Çankaya University Journal of Science and Engineering
Volume 10 (2013), No 2, 153–167.

A Convergence Analysis of Extended Global FOM and Extended Global GMRES For Matrix Equations *AXB* = *F*

A. Rivaz[1,*], A. Tajaddini[1], A. Salarmohammadi[1]

[1]*Department of Mathematics, Faculty of Mathematics & Computer, Shahid Bahonar University of Kerman, 76169-1411, Kerman, Iran*
**Corresponding author: arivaz@uk.ac.ir*

Abstract. In this paper, we present some convergence results of the extended global full orthogonalization and the extended global generalized minimal residual methods. We also present new expressions of the approximate solutions and the corresponding residuals.

Keywords. *Extended block Krylov subspace; Extended global FOM; Extended global GMRES*.

1. Introduction

In the present work, we consider the matrix equation of the form

$$AXB = F \tag{1.1}$$

Where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{s\times s}$ are nonsingular matrices.

Recently there has been an increased interest in solving matrix equations; for details see [4,6] and references therein. For matrix equation $AXB = F$, many iterative methods have been presented in the last few years [2,3,4,5,6]. Among them are the extended block Krylov subspace methods such as the extended global generalized minimal residual (EGl-GMRES) method [9]. The aim of this paper is to present some convergence results for extended block Krylov subspace methods that include the

ISSN 1309 – 6788

extended global full orthogonalization method (EGl-FOM) and the extended global generalized minimal residual method (EGl-GMRES). Here, we exploit the structure of the block Krylov matrix only and we ignore the algorithm that implements the method. Using some properties of the Schur complement and applying the matrix product, we provide expressions for the approximate and corresponding residual. These results will be used to derive convergence properties for the extended global full orthogonalization method (EGl-FOM) and the extended global generalized minimal residual method (EGl-GMRES).

2. Definitions and Preliminaries

The innerproduct $\langle .,. \rangle_F$ for the matrices X and Y in $\mathbb{R}^{n\times s}$ is defined as $\langle X,Y \rangle_F = tr(X^T Y)$ and the corresponding matrix norm is the well-known Frobenius norm. A system of vectors (matrices) of $\mathbb{R}^{n\times s}$ is said to be F-orthonormal if it is orthonormal with respect to $\langle .,. \rangle_F$. The vector $vec(X)$ denotes the vector of $\mathbb{R}^{ns}$ (the set of $ns-$ dimensional real vectors) obtained by stacking the columns of the $n\times s$ matrix X. For two matrices, A and B, of dimensions $n\times p$ and $q\times l$ respectively, the Kronecker product $A\otimes B$ is the $nq\times pl$ matrix defined by $A\otimes B = [a_{ij}B]$. The following are the properties for this product [7]:

(1) $vec(ABC) = (C^T\otimes A)vec(B)$.

(2) $(A\otimes B)(C\otimes D) = (AC\otimes BD)$.

(3) If A and B are nonsingular matrices of dimension $n\times n$ and $p\times p$, respectively, then $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$.

(4) If A and B are $n\times n$ and $p\times p$ matrices, then
$det(A\otimes B) = det(A)^p det(B)^n$ and $tr(A\otimes B) = tr(A)tr(B)$.

(5) $vec(A)^T vec(B) = tr(A^T B)$.

Definition 1 ([1]). Let $A = [A_1, A_2, \dots, A_p]$ and $B = [B_1, B_2, \dots, B_l]$ be matrices of dimension $n\times ps$ and $n\times ls$, respectively, where A_i and B_j $(i = 1,\dots,p; j = 1,\dots,l)$ are $n\times s$ matrices. Then the $p\times l$ matrix $A^T \Diamond B$ is defined by $A^T \Diamond B = [C_{ij}]$ where $C_{ij} = \langle A_i, B_j \rangle_F$.

Remarks ([1]).

(1) If $s = 1$ then $A^T \lozenge B = A^T B$.

(2) If $s = 1$, $p = 1$ and $l = 1$, then setting $A = u \in \mathbb{R}^n$ and $B = v \in \mathbb{R}^n$, we have $A^T \lozenge B = u^T v \in \mathbb{R}$.

(3) The matrix $A = [A_1, A_2, \dots, A_p]$ is F-orthonormal if and only if $A^T \lozenge A = I_p$.

