Sylvester-Type Matrix Equations over Generalized Quaternions with Applications to Color Image Processing

Janthip Jaiprasert and Pattrawut Chansangiam

Abstract—**This paper focuses on solving a generalized Sylvester matrix equation over a generalized quaternion skew-field. We apply a real representation of generalized quaternion matrices and certain vectorizations to transform the matrix equation into a real linear system. Then, we obtain an equivalent condition for the consistency of the matrix equation. We can derive vector representations of the (minimal-norm) least-squares (LS) solution, the (minimal-norm) pure-imaginary LS solution, the (minimal-norm) real LS solution, and the (minimal-norm) LS solution closet to a given matrix. Such solutions are expressed in terms of Kronecker products and Moore-Penrose inverses. When the matrix equation is consistent, such LS solutions become exact solutions. This work includes Sylvester and Stein matrix equations over generalized quaternions, and quaternionic matrix equations. We also propose a gradient-descent iterative (GDI) algorithm to solve the transformed linear system. Moreover, the theory can be applied to a color image processing model.**

Index Terms—**Sylvester-type matrix equation, matrix over a generalized quaternion, least-squares solution, Kronecker product, iterative algorithm, RGB color model**

I. Introduction

THE paper focuses on matrix equations over quaternion-like structures. Let us recall that for any pair (u, v) of nonzero real numbers, we can associate with a four-dimensional algebra $\mathbb{Q}_{u,v}$ over the real number field $\mathbb{R}$. The algebra $\mathbb{Q}_{u,v}$ of generalized quaternions is formed by its ordered basis $\{1, i, j, k\}$ where 1 acts as the multiplicative identity. Besides the addition and the scalar multiplication on $\mathbb{Q}_{u,v}$, the product of any two of i, j, k is defined by the following multiplication rules:

$$i^2 = u,\ j^2 = v,\ k^2 = ijk = -uv,$$
$$ij = -ji = k,\ jk = -kj = -vi,\ ik = -ki = uj.$$

Thus, every element of $\mathbb{Q}_{u,v}$ can be represented as

$$q = q_1 + q_2 i + q_3 j + q_4 k \in \mathbb{Q}_{u,v}$$

where $q_1, q_2, q_3, q_4 \in \mathbb{R}$. We call q_1 the real part of q, while the vector (q_2, q_3, q_4) is called the imaginary part of q. The set $\mathbb{Q}_{u,v}$, together with the addition and the above multiplication, forms a skew field. The famous particular case of $\mathbb{Q}_{u,v}$ is when $(u, v) = (-1, -1)$, known as the (Hamilton) quaternions $\mathbb{Q}$. Other interesting cases of $\mathbb{Q}_{u,v}$ are the split quaternion ring $(u = -1, v = 1)$, the nectarine quaternion ring $(u = 1, v = -1)$, and the conectarine quaternion ring $(u = v = 1)$.

Quaternionic matrix theory is an attractive area in linear algebra; see e.g. [1]. Matrix equations over quaternion-like structures play an important role in computer platform [2], signal processing [3], [4], quantum mechanics [5], [6], and image processing [7], [8], [9]. In color image processing, a color image can be represented as a vector or matrix. According to RGB color model, the color information of a pixel can be represented as

$$q = 0 + q_r i + q_g j + q_b k \in \mathbb{Q},$$

where q_r, q_g, q_b are the red/green/blue component of the color pixel, respectively. The addition on $\mathbb{Q}$ represents the addition between different colors, so that their light spectra adds up. An RGB color image consists of many array pixels, and can be represented as a vector

$$\dot{y} = 0 + y_r i + y_g j + y_b k \in \mathbb{Q}^a,$$

where $y_r, y_g, y_b \in \mathbb{R}^a$. A quaternion-based sparse representation model [10] says that $\dot{y} = \dot{D}\dot{a}$, where

$$\dot{D} = D_s + D_r i + D_g j + D_b k \in \mathbb{Q}^{a\times b} \text{ and}$$
$$\dot{a} = a_1 + a_2 i + a_2 j + a_4 k \in \mathbb{Q}^b.$$

The matrix $\dot{D}$ and the vector $\dot{a}$ are called a dictionary matrix and a sparse coefficient vector.

Linear matrix equations over the field $\mathbb{R}$ arise naturally in pure and applied mathematics, e.g. differential equations, and mathematical control theory. The famous Sylvester matrix equation:

$$AX + XD = E \tag{1}$$

plays an important role in model reduction [11], numerical methods for differential equations [12], [13] and control systems [14], [15]. Besides, there are many researchers studied a generalized Sylvester equation:

$$AXB + CXD = E \tag{2}$$

see e.g. [16], [17], [18], [19], [20]. In particular, if C and D are identity matrices with suitable size, Eq. (2) is reduced to the Stein matrix equation. In the last decent, many author investigated Sylvester-type matrix equation over $\mathbb{Q}$ or $\mathbb{Q}_{u,v}$. In 2014, Shi-Fang Yuan [21] derived explicit forms of the least-squares (LS) solution, the imaginary LS solution, and the real solution of Eq. (2). Later, F. Zhang et al. [22] solved Eq. (2) over $\mathbb{Q}$ for the minimal-norm LS solution, the

Manuscript received December 31, 2024; revised April 7, 2025. This work was supported by King Mongkut's Institute of Technology Ladkrabang.

J. Jaiprasert is a PhD candidate of Mathematics Department, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand (e-mail: 65056013@kmitl.ac.th).

P. Chansangiam is a professor of Mathematics Department, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand (corresponding author to provide phone: 662-329-4800; fax; 662-329-4800; e-mail: pattrawut.ch@kmitl.ac.th).

imaginary LS solution, and the real LS solution. Furthermore, Tian et al. [23] investigated Eq. (2) over $\mathbb{Q}_{u,v}$. Indeed, they provided criterion for an existence of a Hermitian solution, and derived an explicit formula of the solution. Recently, the exact and least-squares solutions of a generalized Sylvester-transpose matrix equation over $\mathbb{Q}_{u,v}$ were studied in [24].

This paper is a continuation of the work [22]. We consider the generalized Sylvester matrix equation (2) where the given coefficients and the unknown are compatible rectangular matrices over $\mathbb{Q}_{u,v}$. We will discuss the following general/specific types of solutions.

Problem 1. *Find the general exact/least-squares solutions* $X = X_1 + X_2 i + X_3 j + X_4 k \in \mathbb{Q}_{u,v}^{a\times b}$ *of Eq.* (2)*. In addition, among the general solutions, find the minimal-norm one.*

Problem 2. *Find all exact/least-squares solutions of Eq.* (2) *that consist only of the imaginary part, i.e.,* $X = X_2 i + X_3 j + X_4 k \in \mathbb{Q}_{u,v}^{a\times b}$*. Among such solutions, find the minimal-norm one.*

Problem 3. *Find all exact/least-squares solutions of Eq.* (2) *that consist only of the real part. Among such solutions, find the minimal-norm one.*

Problem 4. *Let* $Y \in \mathbb{Q}_{u,v}^{a\times b}$ *be given.*

1.4.1 *Find the general exact/least-squares solution of Eq.* (2) *closet to* Y.
1.4.2 *Find the imaginary-part exact/least-squares solution of Eq.* (2) *closet to* Y.
1.4.3 *Find the real exact/least-squares solution of Eq.* (2) *closet to* Y.

Problems 1-4 include both consistent and inconsistent cases. When the associated least-squares error is zero, such least-squares solutions become an exact solution. We use real representations of generalized quaternion matrices and vectorizations of real matrices to transform Eq. (2) into a real linear system. So, we can derive the desired solutions of Problems 1-4 in terms of the Kronecker product and Moore-Penrose inverses; see Sections III and IV. We discuss certain special cases of Eq. (2), namely, Eq. (1), the Stein equation $AXB + X = E$, and the split quaternions case; see Section V. In Section VI, we provide numerical examples to illustrate the theory. In Section VII, we propose an iterative algorithm to solve the linear system associated with Eq. (2). In Section VIII, we apply the theory to a color image processing model. Finally, we summarize the whole work in the last section.

