Journal of
New Theory
ISSN: 2149-1402

47 (2024) 39-51
Journal of New Theory
https://dergipark.org.tr/en/pub/jnt
Open Access

A Global Krylov Subspace Method for the Sylvester Quaternion Matrix Equation

Sinem Şimşek[1]

Article Info
Received: 17 Apr 2024
Accepted: 24 Jun 2024
Published: 30 Jun 2024
doi:10.53570/jnt.1469996
Research Article

Abstract — This study concerns the Sylvester matrix equation in the quaternion setting when the coefficient matrices as well as the unknown matrix have quaternion entries. We propose a global Generalized Minimal Residual (GMRES) method for the solution of such a matrix equation. The proposed approach works directly with the Sylvester operator to generate orthonormal bases for Krylov subspaces formed of matrices. Then, the best approximate matrix solution to the Sylvester equation at hand in such a Krylov subspace is constructed from a matrix minimizing the Frobenius norm of the residual. We describe how this minimization of the residual norm can be carried out efficiently and report numerical results on real examples related to image restoration.

Keywords *Sylvester quaternion matrix equation, quaternion Krylov subspace, global GMRES, quaternion Arnoldi process*

Mathematics Subject Classification (2020) 15A24, 15B33

1. Introduction

The Sylvester matrix equation is of the form

$$AX + XB = C \tag{1.1}$$

where X is the $n \times m$ unknown matrix, and A, B, and C are given matrices with appropriate sizes [1]. Such a matrix equation and its special case, a Lyapunov matrix equation [2], arise in fields, such as control theory, eigenstructure assignment, model reduction, image restoration problems, numerical solutions of ordinary differential equations [3–9]. On the other side, quaternions have applications in various fields, including those from computer science, quantum mechanics, signal and color image processing [10, 11]. Due to these wide ranges of applications for Sylvester equations, as well as quaternions, the problem of obtaining solutions to (1.1), specifically over the skew-field of quaternions, has attracted considerable attention [12–15]. Matrix equations other than Sylvester equations over quaternions have also been studied in the literature [16–21].

In general, direct or iterative numerical methods are employed to find the solutions of (1.1) depending on the size of A and B. When the coefficient matrices have sizes of a few hundred at most, the problem is referred to as a small- or medium-scale problem. For these problems, the most efficient method is a direct method proposed by Bartels and Stewart [22]. This method is based on the Schur decompositions of the coefficient matrices A and B, resulting in a Sylvester matrix equation in a

[1]sinem.simsek@klu.edu.tr (Corresponding Author)
[1]Department of Mathematics, Faculty of Arts and Sciences, Kırklareli University, Kırklareli, Türkiye

simplified form that is easily solved by back substitution. When both coefficient matrices are small, but one is significantly smaller than the other, Golub et al. have presented a variant of Bartels and Stewart algorithm by means of the Hessenberg decomposition of the larger coefficient matrix [23]. A classical alternative approach is to turn the Sylvester matrix equation into a linear system by using the Kronecker product and vec operator. Then, the LU factorization with partial pivoting can be applied to this linear system for finding the solution. Apart from these methods, if only one of n and m is large, several approaches are available based on a decomposition of the smaller coefficient matrix. However, these approaches are not useful when both n and m are large (i.e., typically larger than 200). In the case of such large problems, commonly employed techniques to solve Sylvester equations are alternate direction implicit (ADI) iteration and projection methods. For instance, Krylov subspace methods are commonly employed projection methods for such large Sylvester equations when coefficients matrices are sparse. For large Sylvester matrix equations over real or complex fields, block and global Krylov subspace methods have attracted substantial interest in the literature [24–34]. While a block Arnoldi process constructs orthonormal bases for several subspaces of $\mathbb{C}^n$ or $\mathbb{R}^n$, simultaneously, the global Arnoldi process constructs an orthonormal basis for a subspace of a space of matrices.

Linear matrix equations over quaternions rather than over real or complex numbers come up with additional challenges, especially as the multiplication of two quaternion scalars is not commutative. It is possible to convert a quaternion matrix equation into a real or a complex matrix equation by employing real or complex representations of quaternion matrices. However, this conversion is usually not desirable since it results in matrices in the converted matrix equation that are twice or four times as large as the matrices in the original problem. On the other hand, structure-preserving Krylov subspace methods have become popular recently to overcome the increase in the size of the matrices when such a conversion is applied [19–21]. Some other approaches for solving linear quaternion matrix equations work on a right or a left Hilbert space over quaternions equipped with a proper inner product [17,18].

In this study, we consider (1.1) with non-Hermitian and nonsingular coefficient matrices A and B of size $n \times n$ and $m \times m$, respectively, when n and m are large. We aim to find the solution by means of a global Generalized Minimal Residual (GMRES) algorithm operating on the Krylov subspaces of the space of $n \times m$ quaternion matrices. We directly work with the original quaternion matrices without using real or complex representations of quaternion matrices and exploit a real inner product defined in the space of $n \times m$ quaternion matrices. Our approach applies the Sylvester operator $X \mapsto AX + XB$ at every iteration when adding a new direction to the Krylov subspace.

We present our study in the following order. In Section 2, the preliminaries for quaternion matrices, useful identities, and problem reformulation are presented. In Section 3, the global Arnoldi process to construct an orthonormal basis for a matrix Krylov subspace is described, and then the global GMRES method to retrieve the best approximation in this matrix Krylov subspace is presented. Finally, in Section 4, the efficiency and accuracy of the proposed approach are illustrated with examples related to image restoration.