(4) If $X \in \mathbb{R}^{n \times s}$, then $X^T \lozenge X = \|X\|_F^2$.

Proposition 1 (The relations (1), (2), (3), (4), (5), (6) and (7) are taken from reference [1]). Let $A, B, C \in \mathbb{R}^{n \times ps}$, $D \in \mathbb{R}^{n \times n}$, $L \in \mathbb{R}^{p \times s}$, $E \in \mathbb{R}^{s \times p}$ and $\alpha \in \mathbb{R}$. Then we have

(1) $(A + B)^T \lozenge C = A^T \lozenge C + B^T \lozenge C$.

(2) $A^T \lozenge (B + C) = A^T \lozenge B + A^T \lozenge C$.

(3) $(\alpha A)^T \lozenge C = \alpha(A^T \lozenge C)$.

(4) $(A^T \lozenge B)^T = B^T \lozenge A$.

(5) $(DA)^T \lozenge B = A^T \lozenge (D^T B)$.

(6) $A^T \lozenge \left(B(L \otimes I_s)\right) = (A^T \lozenge B)L$.

(7) $\|(A^T \lozenge B)\|_F \leq \|A\|_F \|B\|_F$.

(8) $A^T \lozenge \left(B(L \otimes E)\right) = \left(A^T \lozenge \left(B\left(I_p \otimes E\right)\right)\right) L$.

Proof. We can, for simplicity, show that the properties (1), (2), (3), (4), (5), (6) and (7) are satisfied by the product $\lozenge$. In the following, we show relation (8) as follows

$$A^T \lozenge \left(B(L \otimes E)\right) = \begin{pmatrix} \langle A_1, \sum_{i=1}^{p} B_i l_{i,1} E \rangle_F & \cdots & \langle A_1, \sum_{i=1}^{p} B_i l_{i,s} E \rangle_F \\ \vdots & \ddots & \vdots \\ \langle A_p, \sum_{i=1}^{p} B_i l_{i,1} E \rangle_F & \cdots & \langle A_p, \sum_{i=1}^{p} B_i l_{i,s} E \rangle_F \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{p} \langle A_1, B_i l_{i,1} E \rangle_F & \cdots & \sum_{i=1}^{p} \langle A_1, B_i l_{i,s} E \rangle_F \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{p} \langle A_p, B_i l_{i,1} E \rangle_F & \cdots & \sum_{i=1}^{p} \langle A_p, B_i l_{i,s} E \rangle_F \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{p} l_{i,1}\langle A_1, B_iE\rangle_F & \cdots & \sum_{i=1}^{p} l_{i,s}\langle A_1, B_iE\rangle_F \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{p} l_{i,1}\langle A_p, B_iE\rangle_F & \cdots & \sum_{i=1}^{p} l_{i,s}\langle A_p, B_iE\rangle_F \end{pmatrix}$$

$$= \begin{pmatrix} \langle A_1, B_1E\rangle_F & \cdots & \langle A_1, B_pE\rangle_F \\ \vdots & \ddots & \vdots \\ \langle A_p, B_1E\rangle_F & \cdots & \langle A_p, B_pE\rangle_F \end{pmatrix} \begin{pmatrix} l_{1,1} & \cdots & l_{1,s} \\ \vdots & \ddots & \vdots \\ l_{p,1} & \cdots & l_{p,s} \end{pmatrix}$$

$$= \left(A^T \lozenge \left[B_1E, \ldots, B_pE\right]\right)L = \left(A^T \lozenge \left(B\left(I_p \otimes E\right)\right)\right)L. \quad \square$$

Definition 2 ([1]). Let M be a matrix partitioned into four blocks

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

Where the submatrix D is assumed be square and nonsingular. The Schur complement of D in M, denoted by (M/D), is defined by

$$(M/D) = A - BD^{-1}C.$$

Proposition 2 ([8]). Assuming that the matrix D is nonsingular, then

$$\left(\begin{pmatrix} A & B \\ C & D \end{pmatrix}/D\right) = \left(\begin{pmatrix} D & C \\ B & A \end{pmatrix}/D\right) = \left(\begin{pmatrix} B & A \\ D & C \end{pmatrix}/D\right) = \left(\begin{pmatrix} C & D \\ A & B \end{pmatrix}/D`\right).$$