Next, we prepare basic notations and recall prerequisite results from classical and quaternionic matrix theory in Section II. These results involve real linear systems, vectorizations, the Kronecker product, and real representations for the generalized quaternion matrices.

II. Preliminaries from classical and quaternionic matrix theory

Throughout this paper, let $u, v \in \mathbb{R} - \{0\}$. Denote the set of all m-by-n real matrices and generalized quaternion matrices by $\mathbb{R}^{m\times n}$ and $\mathbb{Q}_{u,v}^{m\times n}$, respectively. For any matrix A, its transpose, its Moore-Penrose inverse, and its Frobenius norm are denoted by $A^T, A^\dagger$ and $\|A\|$, respectively. Let us denote the i-th column of matrix A by $\mathrm{col}_i(A)$. The identity matrix of order n is denoted by I_n, and we define $e_i^n = \mathrm{col}_i(A)$. Recall the following results.

Lemma 5. *(e.g. [25]) Let* $A \in \mathbb{R}^{m\times n}$ *and* $b \in \mathbb{R}^m$*. Then the linear system*

$$Ax = b, \tag{3}$$

has a solution $x \in \mathbb{R}^n$ *if and only if* $AA^\dagger b = b$*, or equivalently,* $\mathrm{rank}[A\ b] = \mathrm{rank}\ A$*. In both consistent and inconsistent cases, the following statements hold:*

(i) *The general exact/LS solutions of Eq.* (3) *can be expressed by the formula*

$$x = A^\dagger b + (I_n - A^\dagger A)w, \tag{4}$$

where $w \in \mathbb{R}^n$ *is arbitrary.*

(ii) *The minimal-norm exact/LS solution of Eq.* (4) *is given by the formula*

$$x = A^\dagger b. \tag{5}$$

(iii) *Eq.* (3) *has a unique exact/LS solution given by the formula* (5) *if* A *is of full-column rank (i. e.* $\mathrm{rank}[A] = n$*).*

Recall that the operator $\mathrm{V_c}(\cdot)$ transforms any matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ to be a column vector

$$\mathrm{V_c}(A) = \begin{pmatrix}\mathrm{col}_1(A) & \mathrm{col}_2(A) & \dots & \mathrm{col}_n(A)\end{pmatrix}^T \in \mathbb{R}^{mn}.$$

The Kronecker product of $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{s\times t}$ is defined as

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1j}B \\ a_{21}B & a_{22}B & \dots & a_{2j}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix} \in \mathbb{R}^{ms\times nt}.$$

Lemma 6. *(e.g. [26]) For any* $A \in \mathbb{R}^{m\times n}$*,* $X \in \mathbb{R}^{n\times p}$ *and* $B \in \mathbb{R}^{p\times q}$*, we have*

$$\mathrm{V_c}(AXB) = (B^T \otimes A)\,\mathrm{V_c}(X).$$

For any generalized quaternion matrix $A \in \mathbb{Q}_{u,v}^{m\times n}$, we can write

$$A = A_1 + A_2 i + A_3 j + A_4 k$$

with real coefficients $A_1, A_2, A_3, A_4 \in \mathbb{R}^{m\times n}$. We denote the column block of real coefficients by

$$\Gamma_1(A) = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} \in \mathbb{R}^{4m\times n}.$$

Now, consider $X = X_1 + X_2 i + X_3 j + X_4 k \in \mathbb{Q}_{u,v}^{n\times p}$, where $X_1, X_2, X_3, X_4 \in \mathbb{R}^{n\times p}$. A direct computation reveals that

$$\begin{aligned}\Gamma_1(AX) &= \begin{pmatrix} A_1X_1 + uA_2X_2 + vA_3X_3 - uvA_4X_4 \\ A_1X_2 + A_2X_1 - vA_3X_4 + vA_4X_3 \\ A_1X_3 + uA_2X_4 + A_3X_1 - uA_4X_2 \\ A_1X_4 + A_2X_3 - A_3X_2 + A_4X_1 \end{pmatrix} \\ &= \mathcal{R}(A)\Gamma_1(X), \end{aligned} \tag{6}$$

where

$$\mathcal{R}(A) = \begin{pmatrix} A_1 & uA_2 & vA_3 & -uvA_4 \\ A_2 & A_1 & vA_4 & -vA_3 \\ A_3 & -uA_4 & A_1 & uA_2 \\ A_4 & -A_3 & A_2 & A_1 \end{pmatrix}.$$

We call $\mathcal{R}(A)$ a real-matrix representation of A. We define the following representations of A :

$$\Gamma_2(A) = \begin{pmatrix} uA_2 \\ A_1 \\ -uA_4 \\ -A_3 \end{pmatrix},$$

$$\Gamma_3(A) = \begin{pmatrix} vA_3 \\ vA_4 \\ A_1 \\ A_2 \end{pmatrix}, \text{ and}$$

$$\Gamma_4(A) = \begin{pmatrix} -uvA_4 \\ -vA_3 \\ uA_2 \\ A_1 \end{pmatrix} \in \mathbb{R}^{4m\times n}.$$

The representations $\mathrm{V_c}, \Gamma_1, \Gamma_2, \Gamma_3$ and Γ_4 are clearly one-to-one. Moreover, A and $\Gamma_1(A)$ have the same (Frobenius) norm:

$$\begin{aligned}\|A\| &= \sqrt{\|A_1\|^2 + \|A_2\|^2 + \|A_3\|^2 + \|A_4\|^2} \\ &= \|\Gamma_1(A)\|. \end{aligned} \quad (7)$$

Proposition 7. *([1]) The following properties hold:*

(i) $\Gamma_1(A+B) = \Gamma_1(A) + \Gamma_1(B)$, $\Gamma_1(kA) = k\Gamma_1(A)$ *for any* $A, B \in \mathbb{Q}_{u,v}^{m\times n}$ *and* $k \in \mathbb{R}$.

(ii) $\mathcal{R}(A+B) = \mathcal{R}(A) + \mathcal{R}(B)$, $\mathcal{R}(kA) = k\mathcal{R}(A)$ *for any* $A, B \in \mathbb{Q}_{u,v}^{m\times n}$ *and* $k \in \mathbb{R}$.

(iii) $\mathcal{R}(AB) = \mathcal{R}(A)\mathcal{R}(B)$ *for any* $A \in \mathbb{Q}_{u,v}^{m\times n}$ *and* $B \in \mathbb{Q}_{u,v}^{n\times p}$.

III. General solutions of the generalized Sylvester matrix equation

In this section, we investigate Problem 1. From now on, we are given $A \in \mathbb{Q}_{u,v}^{m\times a}, B \in \mathbb{Q}_{u,v}^{b\times p}, C \in \mathbb{Q}_{u,v}^{m\times a}, D \in \mathbb{Q}_{u,v}^{b\times p}$, and $E \in \mathbb{Q}_{u,v}^{m\times p}$. In order to find the general solutions of Eq. (2), the following lemma will be used in a calculation.

Lemma 8. *For any* $X \in \mathbb{Q}_{u,v}^{a\times b}$*, we have*

$$\begin{pmatrix} \mathrm{V_c}(\Gamma_1(X)) \\ \mathrm{V_c}(\Gamma_2(X)) \\ \mathrm{V_c}(\Gamma_3(X)) \\ \mathrm{V_c}(\Gamma_4(X)) \end{pmatrix} = \mathcal{M}_{u,v}\, \mathrm{V_c}(\Gamma_1(X)), \quad (8)$$

where $\mathcal{M}_{u,v} = \begin{pmatrix} I_{4ab} \\ I_b \otimes N \otimes I_a \\ I_b \otimes K \otimes I_a \\ I_b \otimes T \otimes I_a \end{pmatrix} \in \mathbb{R}^{16ab\times 4ab}$.

Here, $N = \left(e_2^4 \;\; ue_1^4 \; -e_4^4 \; -ue_3^4\right)$,
$K = \left(e_3^4 \;\; e_4^4 \;\; ve_1^4 \;\; ve_2^4\right)$, *and*
$T = \left(e_4^4 \;\; ue_3^4 \; -ve_2^4 \; -uve_1^4\right), \in \mathbb{R}^{4\times 4}$.