2. Preliminaries

In this section, we summarize quaternions and some of their properties. The division ring $\mathbb{H}$ of quaternions is given by

$$\mathbb{H} = \{q_0 + q_1 i + q_2 j + q_3 k \mid i^2 = j^2 = k^2 = -1, \;\; ij = -ji = k, \;\; q_0, q_1, q_2, q_3 \in \mathbb{R}\}$$

For $q = q_0 + q_1 i + q_2 j + q_3 k \in \mathbb{H}$, the conjugate and the modulus of q are

$$\bar{q} = q_0 - q_1 i - q_2 j - q_3 k$$

and

$$|q| = \sqrt{q\bar{q}} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}$$

respectively.

Since the multiplication of the quaternion units i, j, and k are non-commutative, the multiplication of a quaternion scalar $p \in \mathbb{H}$ with another quaternion scalar $q \in \mathbb{H}$ is usually not commutative. Thus, the n-tuples of $\mathbb{H}$ denoted by $\mathbb{H}^n$ can be regarded either as a right vector space or a left vector space over the division ring $\mathbb{H}$, depending on whether the multiplication in $\mathbb{H}^n$ with quaternion scalars is defined from the right or from the left, respectively. In this study, we consider $\mathbb{H}^n$ together with the multiplication with scalars from the right, that is, $\mathbb{H}^n$ as a right vector space. A possible real inner product on $\mathbb{H}^n$ is

$$\langle u, v\rangle \ = \ \sum_{i=1}^{n} \mathrm{Re}\,(\overline{v_i}\, u_i) \tag{2.1}$$

for $u, v \in \mathbb{H}^n$. The right vector space $\mathbb{H}^n$ with (2.1) is commonly referred to as a right quaternionic Hilbert space. We define the norm of a vector $u \in \mathbb{H}^n$ in this right quaternionic Hilbert space as

$$\|u\| = \sqrt{\sum_{i=1}^{n} \mathrm{Re}\,(\overline{u_i}\, u_i)} = \sqrt{\sum_{i=1}^{n} \mid u_i \mid^2} \tag{2.2}$$

We denote the set of $n\times m$ matrices with quaternion entries with $\mathbb{H}^{n\times m}$, which can also be regarded as a right vector space over $\mathbb{H}$ together with the multiplication with quaternion scalars from the right. The basic linear algebra terminology, definitions, and standard notations for a vector space over complex numbers also apply to a right vector space over quaternions. In particular, the notations X^* and v^* are reserved for the conjugate transposes of $X \in \mathbb{H}^{n\times m}$ and $v \in \mathbb{H}^n$, respectively. Multiplication of two quaternion matrices of suitable sizes is defined analogously to the multiplication of two complex matrices. If a matrix $X \in \mathbb{H}^{n\times n}$ satisfies $X^*X = I$, then X is called a unitary matrix. On the other hand, if $X \in \mathbb{H}^{n\times n}$ satisfies $X^* = X$, then X is called Hermitian. Moreover, $X \in \mathbb{H}^{n\times n}$ is invertible if there exits $X^{-1} \in \mathbb{H}^{n\times n}$ such that $XX^{-1} = X^{-1}X = I$. In this work, we make use of the real inner product on $\mathbb{H}^{n\times m}$ defined as

$$\langle X, Y\rangle_F = \mathrm{Re}\,(\mathrm{tr}\,(Y^*X)) \tag{2.3}$$

for $X, Y \in \mathbb{H}^{n\times m}$. The norm of $X \in \mathbb{H}^{n\times m}$ induced by (2.3) is

$$\|X\|_F = \sqrt{\mathrm{Re}\,(\mathrm{tr}\,(X^*X))} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} \mid x_{ij} \mid^2} \tag{2.4}$$

Note that (2.4) is an extension of the Frobenius norm defined for complex matrices to quaternion matrices.

We provide a generalization of the definition of a block-partitioned matrix product introduced originally by Bouyouli et al. [35] to the setting of quaternion matrices equipped with (2.3).

Definition 2.1. Let $A = [A_1A_2\cdots A_p] \in \mathbb{H}^{n\times mp}$ and $B = [B_1B_2\cdots B_l] \in \mathbb{H}^{n\times ml}$ with $A_i, B_j \in \mathbb{H}^{n\times m}$ for $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, l\}$. Then, the $p \times l$ matrix $A^*\diamond B$ is defined by $(A^*\diamond B)_{ij} = \langle A_i, B_j\rangle_F$, for $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, l\}$.

If the product of $A^*\diamond A$ is equal to the $p \times p$ identity matrix, that is if

$$\langle A_i, A_j\rangle_F = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

for $i, j \in \{1, 2, \ldots, p\}$, then the matrix A is called *F-orthonormal*.

In the next subsection, we formally introduce our problem, as well as basic notions concerning the important ingredients of the problem, including the Sylvester operator.