Proposition 3 ([1]). Let $A \in \mathbb{R}^{n\times ps}, B \in \mathbb{R}^{n\times ks}, C \in \mathbb{R}^{k\times p}, D \in \mathbb{R}^{k\times k}$ and $E \in \mathbb{R}^{n\times s}$. If the matrix D is nonsingular, then

$$E^T \lozenge \left(\begin{pmatrix} A & B \\ C\otimes I_s & D\otimes I_s \end{pmatrix}/D\otimes I_s\right) = \left(\begin{pmatrix} E^T \lozenge A & E^T \lozenge B \\ C & D \end{pmatrix}/D\right).$$

3. Extended Global FOM and Extended Global GMRES Methods

3.1. An Extended Global FOM Method

Let $\mathcal{G}K_m = \mathcal{G}K_m(A,V,B) = span\{V, AVB, A^2VB^2, \dots, A^{m-1}VB^{m-1}\}$ denote the extended block Krylov subspace of $\mathbb{R}^{n\times s}$, where $V \in \mathbb{R}^{n\times s}$. This subspace is defined in [9].

Note that $Z \in \mathcal{G}K_m(A,V,B)$ means that

$$Z = \sum_{i=1}^{m} \alpha_i A^{i-1} V B^{i-1}, \qquad \alpha_i \in \mathbb{R}, \qquad i = 1, \dots, m.$$

Now consider the matrix equation (1.1) and let X_0 be an initial $n \times s$ matrix with the corresponding residual $R_0 = B - AX_0$. At step m an EGL-FOM constructs the new approximation X_m such that

$$X_m^{EGl-FOM} = X_0 + Z_m, \qquad Z_m \in \mathcal{G}K_m(A, R_0, B) \tag{3.1}$$

and

$$R_m^{EGl-FOM} = R_0 - AZ_mB \perp_F \mathcal{G}K_m(A, R_0, B), \tag{3.2}$$

where the notation $\perp_F$ means the orthogonality with respect to $\langle .,. \rangle_F$.

The relation (3.1) implies

$$X_m^{EGl-FOM} = X_0 + [R_0, AR_0B, A^2R_0B^2, \dots, A^{m-1}R_0B^{m-1}](y \otimes I_s),$$

where $y = [y_1, y_2, \dots, y_m]^T$. Then the residual R_m can be expressed as

$$R_m^{EGl-FOM} = R_0 - [AR_0B, A^2R_0B^2, \dots, A^mR_0B^m](y \otimes I_s) \tag{3.3}$$

We can obtain the parameters y_i, $i = 1, \dots, m$, from the orthogonality condition (3.2), which is equivalent to

$$\langle R_m^{EGl-FOM}, A^iR_0B^i \rangle_F = 0, \quad i = 0, \dots, m-1 \tag{3.4}$$

Let $\mathcal{K}_m = [R_0, AR_0B, A^2R_0B^2, \dots, A^{m-1}R_0B^{m-1}]$ and $\mathcal{W}_m = A\mathcal{K}_m(I_m \otimes B)$. Then from (3.3) and (3.4) we have

$$(\mathcal{K}_m^T \lozenge \mathcal{W}_m)y = \mathcal{K}_m^T \lozenge R_0 \tag{3.5}$$

Theorem 1. Suppose that the matrix $(\mathcal{K}_m^T \lozenge \mathcal{W}_m)$ is nonsingular. Then the approximate solution X_m and the corresponding residual R_m can be expressed as the following Schur complements

$$X_m^{EGl-FOM} = \left(\begin{pmatrix} X_0 & -\mathcal{K}_m \\ (\mathcal{K}_m^T \lozenge R_0)\otimes I_s & (\mathcal{K}_m^T \lozenge \mathcal{W}_m)\otimes I_s \end{pmatrix} / (\mathcal{K}_m^T \lozenge \mathcal{W}_m)\otimes I_s\right) \tag{3.6}$$

and

$$R_m^{EGl-FOM} = \left(\begin{pmatrix} R_0 & \mathcal{W}_m \\ (\mathcal{K}_m^T \lozenge R_0)\otimes I_s & (\mathcal{K}_m^T \lozenge \mathcal{W}_m)\otimes I_s \end{pmatrix} / (\mathcal{K}_m^T \lozenge \mathcal{W}_m)\otimes I_s\right). \tag{3.7}$$

Proof. The proof is similar to that given in [1].