Proof: A direct computations reveals that

$$\mathrm{V_c}(\Gamma_2(X)) = \begin{pmatrix} u\,\mathrm{col}_1(X_2) \\ \mathrm{col}_1(X_1) \\ -u\,\mathrm{col}_1(X_4) \\ -\,\mathrm{col}_1(X_3) \\ \vdots \\ u\,\mathrm{col}_b(X_2) \\ \mathrm{col}_b(X_1) \\ -u\,\mathrm{col}_b(X_4) \\ -\,\mathrm{col}_b(X_3) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & uI_a & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ I_a & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -uI_a & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & -I_a & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & uI_a & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & I_a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & -uI_a \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & -I_a & 0 \end{pmatrix} \times \begin{pmatrix} \mathrm{col}_1(X_1) \\ \mathrm{col}_1(X_2) \\ \mathrm{col}_1(X_3) \\ \mathrm{col}_1(X_4) \\ \vdots \\ \mathrm{col}_b(X_1) \\ \mathrm{col}_b(X_2) \\ \mathrm{col}_b(X_3) \\ \mathrm{col}_b(X_4) \end{pmatrix}$$

$$\begin{aligned} &= I_b \otimes \left[\left(e_2^4 \quad ue_1^4 \quad -e_4^4 \quad -ue_3^4\right) \otimes I_a\right] \mathrm{V_c}(\Gamma_1(X)) \\ &= (I_b \otimes N \otimes I_a)\, \mathrm{V_c}(\Gamma_1(X)). \end{aligned} \quad (9)$$

Similarly, we obtain

$$\begin{aligned} &\mathrm{V_c}(\Gamma_3(X)) \\ &\quad = I_b \otimes \left[\left(e_3^4 \quad e_4^4 \quad ve_1^4 \quad ve_2^4\right) \otimes I_a\right] \mathrm{V_c}(\Gamma_1(X)) \\ &\quad = (I_b \otimes K \otimes I_a)\, \mathrm{V_c}(\Gamma_1(X)) \end{aligned} \quad (10)$$

and

$$\begin{aligned} &\mathrm{V_c}(\Gamma_4(X)) \\ &\quad = I_b \otimes \left[\left(e_4^4 \quad ue_3^4 \quad -ve_2^4 \quad -uve_1^4\right) \otimes I_n\right] \mathrm{V_c}(\Gamma_1(X)) \\ &\quad = (I_b \otimes T \otimes I_a)\, \mathrm{V_c}(\Gamma_1(X)). \end{aligned} \quad (11)$$

From Eqs. (9), (10) and (11), we arrive at Eq. (8). ∎

Theorem 9. *Consider Eq.* (2)*. Let us denote*

$$\mathcal{F} = \left(\Gamma_1(B)^T \otimes \mathcal{R}(A)\right) + \left(\Gamma_1(D)^T \otimes \mathcal{R}(C)\right). \quad (12)$$

Then Eq. (2) *is consistent if and only if the following rank condition holds:*

$$\mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v} \;\; \mathrm{V_c}(\Gamma_1(E))] = \mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}]. \quad (13)$$

In both consistent and inconsistent cases, we have the following:

(i) *Problem 1 has the general exact/LS solutions represented by*

$$\begin{aligned} &\mathrm{V_c}(\Gamma_1(X)) \\ &\quad = (\mathcal{F}\mathcal{M}_{u,v})^\dagger\, \mathrm{V_c}(\Gamma_1(E)) \\ &\qquad + \left[I_{4mp} - (\mathcal{F}\mathcal{M}_{u,v})^\dagger(\mathcal{F}\mathcal{M}_{u,v})\right] w, \end{aligned} \quad (14)$$

where $w \in \mathbb{R}^{4mp}$ *is arbitrary.*

(ii) *Among the general solutions* (14)*, the minimal-norm one is given by*

$$\mathrm{V_c}(\Gamma_1(X)) = (\mathcal{F}\mathcal{M}_{u,v})^\dagger\, \mathrm{V_c}(\Gamma_1(E)). \quad (15)$$

(iii) *Problem 1 has a unique exact/LS solution given by* (15) *if* $\mathcal{F}\mathcal{M}_{u,v}$ *is of full-column rank.*

Proof: From Eqs. (2) and (7), we consider the associated norm-error

$$\|AXB + CXD - E\| = \|\Gamma_1(AXB + CXD - E)\|.$$

Now, Lemma 6 and Proposition 7 imply that

$$\begin{aligned}
&\Gamma_1(AXB + CXD - E)\\
&= \Gamma_1(AXB) + \Gamma_1(CXD) - \Gamma_1(E)\\
&= \mathcal{R}(A)\mathcal{R}(X)\Gamma_1(B) + \mathcal{R}(C)\mathcal{R}(X)\Gamma_1(D) - \Gamma_1(E)\\
&= \mathrm{V_c}\left[\mathcal{R}(A)\mathcal{R}(X)\Gamma_1(B) + \mathcal{R}(C)\mathcal{R}(X)\Gamma_1(D) - \Gamma_1(E)\right]\\
&= \left[\left(\Gamma_1(B)^T \otimes \mathcal{R}(A)\right) + \left(\Gamma_1(D)^T \otimes \mathcal{R}(C)\right)\right] \mathrm{V_c}\left(\mathcal{R}(X)\right)\\
&\quad - \mathrm{V_c}(\Gamma_1(E)).
\end{aligned}$$

Using Lemma 8, we obtain

$$\begin{aligned}
&\|AXB + CXD - E\|\\
&\quad = \|\mathcal{F}\,\mathrm{V_c}\left(\mathcal{R}(X)\right) - \mathrm{V_c}(\Gamma_1(E))\|\\
&\quad = \|\mathcal{F}\mathcal{M}_{u,v}\,\mathrm{V_c}(\Gamma_1(X)) - \mathrm{V_c}(\Gamma_1(E))\|.
\end{aligned}$$

Thus, the matrix equation (2) is equivalent to a real linear system

$$\mathcal{F}\mathcal{M}_{u,v}\,\mathrm{V_c}(\Gamma_1(X)) \;=\; \mathrm{V_c}(\Gamma_1(E)). \qquad (16)$$

According to Lemma 5, the system (16) is consistent if and only if the rank condition (13) holds. In both consistent and inconsistent cases, the same lemma allows us to express the formula of the general exact/LS solution to be Eq. (14). The statements (ii) and (iii) now follow from Lemma 5. ■

IV. Specific solutions of the generalized Sylvester matrix equation

In this section, we investigate Problems 2-4. Indeed, we would like to find imaginary (LS) solutions, real (LS) solutions, and solutions closet to a given matrix. The next lemma provides a real-vector representation of a generalized quaternion matrix.

Lemma 10. *Suppose $X = X_2 i + X_3 j + X_4 k \in \mathbb{Q}_{u,v}^{a\times b}$ where $X_2, X_3, X_4 \in \mathbb{R}^{a\times b}$. Then*

$$\mathrm{V_c}(\Gamma_1(X)) \;=\; \mathcal{K}\,\mathrm{V_c}\begin{pmatrix} X_2\\ X_3\\ X_4\end{pmatrix},$$

where

$$\mathcal{K} = I_b \otimes \mathcal{J}_a \in \mathbb{R}^{4ab\times 3ab} \text{ and}$$
$$\mathcal{J}_a = \begin{pmatrix} e_2^4 & e_3^4 & e_4^4\end{pmatrix} \otimes I_a \in \mathbb{R}^{4a\times 3a}.$$

