P_{s,y}\|^2 - \alpha_s\langle \mathcal{M}\left(P_{s,x}, P_{s,y}\right), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle \\
&\quad = \|P_{s,x}\|^2 + \|P_{s,y}\|^2 - \alpha_s\langle \mathcal{M}(Q_{s,x}, Q_{s,y}), \mathcal{M}(Q_{s,x}, Q_{s,y})\rangle = 0,
\end{aligned}
$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 372-401, July 2015

$$
\begin{aligned}
&\langle \mathcal{M}\left(Q_{s,x},Q_{s,y}\right),\mathcal{M}\left(Q_{s+1,x},Q_{s+1,y}\right)\rangle\\
&\quad=\langle \mathcal{M}\left(Q_{s,x},Q_{s,y}\right),\mathcal{M}\left(P_{s+1,x}+\beta_s Q_{s,x},P_{s+1,y}+\beta_s Q_{s,y}\right)\rangle\\
&\quad=\beta_s\|\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)\|^2+\langle \mathcal{M}\left(Q_{s,x},Q_{s,y}\right),\mathcal{M}\left(P_{s+1,x},P_{s+1,y}\right)\rangle\\
&\quad=\beta_s\|\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)\|^2+\frac{1}{\alpha_s}\langle R(s)-R(s+1),\mathcal{M}\left(P_{s+1,x},P_{s+1,y}\right)\rangle\\
&\quad=\beta_s\|\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)\|^2+\frac{1}{\alpha_s}\langle R(s)-R(s+1),\mathcal{M}\left(P_{s+1,x},P_{s+1,y}\right)\rangle\\
&\quad=\beta_s\|\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)\|^2+\frac{1}{\alpha_s}\langle \mathcal{M}_1^*\left(R(s)-R(s+1)\right),P_{s+1,x}\rangle\\
&\qquad+\frac{1}{\alpha_s}\langle \mathcal{M}_2^*\left(R(s)-R(s+1)\right),P_{s+1,y}\rangle\\
&\quad=\beta_s\|\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)\|^2+\frac{1}{\alpha_s}\langle \mathcal{L}_1^\eta\left(\mathcal{M}_1^*\left(R(s)-R(s+1)\right)\right),P_{s+1,x}\rangle\\
&\qquad+\frac{1}{\alpha_s}\langle \mathcal{L}_2^\eta\left(\mathcal{M}_2^*\left(R(s)-R(s+1)\right)\right),P_{s+1,y}\rangle\\
&\quad=\beta_s\|\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)\|^2+\frac{1}{\alpha_s}\langle P_{s,x}-P_{s+1,x},P_{s+1,x}\rangle\\
&\qquad+\frac{1}{\alpha_s}\langle P_{s,y}-P_{s+1,y},P_{s+1,y}\rangle=0,
\end{aligned}
$$

and,

$$
\begin{aligned}
&\langle Q_{s,x},P_{s+1,x}\rangle+\langle Q_{s,y},P_{s+1,y}\rangle\\
&\quad=\left\langle Q_{s,x},P_{s,x}-\alpha_s\mathcal{L}_1^\eta\left(A^H\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)B^H\right)\right\rangle\\
&\qquad+\left\langle Q_{s,y},P_{s,y}-\alpha_s\mathcal{L}_2^\eta\left(C^H\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)D^H\right)\right\rangle\\
&\quad=\|P_{s,x}\|^2+\|P_{s,y}\|^2-\alpha_s\left\langle Q_{s,x},A^H\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)B^H\right\rangle\\
&\qquad-\alpha_s\left\langle Q_{s,y},C^H\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)D^H\right\rangle\\
&\quad=\|P_{s,x}\|^2+\|P_{s,y}\|^2-\alpha_s\|\mathcal{M}\left(Q_{s,x},Q_{s,y}\right)\|^2=0.
\end{aligned}
$$

Now the result can be concluded from the principle of mathematical induction.