Theorem 2. Suppose that at step m, the matrix $(\mathcal{K}_m^T \lozenge \mathcal{W}_m)$ is nonsingular, then

$$\|R_m^{EGl-FOM}\|_F^2 = \frac{det(\mathcal{K}_m^T \lozenge \mathcal{K}_m)det(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})}{\left(det(\mathcal{K}_m^T \lozenge \mathcal{W}_m)\right)^2},$$

Where $det(X)$ denotes the determinant of the square matrix X.

Proof. We know that $\|R_m^{EGl-FOM}\|_F^2 = (R_m^{EGl-FOM})^T \lozenge R_m^{EGl-FOM}$.

From (3.3), we obtain

$$\begin{aligned}(R_m^{EGl-FOM})^T &\lozenge R_m^{EGl-FOM} \\ &= (R_0 - A\mathcal{K}_m(I_m \otimes B)(y_m \otimes I_s)) \lozenge R_m^{EGl-FOM}.\end{aligned}$$

Using the orthogonality condition (3.4), we have

$$\begin{aligned}(R_m^{EGl-FOM})^T \lozenge R_m^{EGl-FOM} &= -y_m \langle A^m R_0 B^m, R_m\rangle_F \\ &= -y_m((A^m R_0 B^m)^T \lozenge R_m).\end{aligned} \tag{3.8}$$

First, we compute $(A^m R_0 B^m)^T \lozenge R_m$. By (3.7) and Proposition (3) and (2), we obtain

$$(A^m R_0 B^m)^T \lozenge R_m = \left(\begin{pmatrix} \mathcal{K}_m^T \lozenge R_0 & \mathcal{K}_m^T \lozenge \mathcal{W}_m \\ (A^m R_0 B^m)^T \lozenge R_0 & (A^m R_0 B^m)^T \lozenge \mathcal{W}_m \end{pmatrix} / (\mathcal{K}_m^T \lozenge \mathcal{W}_m)\right) \tag{3.9}$$

We know that $\mathcal{K}_{m+1} = [R_0, \mathcal{W}_m]$and $\mathcal{K}_{m+1} = [\mathcal{K}_m, A^m R_0 B^m]$. Then (3.9) can be expressed as

$$(A^m R_0 B^m)^T \lozenge R_m = \left({\mathcal{K}_{m+1}}^T \lozenge \mathcal{K}_{m+1} / {\mathcal{K}_m}^T \lozenge \mathcal{W}_m\right).$$

As $(A^m R_0 B^m)^T \lozenge R_m$ is a scalar, it follows that

$$(A^m R_0 B^m)^T \lozenge R_m = (-1)^m \frac{det\left({\mathcal{K}_{m+1}}^T \lozenge \mathcal{K}_{m+1}\right)}{det\left({\mathcal{K}_m}^T \lozenge \mathcal{W}_m\right)} \tag{3.10}$$

On the other hand, y_m can be computed from (3.5) by the Cramer rule, as

$$y_m = (-1)^m \frac{det(\mathcal{K}_m{}^T \lozenge \mathcal{K}_m)}{det(\mathcal{K}_m{}^T \lozenge \mathcal{W}_m)} \quad (3.11)$$

Therefore, with substitution (3.10) and (3.11) in (3.8), the result follows.

3.2. An Extended Global GMRES Method

An extended global GMRES (EGl-GMRES) method builds the approximate solution $X_m^{EGl-GMRES}$ satisfying the following two relations:

$X_m^{EGl-GMRE} - X_0 \in \mathcal{G}K_m(A, R_0, B)$

and

$R_m^{EGl-GMRES} = F - AX_m^{EGl-GMRE}B \perp_F A\mathcal{G}K_m(A, R_0, B)(I_m \otimes B).$

From the above two relations, we have

$$X_m^{EGl-GMRE} = X_0 + \mathcal{K}_m(\alpha \otimes I_s) \quad (3.12)$$

and

$$R_m^{EGl-GMRE} = R_0 - \mathcal{W}_m(\alpha \otimes I_s) \quad (3.13)$$

Where α is

$$(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m)\alpha = \mathcal{W}_m{}^T \lozenge R_0 \quad (3.14)$$

We have the minimization property

$$\|R_m^{EGl-GMRE}\|_F = \min_{z \in \mathcal{G}K_m(A, R_0, B)} \|R_0 - AZB\|_F$$

The next results show that $X_m^{EGl-GMRE}$ and $R_m^{EGl-GMRE}$ may be expressed as Schur complements.