2.1. Problem Reformulation

Any $A \in \mathbb{H}^{n\times m}$ can be uniquely expressed as

$$A = A_1 + A_2 j \quad \text{or} \quad A = \mathrm{Re}(A_1) + \mathrm{Im}(A_1)i + \mathrm{Re}(A_2)j + \mathrm{Im}(A_2)k$$

for some $A_1, A_2 \in \mathbb{C}^{n\times m}$. For a matrix $A \in \mathbb{H}^{n\times n}$, a scalar $\lambda \in \mathbb{H}$ is called a right eigenvalue of A if

$$Ax = x\lambda$$

holds for some nonzero $x \in \mathbb{H}^n$. We remark that if $\lambda \in \mathbb{H}$ is a non-real right eigenvalue, then $Axs = xs(s^{-1}\lambda s)$, for all nonzero $s \in \mathbb{H}$, therefore $s^{-1}\lambda s$ is also an eigenvalue of A. Hence, we refer to the set

$$\mathcal{E}_A(\lambda) \ := \ \left\{ s^{-1}\lambda s \ : \ s \in \mathbb{H}, \ s \neq 0 \right\}$$

as the equivalence class of $\lambda \in \mathbb{H}$. Consequently, if a quaternion matrix has a non-real eigenvalue, then it has infinitely many non-real eigenvalues. The equivalence class of a non-real eigenvalue has only one pair of complex conjugate scalars, i.e., $\mathcal{E}_A(\lambda) \cap \mathbb{C} = \left\{\lambda, \overline{\lambda}\right\}$. If the imaginary part of a right complex eigenvalue is nonnegative, it is called a standard eigenvalue. Any $n \times n$ quaternion matrix has exactly n standard eigenvalues counting the multiplicities [18, 36].

Consider the Sylvester matrix equation

$$AX + XB = C \tag{2.5}$$

such that $A \in \mathbb{H}^{n\times n}$, $B \in \mathbb{H}^{m\times m}$, and $C \in \mathbb{H}^{n\times m}$. It follows from Theorem 2.2.4.1 in [37] that (2.5) has a unique solution if and only if

$$\Lambda(A) \cap \Lambda(-B) = \emptyset$$

where $\Lambda(\cdot)$ denotes the set of standard eigenvalues of its quaternion matrix argument. In other words, (2.5) has a unique solution if and only if the quaternion matrices A and $-B$ do not have any common standard eigenvalue.

Associated with every Sylvester equation, there is a Sylvester operator. Formally, the Sylvester operator $S : \mathbb{H}^{n\times m} \to \mathbb{H}^{n\times m}$ for given $A \in \mathbb{H}^{n\times n}$ and $B \in \mathbb{H}^{m\times m}$ is defined as

$$S(X) = AX + XB \tag{2.6}$$

From (2.5) and (2.6),

$$S(X) = C \tag{2.7}$$

We define the norm of the operator S by

$$\|S\| = \max_{\|X\|_F=1} \|S(X)\|_F$$

where $\|\cdot\|_F$ is defined by (2.4). The adjoint of S is denoted by S^* and is given by

$$S^*(Y) = A^*Y + YB^*$$

for $Y \in \mathbb{H}^{n\times m}$. Given $X \in \mathbb{H}^{n\times m}$ and $Y \in \mathbb{H}^{n\times m}$, the equality $\langle S(X), Y\rangle_F = \langle X, S^*(Y)\rangle_F$ holds where the inner product $\langle\cdot,\cdot\rangle_F$ is defined as in (2.3).

In the rest of this paper, we focus on the solution of (2.5) by means of a global Krylov subspace method assuming that (2.5) has a unique solution. Our approach makes use of the associated Sylvester operator frequently. In the next section, we describe a global Arnoldi process to construct an orthonormal basis for a Krylov subspace, as well as a global GMRES method to find the best solution of (2.5) in a least-squares sense in an affine space associated with this Krylov subspace.

3. Solution of the Sylvester Quaternion Matrix Equation by Global GMRES

3.1. The Global Arnoldi Process

Suppose that $X_0 \in \mathbb{H}^{n\times m}$ is an approximate solution of (2.7), and $R_0 = C - S(X_0)$ is the corresponding residual. The quaternion matrix Krylov subspace $\mathcal{K}_k(S, R_0) \subset \mathbb{H}^{n\times m}$ associated with (2.6) and the residual R_0 that we will be dealing with is given by

$$\begin{aligned}\mathcal{K}_k(S, R_0) &:= \operatorname{span}\{R_0, S(R_0), \ldots, S^{k-1}(R_0)\}\\ &= \{\alpha_0 R_0 + \alpha_1 S(R_0) + \cdots + \alpha_{k-1} S^{k-1}(R_0) \mid \alpha_0, \alpha_1, \ldots, \alpha_{k-1} \in \mathbb{R}\}\end{aligned} \tag{3.1}$$

for a prescribed integer k. Note that in (3.1) the operator $S^i(R_0)$ is defined recursively by $S(S^{i-1}(R_0))$, for $i \in \{1, 2, \ldots, k-1\}$, and $S^0(R_0) = R_0$. We remark that the set $\mathbb{H}^{n\times m}$ over the field of real numbers is indeed a real vector space. Moreover, $\mathcal{K}_k(S, R_0)$, a subset of $\mathbb{H}^{n\times m}$, equipped with real scalars, is a real vector space as well, hence a subspace of $\mathbb{H}^{n\times m}$.

The global Arnoldi process, described formally in Algorithm 1, is a procedure that constructs an F-orthonormal basis for the Krylov subspace $\mathcal{K}_k(S, R_0)$. At termination, the process generates the set of matrices $\{Q_1, Q_2, \ldots, Q_k\}$ that forms an orthonormal basis for $\mathcal{K}_k(S, R_0)$ with respect to the inner product $\langle X, Y\rangle_F = \operatorname{Re}(\operatorname{tr}(Y^* X))$, for $X, Y \in \mathbb{H}^{n\times m}$.

Algorithm 1 Global Arnoldi Process

1: $R_0 \leftarrow C - S(X_0)$

2: Set $Q_1 = \frac{R_0}{\|R_0\|_F}$

3: **for** $j = 1$ to k **do**

4: $\quad V \leftarrow S(Q_j)$

5: $\quad$ **for** $i = 1$ to j **do**

6: $\quad\quad h_{ij} \leftarrow \langle Q_i, V\rangle_F$

7: $\quad\quad V \leftarrow V - Q_i h_{ij}$

8: $\quad$ **end for**

9: $\quad h_{(j+1)j} \leftarrow \|V\|_F$. If $h_{(j+1)j} = 0$, then stop.