Proof: A direct computations reveals that

$$\mathrm{V_c}(\Gamma_1(X)) \;=\; \begin{pmatrix} \mathrm{col}_1(0)\\ \mathrm{col}_1(X_2)\\ \mathrm{col}_1(X_3)\\ \mathrm{col}_1(X_4)\\ \vdots\\ \mathrm{col}_b(0)\\ \mathrm{col}_b(X_2)\\ \mathrm{col}_b(X_3)\\ \mathrm{col}_b(X_4)\end{pmatrix}$$

$$= \begin{pmatrix}
0 & 0 & 0 & \dots & 0 & 0 & 0\\
I_a & 0 & 0 & \dots & 0 & 0 & 0\\
0 & I_a & 0 & \dots & 0 & 0 & 0\\
0 & 0 & I_a & \dots & 0 & 0 & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\
0 & 0 & 0 & \dots & 0 & 0 & 0\\
0 & 0 & 0 & \dots & I_a & 0 & 0\\
0 & 0 & 0 & \dots & 0 & I_a & 0\\
0 & 0 & 0 & \dots & 0 & 0 & I_a
\end{pmatrix}
\times \begin{pmatrix} \mathrm{col}_1(X_2)\\ \mathrm{col}_1(X_3)\\ \mathrm{col}_1(X_4)\\ \vdots\\ \mathrm{col}_b(X_2)\\ \mathrm{col}_b(X_3)\\ \mathrm{col}_b(X_4)\end{pmatrix}$$

$$= I_b \otimes \left[\begin{pmatrix} e_2^4 & e_3^4 & e_4^4\end{pmatrix} \otimes I_a\right] \mathrm{V_c}\begin{pmatrix} X_2\\ X_3\\ X_4\end{pmatrix}$$

$$= (I_b \otimes \mathcal{J}_a)\,\mathrm{V_c}\begin{pmatrix} X_2\\ X_3\\ X_4\end{pmatrix} = \mathcal{K}\,\mathrm{V_c}\begin{pmatrix} X_2\\ X_3\\ X_4\end{pmatrix}.$$

■

Theorem 11. *Consider Eq.* (2)*. Let us denote $\mathcal{F}$ as in Eq.* (12)*. Then Eq.* (2) *is consistent if and only if*

$$\mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\mathcal{K}\;\; \mathrm{V_c}(\Gamma_1(E))] \;=\; \mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\mathcal{K}].$$

Moreover,

(i) *Problem 2 has the imaginary-part exact/LS solutions $X = X_2 i + X_3 j + X_4 k$ expressed as*

$$\begin{aligned}
&\mathrm{V_c}\begin{pmatrix} X_2\\ X_3\\ X_4\end{pmatrix}\\
&\quad = (\mathcal{F}\mathcal{M}_{u,v}\mathcal{K})^{\dagger}\,\mathrm{V_c}(\Gamma_1(E))\\
&\qquad + \left[I_{4mp} - (\mathcal{F}\mathcal{M}_{u,v}\mathcal{K})^{\dagger}(\mathcal{F}\mathcal{M}_{u,v}\mathcal{K})\right] w, \qquad (17)
\end{aligned}$$

where $w \in \mathbb{R}^{4mp}$ is arbitrary.

(ii) *The minimal-norm exact/LS solution* (17) *is given by same formula:*

$$\mathrm{V_c}\begin{pmatrix} X_2\\ X_3\\ X_4\end{pmatrix} \;=\; (\mathcal{F}\mathcal{M}_{u,v}\mathcal{K})^{\dagger}\,\mathrm{V_c}(\Gamma_1(E)). \qquad (18)$$

(iii) *Problem 2 has a unique exact/LS solution given by the formula* (18) *if $\mathcal{F}\mathcal{M}_{u,v}\mathcal{K}$ is of full-column rank.*

Proof: From the proof Theorem 9 and Lemma 10, we obtain

$$\begin{aligned}
&\|AXB + CXD - E\|\\
&\quad = \|\mathcal{F}\mathcal{M}_{u,v}\,\mathrm{V_c}(\Gamma_1(X)) - \mathrm{V_c}(\Gamma_1(E))\|\\
&\quad = \left\|\mathcal{F}\mathcal{M}_{u,v}\,\mathrm{V_c}\begin{pmatrix} 0\\ X_2\\ X_3\\ X_4\end{pmatrix} - \mathrm{V_c}(\Gamma_1(E))\right\|\\
&\quad = \left\|\mathcal{F}\mathcal{M}_{u,v}\mathcal{K}\,\mathrm{V_c}\begin{pmatrix} X_2\\ X_3\\ X_4\end{pmatrix} - \mathrm{V_c}(\Gamma_1(E))\right\|.
\end{aligned}$$

Thus, Eq. (2) is equivalent to a real linear system

$$\mathcal{F}\mathcal{M}_{u,v}\mathcal{K}\,\mathrm{V_c}\begin{pmatrix}X_2\\X_3\\X_4\end{pmatrix} = \mathrm{V_c}(\Gamma_1(E)). \tag{19}$$

By Lemma 5, the Eq. (19) is consistent if and only if

$$\mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\mathcal{K}\ \mathrm{V_c}(\Gamma_1(E))] = \mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\mathcal{K}].$$

The same lemma implies that the matrix equation (19) has the general exact/LS solutions

$$\begin{aligned}\mathrm{V_c}\begin{pmatrix}X_2\\X_3\\X_4\end{pmatrix} &= (\mathcal{F}\mathcal{M}_{u,v}\mathcal{K})^{\dagger}\,\mathrm{V_c}(\Gamma_1(E))\\ &\quad+\left[I_{4mp}-(\mathcal{F}\mathcal{M}_{u,v}\mathcal{K})^{\dagger}(\mathcal{F}\mathcal{M}_{u,v}\mathcal{K})\right]w,\end{aligned}$$

where $w \in \mathbb{R}^{4mp}$ is arbitrary. The statements (ii) and (iii) now follow from Lemma 5. ■

The next lemma will be used in a calculation involving Problem 3.

Lemma 12. *For any $X \in \mathbb{R}^{a\times b}$, we have*

$$\mathrm{V_c}(\Gamma_1(X)) = \tilde{\mathcal{K}}\,\mathrm{V_c}(X)$$

where $\tilde{\mathcal{K}} = I_b \otimes e_1^4 \otimes I_a \in \mathbb{R}^{4ab\times ab}$.

Proof: Since $X = X + 0i + 0j + 0k$, we have

$$\begin{aligned}\mathrm{V_c}(\Gamma_1(X)) &= \begin{pmatrix}\mathrm{col}_1(X)\\ \mathrm{col}_1(0)\\ \mathrm{col}_1(0)\\ \mathrm{col}_1(0)\\ \vdots\\ \mathrm{col}_b(X)\\ \mathrm{col}_b(0)\\ \mathrm{col}_b(0)\\ \mathrm{col}_b(0)\end{pmatrix}\\ &= \begin{pmatrix}I_a & \dots & 0\\ 0 & \dots & 0\\ 0 & \dots & 0\\ 0 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & I_a\\ 0 & \dots & 0\\ 0 & \dots & 0\\ 0 & \dots & 0\end{pmatrix}\begin{pmatrix}\mathrm{col}_1(X)\\ \vdots\\ \mathrm{col}_b(X)\end{pmatrix}\\ &= (I_b\otimes e_1^4\otimes I_a)\,\mathrm{V_c}(X)\\ &= \tilde{\mathcal{K}}\,\mathrm{V_c}(X).\end{aligned}$$

■

Theorem 13. *Consider Eq.* (2). *Let us denote $\mathcal{F}$ as in Eq.* (12). *Then Eq.* (2) *is consistent if and only if*

$$\mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}}\ \mathrm{V_c}(\Gamma_1(E))] = \mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}}].$$

Morever,

(i) *Problem 3 has the real exact/LS solutions $X \in \mathbb{R}^{a\times b}$ expressed as*

$$\begin{aligned}\mathrm{V_c}(X) &= (\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}})^{\dagger}\,\mathrm{V_c}(\Gamma_1(E))\\ &\quad+\left[I_{4mp}-(\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}})^{\dagger}(\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}})\right]y, \end{aligned}\tag{20}$$

where $y \in \mathbb{R}^{16bp}$ is an arbitrary vector.