Theorem 3 ([1]). Assume that $det(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m) \neq 0$. Then the approximate solution $X_m^{EGl-GMRE}$ and the corresponding residual $R_m^{EGl-GMRE}$ are expressed as the following Schur complements:

$$X_m^{EGl-GMRES} = \left(\begin{pmatrix} X_0 & -\mathcal{K}_m \\ (\mathcal{W}_m^T \lozenge R_0) \otimes I_s & (\mathcal{W}_m^T \lozenge \mathcal{W}_m) \otimes I_s \end{pmatrix} / (\mathcal{W}_m^T \lozenge \mathcal{W}_m) \otimes I_s \right) \quad (3.15)$$

and

$$R_m^{EGl-GMRES} = \left(\begin{pmatrix} R_0 & \mathcal{W}_m \\ (\mathcal{W}_m^T \lozenge R_0)\otimes I_s & (\mathcal{W}_m^T \lozenge \mathcal{W}_m)\otimes I_s \end{pmatrix} / (\mathcal{W}_m^T \lozenge \mathcal{W}_m)\otimes I_s\right) \tag{3.16}$$

In the following result, we provide an expression of the residual norm of the EGl-GMRES method.

Theorem 4. If $det(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m) \neq 0$, then we have

$$\|R_m^{EGl-GMRES}\|_F^2 = \frac{det(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})}{det(\mathcal{W}_m^T \lozenge \mathcal{W}_m)} \tag{3.17}$$

Proof. We have

$$\|R_m^{EGl-GMRES}\|_F^2 = (R_m^{EGl-GMRES})^T \lozenge R_m^{EGl-GMRES}$$
$$= \left(R_0 - \mathcal{W}_m(\alpha\otimes I_s)\right)^T \lozenge \left(R_0 - \mathcal{W}_m(\alpha\otimes I_s)\right)$$
$$= R_0{}^T \lozenge R_0 - R_0{}^T \lozenge \left(\mathcal{W}_m(\alpha\otimes I_s)\right) - \left(R_0{}^T \lozenge \left(\mathcal{W}_m(\alpha\otimes I_s)\right)\right)^T +$$
$$\left(\mathcal{W}_m(\alpha\otimes I_s)\right)^T \left(\mathcal{W}_m(\alpha\otimes I_s)\right)$$

Using the relation (6) of Proposition 1, we obtain

$$\|R_m^{EGl-GMRES}\|_F^2 = R_0{}^T \lozenge R_0 - \left(R_0{}^T \lozenge \mathcal{W}_m\right)\alpha - \left(\left(R_0{}^T \lozenge \mathcal{W}_m\right)\alpha\right)^T$$
$$-\left(\left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)\alpha\right)^T \alpha \tag{3.18}$$

As $\det(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m) \neq 0$, from (3.14) we have $\alpha = \left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)^{-1}\left(\mathcal{W}_m{}^T \lozenge R_0\right)$.

Now, by substituting $\alpha = \left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)^{-1}\left(\mathcal{W}_m{}^T \lozenge R_0\right)$ in (3.18), we get

$$\|R_m^{EGl-GMRES}\|_F^2 = R_0{}^T \lozenge R_0 - \left(R_0{}^T \lozenge \mathcal{W}_m\right)\left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)^{-1}\left(\mathcal{W}_m{}^T \lozenge R_0\right)$$

Then, using the definition of the Schur complement, it follows that

$$\|R_m^{EGl-GMRES}\|_F^2 = \left(\begin{pmatrix} R_0{}^T \lozenge R_0 & R_0{}^T \lozenge \mathcal{W}_m \\ \mathcal{W}_m^T \lozenge R_0 & \mathcal{W}_m^T \lozenge \mathcal{W}_m \end{pmatrix} / \mathcal{W}_m^T \lozenge \mathcal{W}_m\right) \tag{3.19}$$