10: $\quad Q_{j+1} \leftarrow \frac{V}{h_{(j+1)j}}$

11: **end for**

12: $\widetilde{Q}_k \leftarrow [Q_1 Q_2 \ldots Q_k]$, $\widetilde{Q}_{k+1} \leftarrow [\widetilde{Q}_k\, Q_{k+1}]$ and $\widetilde{H}_k$ is as in (3.4).

From the global Arnoldi process above, the recurrence for $j \in \{1, 2, \ldots, k\}$,

$$S(Q_j) \;=\; \sum_{i=1}^{j+1} Q_i h_{ij} \tag{3.2}$$

is immediate. Moreover, it can be verified in a straightforward manner that (3.2) above yields the relation

$$S(\widetilde{Q}_k) \;=\; \widetilde{Q}_{k+1}(\widetilde{H}_k \otimes I) \tag{3.3}$$

where I is the $m \times m$ identity matrix and $\widetilde{H}_k \in \mathbb{R}^{(k+1)\times k}$ is the Hessenberg matrix whose the entry (i, j) is h_{ij} produced by the global Arnoldi process,

$$\widetilde{H}_k = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1(k-1)} & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2(k-1)} & h_{2k} \\ 0 & h_{32} & \cdots & h_{3(k-1)} & h_{3k} \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & h_{k(k-1)} & h_{kk} \\ 0 & 0 & \cdots & 0 & h_{(k+1)k} \end{bmatrix} \tag{3.4}$$

Here and throughout the rest of this paper, $F \otimes G$ represents the Kronecker product of the real matrices F and G. In the next subsection, we present the global GMRES method for retrieving the solution of (2.7) by making use of $\widetilde{Q}_k$ and $\widetilde{H}_k$ generated by the global Arnoldi process.

3.2. The Global GMRES Method

For a given initial estimate $X_0 \in \mathbb{H}^{n\times m}$ for the solution of (2.7), our global GMRES method at the kth iteration finds X_k minimizing $\|C - S(X)\|_F$ over all $X \in X_0 + \mathcal{K}_k(S, R_0)$. For all $X_k \in X_0 + \mathcal{K}_k(S, R_0)$ can be expressed as

$$X_k = X_0 + \sum_{i=1}^{k} y_i^{(k)} Q_i$$

for some real scalars $y_i^{(k)}$ for $i \in \{1, \ldots, k\}$, or equivalently

$$X_k = X_0 + \widetilde{Q}_k Y_k$$

for $Y_k = y^{(k)} \otimes I \in \mathbb{R}^{km\times m}$ where $y^{(k)} := \left[y_1^{(k)} y_2^{(k)} \ldots y_k^{(k)}\right]^T$. Thus, recalling $R_0 = C - S(X_0)$, the residual $R_k = C - S(X_k)$ can be written as

$$R_k = R_0 - S(\widetilde{Q}_k Y_k) \tag{3.5}$$

The minimization of $\|C - S(X)\|_F$ overall $X \in X_0 + \mathcal{K}_k(S, R_0)$ is equivalent to the minimization of $\|R_k\|_F$ with R_k of (3.5) and $Y_k = y^{(k)} \otimes I$, for some $y^{(k)} \in \mathbb{R}^k$. In other words, we would like to solve the following minimization problem over $y^{(k)} \in \mathbb{R}^k$:

$$\left\| R_0 - S\left(\widetilde{Q}_k (y^{(k)} \otimes I)\right) \right\|_F = minimum$$

Using the linearity of the operator S and (3.3), the last minimization can be rewritten as

$$\left\| R_0 - \widetilde{Q}_{k+1}(\widetilde{H}_k \otimes I)(y^{(k)} \otimes I) \right\|_F = minimum \tag{3.6}$$

It follows from the description in Algorithm 1 that $R_0 = \beta Q_1$ for $\beta := \|R_0\|_F$, or equivalently $R_0 = \widetilde{Q}_{k+1}(\beta e_1 \otimes I)$ where e_1 is the first column of the $(k+1) \times (k+1)$ identity matrix. Hence, (3.6) can further be simplified as

$$\left\| \widetilde{Q}_{k+1} \left((\beta e_1 - \widetilde{H}_k y^{(k)}) \otimes I \right) \right\|_F = \left\| (\beta e_1 - \widetilde{H}_k y^{(k)}) \otimes I \right\|_F = minimum \tag{3.7}$$

The first equality in (3.7) follows from the fact that $\widetilde{Q}_{k+1} = [Q_1 Q_2 \cdots \ Q_{k+1}]$ where $\{Q_1, Q_2, \cdots, Q_{k+1}\}$ is an orthonormal set with respect to the inner product $\langle \cdot, \cdot \rangle_F$, inducing the Frobenius norm $\| \cdot \|_F$. As a result, our least squares problem reduces to finding $y^{(k)}$ such that

$$\left\| \beta e_1 - \widetilde{H}_k y^{(k)} \right\| = minimum \tag{3.8}$$

where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^{k+1}$. At iteration k of the global GMRES method, we solve this real least-squares problem over the variable $y^{(k)} \in \mathbb{R}^k$.