(ii) *The minimal-norm exact/LS solution* (20) *is given by the same formula:*

$$\mathrm{V_c}(X) = (\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}})^{\dagger}\,\mathrm{V_c}(\Gamma_1(E)). \tag{21}$$

(iii) *Problem 3 has a unique exact/LS solution given by the formula* (21) *if $\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}}$ is of full-column rank.*

Proof: Since $X = X + 0i + 0j + 0k$ and from the proof Theorem 9 and Lemma 12, we obtain

$$\begin{aligned}&\|AXB+CXD-E\|\\ &\quad= \|\mathcal{F}\mathcal{M}_{u,v}\,\mathrm{V_c}(\Gamma_1(X))-\mathrm{V_c}(\Gamma_1(E))\|\\ &\quad= \left\|\mathcal{F}\mathcal{M}_{u,v}\,\mathrm{V_c}\begin{pmatrix}X_1\\0\\0\\0\end{pmatrix}-\mathrm{V_c}(\Gamma_1(E))\right\|\\ &\quad= \left\|\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}}\,\mathrm{V_c}(X)-\mathrm{V_c}(\Gamma_1(E))\right\|.\end{aligned}$$

Thus, Eq. (2) is equivalent to a real linear system

$$\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}}\,\mathrm{V_c}(\Gamma_1(X)) = \mathrm{V_c}(\Gamma_1(E)). \tag{22}$$

By Lemma 5, Eq. (22) is consistent if and only if

$$\mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}}\ \mathrm{V_c}(\Gamma_1(E))] = \mathrm{rank}[\mathcal{F}\mathcal{M}_{u,v}\tilde{\mathcal{K}}].$$

The assertions $(i)-(iii)$ now follow from Lemma 5. ■

Now, we investigate Problem 4.

Theorem 14. *Consider Eq.* (2). *Let $Y \in \mathbb{Q}_{u,v}^{a\times b}$ be given.*

(i) *Problem 4.1 has the solution $X = Y + Z \in \mathbb{Q}_{u,v}^{n\times p}$, where Z is the general exact/LS solution of the associated matrix equation*

$$AZB + CZD = E - (AYB + CYD). \tag{23}$$

(ii) *Problem 4.2 has the solution $X = Y + Z$, where Z is the imaginary-part exact/LS solution of Eq.* (23).

(iii) *Problem 4.3 has the solution $X = Y + Z$, where Z is the real exact/LS solution of Eq.* (23).

Proof: Denote $\breve{E} = E-(AYB+CYD)$. Let us denote by $G_{\mathcal{S}}$ the set of general exact/LS solutions of the equation $AXB + CXD = E$. Consider the following error

$$\begin{aligned}&AXB+CXD-E\\ &\quad= AXB+CXD-E-AYB\\ &\qquad -CYD+AYB+CYD\\ &\quad= A(X-Y)B+C(X-Y)D\\ &\qquad -E+AYB+CYD\\ &\quad= AZB+CZD-\breve{E}.\end{aligned}\tag{24}$$

By letting $Z = X - Y$, we have that the Problem 4.1 is equivalent to the following minimization

$$\begin{aligned}\min_{X\in G_{\mathcal{S}}}\|X-Y\| &= \min_{\|AXB+CXD-E\|=\min}\|X-Y\|\\ &= \min_{\|AXB+CXD-E\|=\min}\|Z\|\\ &= \min_{\|AZB+CZD-\breve{E}\|=\min}\|Z\|.\end{aligned}$$

Similarly, we obtain the statements (ii) and (iii). ■

V. Sylvester and Stein matrix equations, and quaternionic matrix equations

The generalized Sylvester equation (2) includes the following special cases.

Corollary 15. *Consider the Sylvester matrix equation*

$$AX + XD = E \tag{25}$$

in an unknown $X \in \mathbb{Q}_{u,v}^{a\times b}$. Here, the matrices $A \in \mathbb{Q}_{u,v}^{m\times a}, D \in \mathbb{Q}_{u,v}^{b\times p}$, and $E \in \mathbb{Q}_{u,v}^{m\times p}$ are given. Then the conclusions of Theorems 9, 11, 13, and 14 hold, where the matrix $\mathcal{F}$ is given by

$$\mathcal{F} = \left((e_1^4)^T \otimes I_b \otimes \mathcal{R}(A)\right) + \left(\Gamma_1(D)^T \otimes I_{4a}\right). \tag{26}$$

Proof: We set $B = I_b$ and $C = I_a$ in those theorems. Note that $\Gamma_1(I_b) = e_1^4 \otimes I_b$. So, the matrix $\mathcal{F}$ in (12) *is reduced to* (26). ∎

In the next corollary, we consider the Stein matrix equation.

Corollary 16. *Consider the Stein matrix equation*

$$AXB + X = E$$

in an unknown $X \in \mathbb{Q}_{u,v}^{n\times p}$. Here, $A \in \mathbb{Q}_{u,v}^{m\times a}, B \in \mathbb{Q}_{u,v}^{b\times p}$, and $E \in \mathbb{Q}_{u,v}^{m\times p}$ are given. Then the conclusions of Theorems 9, 11, 13, and 14 hold, where the matrix $\mathcal{F}$ is given by

$$\mathcal{F} = \left(\Gamma_1(B)^T \otimes \mathcal{R}(A)\right) + \left((e_1^4)^T \otimes I_{4ab}\right).$$

Proof: We set $C = I_a$ and $D = I_b$ in those theorems. ∎

In particular when $u = v = -1$, the previous results in Sections III and IV become those for matrices over the Hamilton quaternions $\mathbb{Q}$.

Corollary 17. *Consider the matrix equation*

$$AXB + CXD = E,$$

where $A, C \in \mathbb{Q}^{m\times a}, B, D \in \mathbb{Q}^{p\times b}$, and $E \in \mathbb{Q}^{m\times p}$ are given. Then the conclusions of Theorems 9, 11, 13, and 14 hold, where the matrix $\mathcal{M}_{u,v}$ is given explicitly by

$$\mathcal{M}_{-1,-1} = \begin{pmatrix} I_{4ab} \\ I_b \otimes \hat{N} \\ I_b \otimes \hat{K} \\ I_b \otimes \hat{T} \end{pmatrix} \in \mathbb{R}^{16ab\times 4ab}$$

and

$$\hat{N} = \begin{pmatrix} 0 & -I_a & 0 & 0 \\ I_a & 0 & 0 & 0 \\ 0 & 0 & 0 & I_a \\ 0 & 0 & -I_a & 0 \end{pmatrix},$$

$$\hat{K} = \begin{pmatrix} 0 & 0 & -I_a & 0 \\ 0 & 0 & 0 & -I_a \\ I_a & 0 & 0 & 0 \\ 0 & I_a & 0 & 0 \end{pmatrix},$$

$$\hat{T} = \begin{pmatrix} 0 & 0 & 0 & -I_a \\ 0 & 0 & I_a & 0 \\ 0 & -I_a & 0 & 0 \\ I_a & 0 & 0 & 0 \end{pmatrix}.$$

Proof: Set $u = v = -1$ in those theorems. The conclusions of Theorems 9, 11, and 13 were investigated in [22]. ∎

VI. Numerical examples

In this section, we provide numerical examples to illustrate our results.