We know that $\mathcal{K}_{m+1} = [R_0, \mathcal{W}_m]$, then (3.19) can be expressed as

$$\|R_m^{EGl-GMRES}\|_F^2 = (\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})/(\mathcal{W}_m^T \lozenge \mathcal{W}_m)$$

and as $\|R_m^{EGl-GMRES}\|_F^2$ is a scalar, it follows that

$$\|R_m^{EGl-GMRES}\|_F^2 = \frac{det(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})}{det(\mathcal{W}_m^T \lozenge \mathcal{W}_m)}.$$

4. Analysis of Convergence Of The Extended Global FOM And The Extended Global GMRES Methods

In this section, we present some convergence results for the extended global FOM and the extended global GMRES methods. Applying the global QR decomposition $\mathcal{K}_{m+1}$ and $\mathcal{K}_m$ yields

$$\mathcal{K}_{m+1} = \mathcal{Q}_{m+1}(R_{m+1} \otimes I_s) \text{ and } \mathcal{K}_m = \mathcal{Q}_m(R_m \otimes I_s) \tag{4.1}$$

with $\mathcal{Q}_{m+1} \in \mathbb{R}^{n\times(m+1)s}$, $R_{m+1} \in \mathbb{R}^{(m+1)\times(m+1)}$, $\mathcal{Q}_m \in \mathbb{R}^{n\times ms}$ and $R_m \in \mathbb{R}^{m\times m}$.

$\mathcal{Q}_{m+1}$ and $\mathcal{Q}_m$ are F-orthonormal (i.e., orthonormal with respect to the $\lozenge$ product); R_{m+1} and R_m are two upper triangular matrices. Note that

$$\mathcal{K}_{m+1} = \begin{bmatrix} 0_{s\times ms} \\ I_{ms} \end{bmatrix} = A\mathcal{K}_m(I_m \otimes B) \tag{4.2}$$

By substituting (4.1) in (4.2), we have

$$\mathcal{Q}_{m+1}(R_{m+1} \otimes I_s)\begin{bmatrix} 0_{s\times ms} \\ I_{ms} \end{bmatrix} = A\mathcal{Q}_m(R_m \otimes I_s)(I_m \otimes B) \tag{4.3}$$

Hence by using the $\lozenge$ product (with $\mathcal{Q}_{m+1}$) for (4.3) and the relation (8) of Proposition 1, we obtain

$$\left(\mathcal{Q}_{m+1}{}^T \lozenge A\mathcal{Q}_m(I_m \otimes B)\right) R_m = R_{m+1}\begin{bmatrix} 0_{1\times m} \\ I_m \end{bmatrix} \tag{4.4}$$

By multiplying (4.4) from the right by R_m^{-1}, it follows that

$$\mathcal{Q}_{m+1}{}^T \lozenge A\mathcal{Q}_m(I_m \otimes B) = R_{m+1}\begin{bmatrix} 0_{1\times m} \\ I_m \end{bmatrix} R_m^{-1} \tag{4.5}$$

We define the $(m+1)\times m$ matrix $\bar{H}_m$ as $\bar{H}_m = \mathcal{Q}_{m+1}{}^T \lozenge A\mathcal{Q}_m(I_m \otimes B)$.

Since R_{m+1} and R_m are both upper triangular matrices, it follows that $\bar{H}_m$ is an upper Hessenberg matrix. Using the fact that $\mathcal{Q}_{m+1} = [\mathcal{Q}_m, Q_{m+1}]$, we obtain

$$\begin{aligned}\bar{H}_m &= \mathcal{Q}_{m+1}{}^T \lozenge A\mathcal{Q}_m(I_m \otimes B) \\ &= \begin{bmatrix} \mathcal{Q}_m{}^T \\ Q_{m+1}{}^T \end{bmatrix} \lozenge A\mathcal{Q}_m(I_m \otimes B) = \begin{bmatrix} \mathcal{Q}_m{}^T \lozenge A\mathcal{Q}_m \\ Q_{m+1}{}^T \lozenge A\mathcal{Q}_m(I_m \otimes B) \end{bmatrix}.\end{aligned} \tag{4.6}$$