A typical approach to solve (3.8) efficiently is triangularizing $\widetilde{H}_k$ unitarily. Specifically, we transform the Hessenberg matrix $\widetilde{H}_k$ into an upper triangular matrix $\widetilde{U}$ by applying k square unitary matrices $W_1, W_2, \ldots, W_k$ from left, that is

$$W_k W_{k-1} \ldots W_1 \widetilde{H}_k = \widetilde{U} = \begin{bmatrix} \widetilde{u}_{11} & \widetilde{u}_{12} & \cdots & \widetilde{u}_{1k} \\ 0 & \widetilde{u}_{22} & \ldots & \widetilde{u}_{2k} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \cdots & \widetilde{u}_{kk} \\ 0 & 0 & \ldots & 0 \end{bmatrix}$$

where $W_j \in \mathbb{R}^{(k+1)\times(k+1)}$, for $j \in \{1, 2, \ldots, k\}$, is given by

$$W_j = \begin{bmatrix} I_{j-1} & 0 & 0 \\ 0 & P_j & 0 \\ 0 & 0 & I_{k-j} \end{bmatrix} \tag{3.9}$$

for a Givens rotator $P_j \in \mathbb{R}^{2\times 2}$ and I_ℓ denoting the identity matrix of size $\ell \times \ell$. For completeness, an efficient realization of these ideas to turn the Hessenberg matrix $\widetilde{H}_k$ into an upper triangular form $\widetilde{U}$ is given in Algorithm 2.

Algorithm 2 Triangularization of the Hessenberg Matrix $\widetilde{H}_k$

for $j = 1$ to k **do**
  $y^{(j)} \leftarrow \widetilde{H}_k(j, j+1 : j)$
  $\widetilde{y} \leftarrow \frac{y^{(j)}}{\|y^{(j)}\|}$
  $u \leftarrow \widetilde{H}_k(j+1, j : k)$
  $\widetilde{H}_k(j+1, j+1 : k) \leftarrow -\widetilde{y}_2 \widetilde{H}_k(j, j+1 : k) + \widetilde{y}_1 \widetilde{H}_k(j+1, j+1 : k))$
  $\widetilde{H}_k(j+1, j) \leftarrow 0$
  $\widetilde{H}_k(j, j : k) \leftarrow \widetilde{y}_1 \widetilde{H}_k(j, j : k) + \widetilde{y}_2 u$
end for
$\widetilde{U} \leftarrow \widetilde{H}_k$

In the description in Algorithm 2, the notation $\widetilde{H}_k(\ell, \ell_1 : \ell_2)$ is reserved for the row vector formed of the entries of $\widetilde{H}_k$ on the ℓth row with column indices from ℓ_1 to ℓ_2. Moreover, the unitary matrices $W_1, W_2, \dots, W_k$ triangularizing $\widetilde{H}_k$ can be formed using the vectors $y^{(1)}, y^{(2)}, \dots, y^{(k)}$ generated by Algorithm 2. Specifically, W_j is as in (3.9) with the Givens rotator P_j defined as

$$P_j = \frac{1}{\|y^{(j)}\|}\begin{bmatrix} y_1^{(j)} & y_2^{(j)} \\ -y_2^{(j)} & y_1^{(j)} \end{bmatrix}$$

for $j \in \{1, 2, \dots, k\}$.

Once $\widetilde{H}_k$ is triangularized into $\widetilde{U}$, (3.8) can be solved efficiently. In particular, as (2.2) is invariant under unitary transformations, (3.8) can equivalently be expressed as

$$\left\| \begin{bmatrix} \widehat{u}_1 \\ \widehat{u}_2 \\ \vdots \\ \widehat{u}_k \\ \widehat{u}_{k+1} \end{bmatrix} - \begin{bmatrix} \widetilde{u}_{11} & \widetilde{u}_{12} & \cdots & \widetilde{u}_{1k} \\ 0 & \widetilde{u}_{22} & \dots & \widetilde{u}_{2k} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \dots & \widetilde{u}_{kk} \\ 0 & 0 & \dots & 0 \end{bmatrix} y^{(k)} \right\| = minimum$$

where

$$[\widehat{u}_1 \; \widehat{u}_2 \; \dots \; \widehat{u}_{k+1}]^T := W_k W_{k-1} \dots W_1(\beta e_1)$$

can be obtained by applying the rotators $P_1, P_2, \dots, P_k$ in this order to βe_1. It follows that the solution $y_*^{(k)}$ of (3.8) is the solution of the upper triangular system

$$\begin{bmatrix} \widetilde{u}_{11} & \widetilde{u}_{12} & \dots & \widetilde{u}_{1k} \\ 0 & \widetilde{u}_{22} & \dots & \widetilde{u}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{u}_{kk} \end{bmatrix} y \;\; = \;\; \begin{bmatrix} \widehat{u}_1 \\ \widehat{u}_2 \\ \vdots \\ \widehat{u}_k \end{bmatrix} \tag{3.10}$$

and can be retrieved by back substitution.

Once we have $y_*^{(k)}$ at hand, the best approximate solution X_k in $X_0 + \mathcal{K}_k(S, R_0)$ for $AX + XB = C$ that is X_k minimizing $\|C - (AX + XB)\|_F$ over all $X \in X_0 + \mathcal{K}_k(S, R_0))$, is given by $X_k = X_0 + \widetilde{Q}_k(y_*^{(k)} \otimes I)$. An outline of the overall global GMRES method is provided in Algorithm 3.