Example 18. *Consider the generalized Sylvester matrix equation $AXB + CXD = E$ over the split quaternions (i.e., $(u, v) = (-1, 1)$), where*

$$A = \begin{pmatrix} 1 & -j+2k \end{pmatrix}_{1\times 2}, \quad C = \begin{pmatrix} i-j-k & 2 \end{pmatrix}_{1\times 2},$$
$$B = \begin{pmatrix} j-k \\ 3-i \end{pmatrix}_{2\times 1}, \quad D = \begin{pmatrix} 3k \\ 2-j \end{pmatrix}_{2\times 1},$$
$$E = \begin{pmatrix} -2+i-4j+k \end{pmatrix}_{1\times 1}.$$

Then we have

$$\mathcal{R}(A) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 2 & 0 & 1 \\ 0 & -1 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix},$$
$$\mathcal{R}(C) = \begin{pmatrix} 2 & 0 & -1 & 0 & -1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 2 & -1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 2 & -1 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$
$$\Gamma_1(B)^T = \begin{pmatrix} 0 & 3 & 0 & -1 & 1 & 0 & -1 & 0 \end{pmatrix},$$
$$\Gamma_1(D)^T = \begin{pmatrix} 0 & 2 & 0 & 0 & 0 & -1 & 3 & 0 \end{pmatrix},$$
$$\Gamma_1(E) = \begin{pmatrix} -2 & 1 & -4 & 1 \end{pmatrix}^T,$$

and

$$\mathcal{F} = \left(\Gamma_1(B)^T \otimes \mathcal{R}(A)\right) + \left(\Gamma_1(D)^T \otimes \mathcal{R}(C)\right).$$

We see that

$$\operatorname{rank}\left[\mathcal{F}\mathcal{M}_{-1,1} \;\; \mathrm{V_c}(\Gamma_1(E))\right] = \operatorname{rank}\left[\mathcal{F}\mathcal{M}_{-1,1}\right] = 4,$$

thus Eq. (2) *is consistent. According to Theorem 9, the matrix equation has a minimal-norm solution, computed via MATLAB as follows:*

$$X = \begin{pmatrix} 0.0919 & -0.0157 \\ 0.0815 & 0.1384 \end{pmatrix} + \begin{pmatrix} -0.0595 & 0.0449 \\ 0.1372 & -0.0747 \end{pmatrix} i + \begin{pmatrix} -0.1335 & -0.0537 \\ -0.0325 & -0.0181 \end{pmatrix} j + \begin{pmatrix} 0.0809 & 0.0804 \\ -0.1514 & -0.0696 \end{pmatrix} k.$$

Example 19. *Consider the generalized Sylvester matrix equation $AXB + CXD = E$ over the split quaternions (i.e., $(u, v) = (-1, 1)$), where*

$$A = \begin{pmatrix} 2i & 1+j \\ i-k & 0 \end{pmatrix}_{2\times 2}, \quad C = \begin{pmatrix} i+3k & -k \\ 1+j & 0 \end{pmatrix}_{2\times 2},$$
$$B = \begin{pmatrix} 1 \\ -2j \end{pmatrix}_{2\times 1}, \quad D = \begin{pmatrix} -1 \\ -k \end{pmatrix}_{2\times 1}, \quad E = \begin{pmatrix} -1 \\ -2i \end{pmatrix}_{2\times 1}.$$

Then we have

$$\mathcal{R}(A) = \begin{pmatrix} 0 & 1 & -2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathcal{R}(C) = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 3 & -1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 3 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 3 & -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

$$\Gamma_1(B)^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -2 & 0 & 0 \end{pmatrix},$$
$$\Gamma_1(D)^T = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix},$$
$$\Gamma(E) = \begin{pmatrix} -1 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \end{pmatrix}^T,$$

and

$$\mathcal{F} = \left(\Gamma_1(B)^T \otimes \mathcal{R}(A)\right) + \left(\Gamma_1(D)^T \otimes \mathcal{R}(C)\right).$$

We see that

$$\operatorname{rank}[\mathcal{FM}_{-1,1}\mathcal{K}\ \mathrm{V_c}(\Gamma_1(E))] = 7 \quad \textit{and}$$
$$\operatorname{rank}[\mathcal{FM}_{-1,1}\mathcal{S}] = 6,$$

thus Eq. (2) *is inconsistent. According to Theorem 11, the matrix equation has a minimal-norm LS solution, computed via MATLAB as follows:*

$$X = \begin{pmatrix} 0.0139 & -0.0765 \\ -0.1781 & 0.0084 \end{pmatrix} i + \begin{pmatrix} 0.0551 & 0.0575 \\ 0.0007 & -0.3225 \end{pmatrix} j + \begin{pmatrix} -0.2938 & -0.3269 \\ 0.2865 & -0.0352 \end{pmatrix} k.$$

Example 20. *Consider the matrix equation $AXB+CXD = E$ over the split quaternions, i.e., $(u,v) = (-1,1)$. Here, we are given the matrices A, B, C, D, E as in Example 6.1, and we will find a solution X closest to a given matrix*

$$Y = \begin{pmatrix} -i & 1 \\ 0 & j \end{pmatrix}.$$

We obtain

$$\check{E} = E - (AYB + CYD) = \left(1 + i + -3j\right).$$

and

$$\operatorname{rank}[\mathcal{FM}_{-1,1}\ \mathrm{V_c}(\Gamma_1(\check{E}))] = \operatorname{rank}[\mathcal{FM}_{-1,1}] = 4.$$

Using Theorem 14 and MATLAB, we obtain:

$$Z = \begin{pmatrix} 0.0329 & 0.0573 \\ -0.0002 & 0.1340 \end{pmatrix} + \begin{pmatrix} 0.0132 & 0.0298 \\ 0.1151 & -0.0503 \end{pmatrix} i + \begin{pmatrix} -0.0260 & -0.0834 \\ -0.0462 & -0.0356 \end{pmatrix} j + \begin{pmatrix} 0.0609 & 0.0406 \\ -0.0197 & 0.0325 \end{pmatrix} k.$$

Thus, we get the desire exact solution $X = Z + Y$:

$$X = \begin{pmatrix} 0.0329 & 1.0573 \\ -0.0002 & 0.1340 \end{pmatrix} + \begin{pmatrix} -0.9868 & 0.0298 \\ 0.1151 & -0.0503 \end{pmatrix} i + \begin{pmatrix} -0.0260 & -0.0834 \\ -0.0462 & 0.9644 \end{pmatrix} j + \begin{pmatrix} 0.0609 & 0.0406 \\ -0.0197 & 0.0325 \end{pmatrix} k.$$

VII. GRADIENT-DESCENT ITERATIVE (GDI) ALGORITHM FOR THE ASSOCIATED LINEAR SYSTEM

From the discussion in Section VII, we see that the LS solution of Eq. (2) is equivalent to a real linear system (16). In order to resolve this system, we utilize a technique for enhancing the gradient descent optimization, as described in [27]. The core concept is to reduce the residual error $\|\mathcal{FM}_{u,v}\,\mathrm{V_c}(\Gamma_1(X)) - \mathrm{V_c}(\Gamma_1(E))\|$ at each interation. Consequently, we derive the following gradient-descent iterative (GDI) algorithm:

Algorithm 1: GDI Algorithm for system (16)

$A \in \mathbb{Q}_{u,v}^{m\times a}$, $B \in \mathbb{Q}_{u,v}^{b\times p}$, $C \in \mathbb{Q}_{u,v}^{m\times a}$, $D \in \mathbb{Q}_{u,v}^{b\times p}$, and $E \in \mathbb{Q}_{u,v}^{m\times p}$. ;
Set $i = 0$. Choose $x^{(0)} \in \mathbb{R}^{4ab}$. Compute
$\hat{A} = \mathcal{FM}_{u,v}$, $\hat{f} = \mathrm{V_c}(\Gamma_1(E))$, $\hat{P} = \hat{A}\hat{A}^T$.
for $i = 0, 1, 2, 3, \ldots$ **do**
  $r^{(i)} = \hat{f} - \hat{A}\mathrm{x}^{(i)}$;
  if $\|r^{(i)}\| \leqslant \epsilon$ **then**
    $\mathrm{x}^{(i)}$ is a solution; break;
  else
    $m_i = \hat{P}r^{(i)}$;
    $\alpha_{i+1} = m_i^T r^{(i)}/(2m_i^T m_i)$;
    $\mathrm{x}^{(i+1)} = \mathrm{x}^{(i)} + \alpha_{i+1}\hat{A}^T r^{(i)}$;
  end
  update i;
end

We implement all simulations using MATLAB R2017a on the same PC environment: AMD A9-9425 RADEON R5 @3.10GHz with RAM 4 GB.