Algorithm 3 The Global GMRES Method to Solve the Sylvester Quaternion Matrix Equation

1: **Apply Algorithm 1.**
In particular, form the matrix $\widetilde{Q}_k = [Q_1 Q_2 \cdots Q_k]$ such that $\{Q_1, Q_2, \dots Q_k\}$ forms an orthonormal basis for $\mathcal{K}_k(A, R_0)$, as well as the Hessenberg matrix $\widetilde{H}_k$ as in (3.4) satisfying (3.3).

2: **Use Algorithm 2 to Triangularize $\widetilde{H}_k$.**
Specifically, unitarily transform $\widetilde{H}_k$ into the upper triangular matrix $\widetilde{U}$, and keep also the rotation vectors $y^{(1)}, y^{(2)}, \dots, y^{(k)}$ that define the unitary transformation.

3: **Apply Unitary Transformation from Step 2 to βe_1.**
Apply the unitary transformation from the previous step to βe_1 by making use of $y^{(1)}, y^{(2)}, \dots, y^{(k)}$ to obtain $[\widehat{u}_1 \; \widehat{u}_2 \; \dots \widehat{u}_{k+1}]^T$.

4: **Find the Solution $y_*^{(k)}$ of the Upper Triangular System in (3.10).**

5: **Form $X_k = X_0 + \widetilde{Q}_k(y_*^{(k)} \otimes I)$, Which is the Best Approximate Solution of $AX + XB = C$ in $X_0 + \mathcal{K}_k(S, R_0)$.**

4. Numerical Examples

In this section, we demonstrate the effectiveness of the proposed global GMRES approach for (2.5) by conducting numerical experiments on examples related to color image restoration.

A color image can be encoded as an $n \times m$ quaternion matrix of the form

$$Q = Ri + Gj + Bk$$

where R, G and B are $n \times m$ real matrices represent the color image's red, green, and blue components. Let A and B be the blurring quaternion matrices of the form

$$A = A_1 + A_2 i + A_3 j + A_4 k \quad \text{and} \quad B = B_1 + B_2 i + B_3 j + B_4 k$$

for some real matrices A_j and B_j such that $j \in \{1, 2, 3, 4\}$. The constant parts A_1 and B_1 of blurring matrices A and B are specified as

$$a_{ij}, d_{ij} = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(i-j)^2}{2\sigma^2}}, & |i-j| \leq r \\ 0, & \text{otherwise} \end{cases}$$

whereas the non-constant parts, i.e., the coefficient matrices for i, j, and k parts, of A and B are given by

$$a_{ij}, d_{ij} = \begin{cases} \frac{1}{10} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(i-j)^2}{2\sigma^2}}, & |i-j| \leq r \\ 0, & \text{otherwise} \end{cases}$$

for prescribed positive real number r and σ. By applying the Sylvester operator we obtain $AX + XB = C$ where C is the quaternion matrix corresponding to the blurred image. On the other hand, given A, B, and C, the solution X to the Sylvester equation $AX + XB = C$ is the quaternion matrix corresponding to the original color image that we would like to restore back.a as as a sas a sa

Example 4.1. We report results illustrating the effectiveness of our proposed approach on such Sylvester equations obtained from the two original color images depicted in Figures 1(a) and 2(a). The images are stored as $n \times m$ quaternion matrices, with sizes 583×500 and 500×752, respectively. We set $\sigma = 10$ and $r = 10$ in both of the examples, and the resulting blurred images C by the application of the Sylvester operator are shown in Figures 1(b) and 2(b). We apply Algorithm 3 to solve $AX + XB = C$ approximately by setting the number of iterations equal to $k = 2$, $k = 4$, $k = 10$, and $k = 50$. The restored images corresponding to the approximate solutions after so many iterations are shown in 1(c)-(f) and 2(c)-(f). Finally, the convergence of the algorithm is illustrated in Figure 3 by plotting the residual norms as a function of number of iterations k. To be precise, the plot on the top in Figure 3 depicts the residual norm $\|R_k\|_F$ for the approximate solution X_k by Algorithm 3 as a function of the number of iterations k for the parrot example. The plot at the bottom does the same for the tiger example.

(a) Original image **(b)** Degraded image **(c)** Restored image for k=2 **(d)** Restored image for k=4 **(e)** Restored image for k=10 **(f)** Restored image for k=50

Figure 1. This concerns the parrot example. Original and blurred images are illustrated in (a) and (b), respectively. The restored images obtained by applications of Algorithm 3 with $k = 2$, $k = 4$, $k = 10$, and $k = 50$ iterations are depicted in (c)-(f)

(a) Original image **(b)** Degraded image **(c)** Restored image for k=2 **(d)** Restored image for k=4 **(e)** Restored image for k=10 **(f)** Restored image for k=50

Figure 2. The images are similar to those in Figure 1 but now for the tiger example. In particular, (a) and (b) are the original and blurred images, whereas (c)-(f) are restored images retrieved by applying Algorithm 3 with k iterations

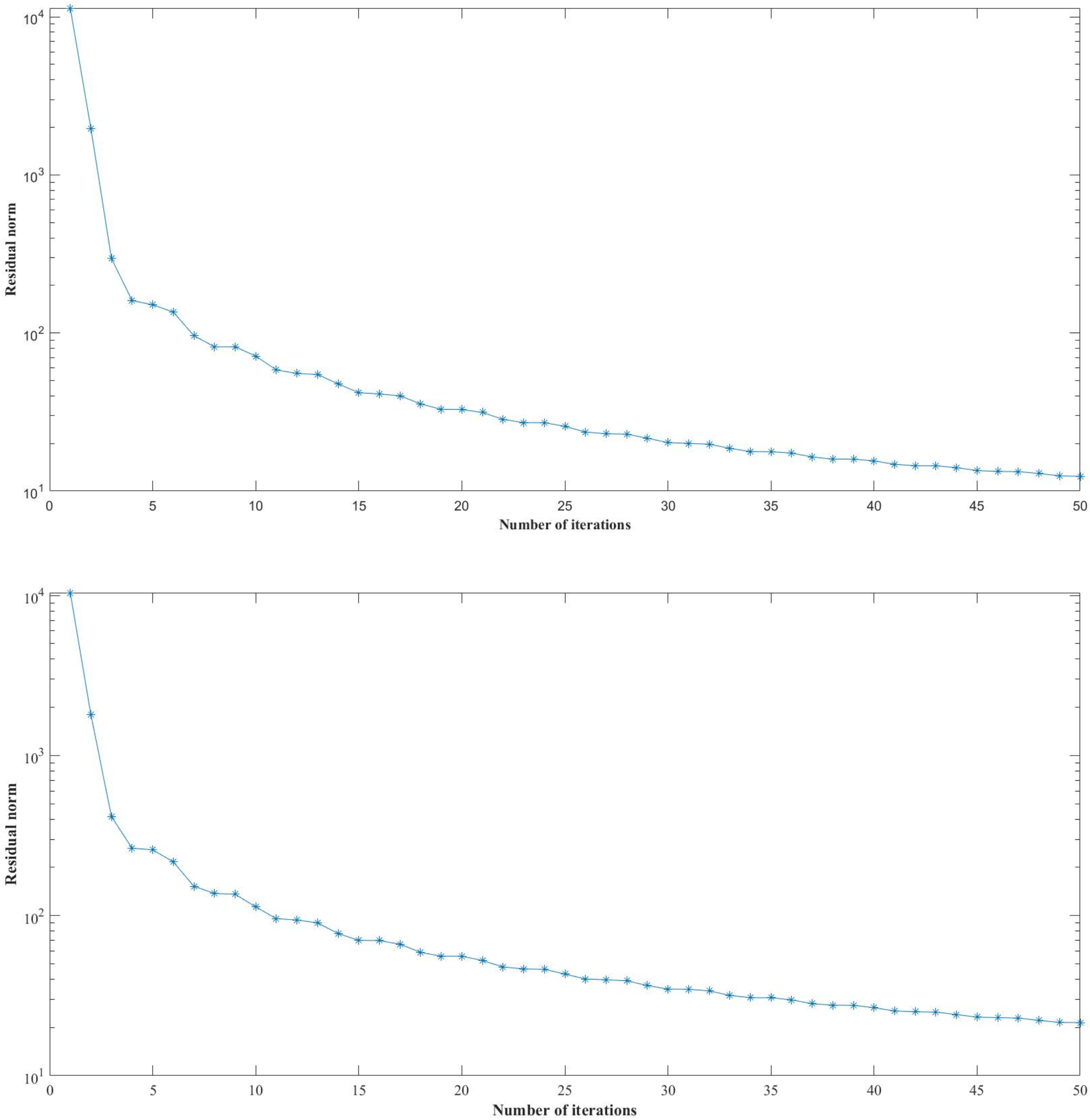

Figure 3. The residual norms for the approximate solutions by Algorithm 3 are plotted as a function of number of iterations k. The top and bottom plots concern the parrot and tiger examples, respectively

5. Conclusion

We have proposed an iterative algorithm for solving a Sylvester quaternion matrix equation, especially the large-scale setting when at least one of the coefficient matrices is large. The proposed algorithm is a global GMRES method operating on a Krylov subspace of a vector space of quaternion matrices. An Arnoldi process based on repeated applications of the Sylvester operator is presented to construct an orthonormal basis for this Krylov subspace. We have also discussed how the determination of the best solution to the Sylvester equation in this Krylov subspace minimizing the Frobenius norm of the residual can be converted into a standard least-squares problem in $\mathbb{R}^n$, which in turn paves the way for efficient computation of the best solution. Finally, we have illustrated on numerical examples that the proposed approach works effectively on Sylvester equations that need to be solved to retrieve the originals of degraded color images. A natural extension of the approach introduced here that could be considered as future work is a conjugate gradient method for solving Lyapunov quaternion matrix equations.

Author Contributions

The author read and approved the final version of the paper.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] J. J. Sylvester, *Sur l'équation en matrices $px = xq$*, Comptes Rendus de l'Académie des Sciences Paris 99 (2) (1884) 67–71, 115–116.

[2] A. M. Lypunov, *The general problem of the stability of motion*, International Journal of Control 55 (3) (1992) 531–534.

[3] E. B. Castelan, V. G. Silva, *On the solution of a Sylvester equation appearing in descriptor systems control theory*, Systems Control Letters 54 (2) (2005) 109–117.

[4] G. R. Duan, *Eigenstructure assignment in descriptor systems via output feedback – A new complete parametric approach*, International Journal of Systems Science 72 (4) (1999) 345–364.

[5] U. Baur, P. Benner, *Cross-Gramian based model reduction for data-sparse systems*, Electronic Transactions on Numerical Analysis 31 (2008) 256–270.

[6] D. Calvetti, L. Reichel, *Application of ADI iterative methods to the restoration of noisy images*, Society for Industrial and Applied Mathematics Journal on Matrix Analysis and Applications 17 (1) (1995) 165–186.

[7] A. Bouhamidi, K. Jbilou, *Sylvester Tikhonov-regularization methods in image restoration*, Journal of Computational and Applied Mathematics 206 (1) (2007) 86–98.

[8] M. Epton, *Methods for the solution of $AXD - BXC = E$ and its application in the numerical solution of implicit ordinary differential equations*, BIT Numerical Mathematics 20 (3) (1980) 341–345.