Example 21. *Let $(u,v) = (-1,-1)$. Consider the generalized Sylvester matrix equation $AXB + CXD = E$, where $A, B, C, D, E \in \mathbb{Q}_{u,v}^{2\times 2}$ are given randomly as follows:*

$$A = \begin{pmatrix} 0.8147 & 0.1270 \\ 0.9058 & 0.9134 \end{pmatrix} + \begin{pmatrix} 0.6324 & 0.2785 \\ 0.0975 & 0.5469 \end{pmatrix} i + \begin{pmatrix} 0.9575 & 0.1576 \\ 0.9649 & 0.9706 \end{pmatrix} j + \begin{pmatrix} 0.9572 & 0.8003 \\ 0.4854 & 0.1419 \end{pmatrix} k,$$

$$B = \begin{pmatrix} 0.4218 & 0.7922 \\ 0.9157 & 0.9595 \end{pmatrix} + \begin{pmatrix} 0.6557 & 0.8491 \\ 0.0357 & 0.9340 \end{pmatrix} i + \begin{pmatrix} 0.6787 & 0.7431 \\ 0.7577 & 0.3922 \end{pmatrix} j + \begin{pmatrix} 0.6555 & 0.7060 \\ 0.1712 & 0.0318 \end{pmatrix} k,$$

$$C = \begin{pmatrix} 0.2769 & 0.0971 \\ 0.0462 & 0.8235 \end{pmatrix} + \begin{pmatrix} 0.6948 & 0.9502 \\ 0.3171 & 0.0344 \end{pmatrix} i + \begin{pmatrix} 0.4387 & 0.7655 \\ 0.3816 & 0.7952 \end{pmatrix} j + \begin{pmatrix} 0.1869 & 0.4456 \\ 0.4898 & 0.6463 \end{pmatrix} k,$$

$$D = \begin{pmatrix} 0.7094 & 0.2760 \\ 0.7547 & 0.6797 \end{pmatrix} + \begin{pmatrix} 0.6551 & 0.1190 \\ 0.1626 & 0.4984 \end{pmatrix} i + \begin{pmatrix} 0.9597 & 0.5853 \\ 0.3404 & 0.2238 \end{pmatrix} j + \begin{pmatrix} 0.7513 & 0.5060 \\ 0.2551 & 0.6991 \end{pmatrix} k,$$

$$E = \begin{pmatrix} 0.8909 & 0.5472 \\ 0.9593 & 0.1386 \end{pmatrix} + \begin{pmatrix} 0.1493 & 0.8407 \\ 0.2575 & 0.2543 \end{pmatrix} i + \begin{pmatrix} 0.8143 & 0.9293 \\ 0.2435 & 0.3500 \end{pmatrix} j + \begin{pmatrix} 0.1966 & 0.6160 \\ 0.2511 & 0.4733 \end{pmatrix} k.$$

We would like to find an LS solution

$$X = X_1 + X_2 i + X_3 j + X_4 k.$$

We apply Algorithm 1 with an intial guess $X^{(0)}$ *is a zero matrix and a tolerance error* $\epsilon = 0.005$. *The relative error* $\|r^{(i)}\|$ *at each iteration is illustrate in Figure 1.*

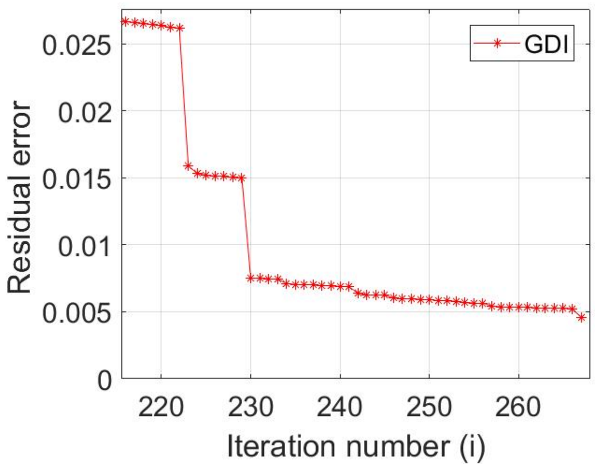

Fig. 1: The relative error at each iteration for Example 20.

It turns out that Algorithm 1 requires 267 iterations and an approximate times of 0.161798 seconds to arrive a desire solution.

$$X^{(267)} = \begin{pmatrix} 0.4155 & -0.0476 \\ -0.2244 & -0.1754 \end{pmatrix} + \begin{pmatrix} -0.0331 & -0.7911 \\ 0.1011 & 0.3502 \end{pmatrix} i + \begin{pmatrix} -0.3315 & 0.5111 \\ 0.0976 & -0.6186 \end{pmatrix} j + \begin{pmatrix} -0.0077 & 0.5213 \\ -0.0355 & -0.3490 \end{pmatrix} k.$$

VIII. Application to color image processing

In color image processing, a color image can be represented by a vector. According to RGB color model, the color information of a pixel can be represented as

$$w = 0 + w_r i + w_g j + w_b k \in \mathbb{Q},$$

where w_r, w_g, w_b are the red/green/blue component of the color pixel, respectively. We can see that w_r, w_g and w_b are the imaginary part of w. In [28], the general image degradation model is provided by

$$Aw + b = f, \tag{27}$$

where f is the observation, w is the desired image, b is the additive noise. The matrix A acts as a linear operator related to the task. For example, A is the an identity matrix for the image denoising task, A is a projection matrix for image inpainting, A is the blur matrix related to the blur kernel for the image deblurring task. Our task is to restore the desired image w from the observation f.

Note that Eq. (27) can be written as $Aw = f - b$. Thus, Eq. (27) is a special case of Eq. (2) when $E = f - b, B = I_1, C = 0$ and $X = w$. We can solve for a minimal-norm least-squares solution solution by using Theorem 11 as follows:

$$\begin{aligned} \mathrm{V_c}\begin{pmatrix} w_r \\ w_g \\ w_b \end{pmatrix} &= (\mathcal{F}\mathcal{M}_{-1,-1}\mathcal{K})^\dagger \mathrm{V_c}(\Gamma_1(f-b)) \\ &= \left[\left(\Gamma_1(I_1)^T \otimes \mathcal{R}(A)\right)\mathcal{M}_{-1,-1}\mathcal{K}\right]^\dagger \\ &\quad \times \mathrm{V_c}(\Gamma_1(f-b)) \\ &= \left[\left((e_4^1)^T \otimes \mathcal{R}(A)\right)\mathcal{M}_{-1,-1}\mathcal{J}_a\right]^\dagger \\ &\quad \times \mathrm{V_c}(\Gamma_1(f-b)). \end{aligned} \tag{28}$$

We then get w_r, w_g and w_b due to the injectivity of the operator $\mathrm{V_c}$.

We summarize the process of image deblurring as in the following algorithm.

Algorithm 2: Algorithm for image deblurring

---

(1) Import an original image.
(2) Blur the image using the option in MATLAB R2019b, namely,

f = fspecial('motion', len, theta)

where the parameters *len* and *theta* indicate the length and the angle of motion in degrees in a counter-clockwise direction, respectively.
(3) Determine the value of A from image bluring.
(4) Calculate w according to Eq. (28).

Note that to simplify computations on a small-scale computer, an original image could be resized to a smaller dimension.

Example 22. *Given an original color image of size* 100×100 *pixels. In order to reduce time and memory for computations, we modify the pixels of the image to a smaller dimension* 80×80, *as in Fig. 2. We represent the image Fig. 2 as a* 3*-tuples of vector* $w = (w_r, w_g, w_b)$. *Then, we blur this image using len = 10 and theta = 40 to get the image* $f = (f_r, f_g, f_b)$ *as in Fig. 3. The image Fig. 2 can be recovered as the minimal-norm least-squares solution of the model* (27). *Indeed, by using* (28), *we get the restored image as shown in Fig. 4.*

Fig. 2: The (80×80)-pixels image.

Fig. 3: The blurred image.

Fig. 4: The restored image.

Example 23. *Given an original color image of size* 350×350 *pixels. In order to reduce time and memory for computations, we modify the pixels of the image to a smaller dimension* 80×80*, as in Fig. 5. We apply the command fspecial('motion', len, theta) where len = 30 and theta = 11 to disturb Fig. 5. So, we get the blurred image as in Fig. 6. From the color model* (27)*, we can restore the image Fig. 7 as the minimal-norm least-squares solution given by* (28)*. Indeed, the restored image is shown in Fig. 7.*

Fig. 5: The (80×80)-pixels image.