[9] V. Simoncini, *Computational methods for linear matrix equations*, Society for Industrial and Applied Mathematics Review 58 (3) (2006) 377–441.

[10] L. Rodman, Topics in quaternion linear algebra, Princeton University Press, Princeton, 2014.

[11] M. Wei, Y. Li, F. Zhang, J. Zhao, Quaternion matrix computations, Nova Science Publishers, New York, 2018.

[12] S. Şimşek, *Least–squares solutions of generalized Sylvester–type quaternion matrix equations*, Advances in Applied Clifford Algebras 33 (3) (2023) Article Number 28 23 pages.

[13] S. Yuan, Q. Wang, *L-structured quaternion matrices and quaternion linear matrix equations*, Linear and Multilinear Algebra 64 (2) (2016) 321–339.

[14] I. Kyrchei, *Cramer's rules for Sylvester quaternion matrix equation and its special cases*, Advances in Applied Clifford Algebras 28 (5) (2018) Article Number 90 26 pages.

[15] I. Kyrchei, *Cramer's Rules of η-Skew-hermitian solutions to the quaternion Sylvester-type matrix equations*, Advances in Applied Clifford Algebras 29 (3) (2019) Article Number 56 31 pages.

[16] S. Şimşek, M. Sarduvan, H. Özdemir, *Centrohermitian and skew-centrohermitian solutions to the minimum residual and matrix nearness problems of the quaternion matrix equation* $(AXB, DXE) = (C, F)$, Advances in Applied Clifford Algebras 27 (3) (2017) 2201–2214.

[17] F. P. A. Beik, S. Ahmadl-Asl, *An iterative algorithm for η-anti-hermitian least squares solutions of quaternion matrix equations*, Electronic Journal of Linear Algebra 30 (2015) 372–401.

[18] S. Şimşek, *A block quaternion GMRES method and its convergence analysis*, Calcolo 61 (2) (2024) Article Number 33 27 pages.

[19] S. Şimşek, A. Körükçü, *A block conjugate gradient method for quaternion linear systems*, Yuzuncu Yil University Journal of the Institute of Natural and Applied Sciences 28 (2) (2023) 394–403.

[20] Z. Jia, M. K. Ng, *Structure preserving quaternion generalized minimal residual method*, Society for Industrial and Applied Mathematics Journal of Matrix Analysis and Applications 42 (2) (2021) 616–634.

[21] T. Li, Q. W. Wang, *Structure preserving quaternion full orthogonalization method with applications*, Numerical Linear Algebra with Applications 30 (5) (2023) e2495 15 pages.

[22] R. H. Bartels, G. W. Stewart, *Algorithm 432: Solution of the matrix equation $AX + XB = C$*, Communications of the ACM 15 (9) (1972) 820–826.

[23] G. H. Golub, S. Nash, C. Van Loan, *A Hessenberg-Schur method for the problem $AX + XB = C$*, IEEE Transactions on Automatic Control 24 (6) (1979) 909–913.

[24] K. Jbilou, A. Messaoudi, H. Sadok, *Global FOM and GMRES algorithms for matrix equations*, Applied Numerical Mathematics 31 (1) (1999) 49–63.

[25] A. El Guennouni, K. Jbilou, A. J. Riquet, *Block Krylov subspace methods for solving large Sylvester equations*, Numerical Algorithms 29 (1-3) (2002) 75–96.

[26] M. Heyouni, K. Jbilou, *Matrix Krylov subspace methods for large scale model reduction problems*, Applied Mathematics and Computation 181 (2) (2006) 1215–1228.

[27] K. Jbilou, A. J. Riquet, *Projection methods for large Lyapunov matrix equations*, Linear Algebra and its Applications 415 (2-3) (2006) 344–358.

[28] A. Bouhamidi, K. Jbilou, *A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications*, Applied Mathematics and Computation 206 (2) (2008) 687–694.

[29] J. Zhang, H. Dai, J. Zhao, *A new family of global methods for linear systems with multiple right-hand sides*, Journal of Computational and Applied Mathematics 236 (6) (2011) 1562–1575.

[30] F. P. A. Beik, D. K. Salkuyeh, *On the global Krylov subspace methods for solving general coupled matrix equations*, Computers and Mathematics with Applications 62 (12) (2011) 4605–4613.

[31] A. Kaabi, *On the numerical solution of generalized Sylvester matrix equations*, Bulletin of the Iranian Mathematical Society 40 (1) (2014) 101–113.

[32] F. P. A. Beik, *Theoretical results on the global GMRES method for solving generalized Sylvester matrix equations*, Bulletin of the Iranian Mathematical Society 40 (5) (2014) 1097–1117.

[33] S. K. Li, T. Z. Huang, *Global FOM and GMRES algorithms for a class complex matrix equations*, Journal of Computational and Applied Mathematics 335 (2018) 227–241.

[34] S. K. Li, M. X. Wang, G. Liu, *A global variant of the COCR method for the complex symmetric Sylvester matrix equation* $AX+XB=C$, Computers and Mathematics with Applications 94 (2021) 104–113.

[35] A. Bouhamidi, K. Jbilou, R. Sadaka, H. Sadok, *Convergence properties of some block Krylov subspace methods for multiple linear systems*, Journal of Computational and Applied Mathematics 196 (2) (2006) 498–511.

[36] D. R. Farenick, B. A. F. Pidkowich, *The spectral theorem in quaternions*, Linear Algebra and Its Applications 371 (2003) 75–102.

[37] R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.