Fig. 6: The blurred image.

Fig. 7: The restored image.

Example 24. *Given an original color image of size* 3000×3000 *pixels. In order to reduce time and memory for computations, we modify the pixels of the image to a smaller dimension* 80×80*, as in Fig. 8. We apply the command fspecial('motion', len, theta) to disturb Fig. 8. So, we get the blurred image as in Fig. 9 and Fig. 10. The image Fig. 2 can be recovered as the minimal-norm least-squares solution of the model* (27)*.*

Fig. 8: The (80×80)-pixels image.

Fig. 9: The blurred image using *len = 10* and *theta = 40*.

Fig. 10: The restored image.

Fig. 11: The blurred image using *len = 30* and *theta = 11*.

Fig. 12: The restored image.

TABLE I: The LS errors of the restored the image and CPU times

	$\|f_r - w_r\|^2$	$\|f_r - w_r\|^2$	$\|f_r - w_r\|^2$	CPU times
Fig. 4	6.3029e-13	6.2912e-13	2.0018e-13	8549.952s
Fig. 7	2.2437e-12	3.5017e-13	4.2708e-13	7937.218s
Fig. 10	3.7830e-13	5.2965e-13	5.4408e-13	8584.374s
Fig. 12	4.7741e-13	7.0398e-13	5.3691e-13	8404.744s

IX. Conclusion

We investigate Problems 1-4 to find general and specific exact/LS solutions of the generalized Sylvester matrix equation (2). All matrices considered here are rectangular compatible matrices over a generalized quaternion. We apply the techniques of real representations and vectorizations of generalized quaternion matrices to reduce the matrix equation (2) to a real linear system. Thus, we can extract a solvability criterion for the matrix equation. Moreover, we can derive formulas of the (minimal-norm) exact/LS solution, the (minimal-norm) pure-imaginary exact/LS solution, the (minimal-norm) real exact/LS solution, and the (minimal-norm) exact/LS solution closet to a given matrix. Such solutions are expressed in terms of Kronecker products and Moore-Penrose inverses. This work includes the studies of Sylvester and Stein equations over generalized quaternions, quaternionic matrix equations, and particularly the work [22]. Moreover, we propose gradient-descent iterative (GDI) algorithm to solve the linear system associated with Eq. (2). In color image processing, we can apply our theory to get an algorithm for image deblurring.

References

[1] X. Liu and Y. Zhang, "Matrices over Quaternion Algebras", in *Matrix and Operator Equations and Applications 2023*, pp. 139–183

[2] J. Ping and H. T. Wu, "A closed-form forward kinematics solution for the 6-6^p Stewart platform", *IEEE Trans. Robot. Autom.*, vol. 17, pp. 522–526, Aug. 2001.

[3] C. E. Moxey, S. J. Sangwine, and T. A. Ell, "Hypercomplex correlation techniques for vector imagines", *IEEE Trans. Signal Process.*, vol. 51, pp. 1941–1953, Jun. 2003.

[4] N.L. Bihan and J. Mars, "Singular value decomposition of quaternion matrices: A new tool for vector-tensor signal processing", *IEEE Trans. Signal Process.*, vol. 84, pp. 1177–1199, Jul. 2004.

[5] S. L. Adler, "Scattering and decay theory for quaternionic quantum mechanics and structure of induced t nonconservation", *Phys. Rev. D*, vol. 37, pp. 3654–3662, Jun. 1988.

[6] M. Denielewski and L. Sapa, "Foundations of the quaternion quantum mechanics", *Entropy*, vol. 22, pp. 1424, Mar. 2020.

[7] F. X. Zhang, M. S. Wei, Y. Li, and J. L. Zhao, "Special least squares solutions of the quaternion matrix equation $AX = B$ with applications", *Appl. Math. Comput.*, vol. 270, pp. 425–433, Nov. 2015.

[8] F. Caccavale, C. Natale, B. Siciliano, and L. Villani, "Six-DOF impedance control based on angle/axis representations", *IEEE Trans. Robot. Autom.*, vol.15, pp. 289–300, Apr. 1999.

[9] M. Ma, "Color image restoration via quaternion-based hybrid regularization method", *IAENG International Journal of Applied Mathematics*, vol. 54, pp. 2176–2182, Nov. 2024.

[10] Y. Xu, L. Yu, H. Xu, H . Zhang and T. Nguyen, "Vector sparse representation of color image using quaternion matrix analysis", *IEEE Trans. Image Process.*, vol. 24, pp. 1315–1329, Jan. 2015.

[11] G. Obinata and B. D. O. Anderson, *Model reduction for control system design*, London, UK: Springer London, 2001.

[12] A. Bouhamidi and K. Jbilou, "A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications", *Appl. Math. Comput.*, vol. 206, pp. 687–694, Dec. 2008.

[13] Z. Z. Bai, M. Benzi, and F. Chen, "Modified HSS iteration methods for a class of complex symmetric linear systems", *Computing*, vol. 87,pp. 93–111, Feb. 2010.

[14] T. W. Chen and B. A. Franci, *Optimal Sampled-Data control systems*, London, UK: Springer London, 1995.

[15] B. Datta, *Numerical methods for linear control systems*, Amsterdam, Netherlands: Elsevier Inc., 2004.

[16] Q. W. Wang, A. Rehman, Z. H. He, et al., "Constraint generalized Sylvester matrix equations", *Automatica*, vol. 69, pp. 60–64, Jul. 2016.

[17] Q. W. Wang and Z. H. He, "Systems of coupled generalized Sylvester matrix equations", *Automatica*, vol. 50, pp. 2840–2844, Nov. 2014.

[18] M. Dehghan and A. Shirilord, "The use of homotopy analysis method for solving generalized Sylvester matrix equation with applications", *Engineering with Computers*, vol. 38, pp. 1–18, Jan. 2022.

[19] S. G. Shafiei and M. Hajarian, "An iterative method based on ADMM for solving generalized Sylvester matrix equations", *Journal of the Franklin Institute*, vol. 359, pp. 8155-8170, Oct. 2022.

[20] N. Sasaki and P. Chansangiam, "Modified Jacobi-Gradient iterative method for generalized Sylvester matrix equation", *Symmetry*, vol. 12, pp. 1831, Nov. 2020.

[21] S. F. Yuan, "Least squares pure imaginary solution and real solution of quaternion matrix equation $AXB+CXD = E$ with the least norm", *Journal of Applied Mathematics*, vol. 2014, pp. 1–9, Apr. 2014.

[22] F. Zhang, W. Mu, Y. Li, and J. Zhao, "Special least squares solutions of the quaternion matrix equation $AXB + CXD = E$", *Comput. Math. Appl.*, vol. 72, pp. 1426–1435, Sep. 2016.

[23] Y. Tian, X. Liu, and S. F. Yuan, "On Hermitian solutions of the generalized quaternion matrix equation $AXB + CXD = E$", *Mathematical Problems in Engineering*, vol. 2021, pp. 1–10, Dec. 2021.

[24] J. Jaiprasert and P. Chansangiam, "Exact and least-squares solutions of the generalized Sylvester-transpose matrix equation over generalized quaternions", *Electronic Research Archive*, vol. 32, pp. 2789–2804, Apr. 2024.

[25] A. B. Israel and T. N. E. Greville, *Generalized Inverses: Theory and applications*, 3rd ed. New York, US: Springer, 2003.

[26] D. A. Turkington, *Matrix Calculus & Zero-One Matrices: Statistical and Econometric Applications*, Cambridge, England: Cambridge University Press, 2002.

[27] K. Tansri, P. Chansangiam, "Gradient-descent iterative algorithm for solving exact and weighted least-squares solutions of rectangular linear systems," *AIMS Mathematics*, vol. 8, pp. 11781–11798, Mar. 2023.

[28] C. Huang, J. Li, and G. Gao, "Review of quaternion-based color image processing methods", *Math.*, vol. 11, pp. 2056, Apr. 2023.