The matrix $\bar{H}_m$ is obtained from $\bar{H}_m$ by deleting its last row. $\bar{H}_m$ is also an upper Hessenberg matrix given by

$$H_m = \mathcal{Q}_m{}^T \lozenge A\mathcal{Q}_m \tag{4.7}$$

By substituting (4.7) in (4.6), we get

$$\overline{H}_m = \begin{bmatrix} H_m \\ {Q_{m+1}}^T \lozenge AQ_m(I_m \otimes B) \end{bmatrix}.$$

Since $\overline{H}_m$ is an upper Hessenberg matrix, then

$${Q_{m+1}}^T \lozenge AQ_m(I_m \otimes B) = h_{m+1,m} e_m^T.$$

It follows that

$$\overline{H}_m = \begin{bmatrix} H_m \\ h_{m+1,m} e_m^T \end{bmatrix}. \tag{4.8}$$

By applying the Kronecker product (with I_s) for (4.8), we find that

$$\overline{H}_m \otimes I_s = \left(\mathcal{Q}_{m+1}{}^T \lozenge AQ_m(I_m \otimes B) \right) \otimes I_s = \begin{bmatrix} H_m \otimes I_s \\ h_{m+1,m} E_m^T \end{bmatrix} \tag{4.9}$$

By multiplying (4.9) from the left by $\mathcal{Q}_{m+1}$, we obtain

$$\mathcal{Q}_{m+1}(\overline{H}_m \otimes I_s) = \mathcal{Q}_{m+1} \begin{bmatrix} H_m \otimes I_s \\ h_{m+1,m} E_m^T \end{bmatrix}.$$

Since $\mathcal{Q}_{m+1} = [\mathcal{Q}_m, Q_{m+1}]$, we have

$$\mathcal{Q}_{m+1}(\overline{H}_m \otimes I_s) = \mathcal{Q}_m(H_m \otimes I_s) + Q_{m+1} h_{m+1,m} E_m^T. \tag{4.10}$$

We know that $A\mathcal{K}_m B \subseteq \mathcal{K}_{m+1}$, then there exists a matrix L of size $(m+1) \times m$ such that

$$AQ_m(I_m \otimes B) = \mathcal{Q}_{m+1}(L \otimes I_s). \tag{4.11}$$

By applying the $\lozenge$ product (with $\mathcal{Q}_{m+1}^T$) for (4.11) from the left, we find that

$$\mathcal{Q}_{m+1}{}^T \lozenge AQ_m(I_m \otimes B) = L. \tag{4.12}$$

From (4.6), (4.11) and (4.12), implies that

$$AQ_m(I_m \otimes B) = \mathcal{Q}_{m+1}(\overline{H}_m \otimes I_s). \tag{4.13}$$

By substituting (4.13) in (4.10), it follows that

$$AQ_m(I_m \otimes B) = \mathcal{Q}_m(H_m \otimes I_s) + Q_{m+1} h_{m+1,m} E_m^T. \tag{4.14}$$

Where $E_m^T = [0_s, \ldots, 0_s, I_s]$ and $h_{m+1,m} = {Q_{m+1}}^T \lozenge AQ_m(I_m \otimes B) = (r_{m+1,m+1})/(r_{m,m})$.

Theorem 5. At step m, let $R_m^{EGl-GMRES}$ and $R_m^{EGl-FOM}$ be the residual produced by the EGl-GMRES and the EGl-FOM methods, respectively. Then

$$\frac{\|R_m^{EGl-GMRES}\|_F^2}{\left\|R_{m-1}^{EGl-GMRES}\right\|_F^2} = \frac{det(\bar{H}_{m-1}^T \bar{H}_{m-1})}{det(\bar{H}_m^T \bar{H}_m)} h_{m+1,m}^2 \tag{4.15}$$

and

$$\frac{\|R_m^{EGl-FOM}\|_F^2}{\left\|R_{m-1}^{EGl-FOM}\right\|_F^2} = \frac{det(H_{m-1}^T H_{m-1})}{det(H_m^T H_m)} h_{m+1,m}^2 \tag{4.16}$$

Proof. We have

$$\mathcal{W}_m^T \lozenge \mathcal{W}_m = \left(A\mathcal{K}_m(I_m \otimes B)\right)^T \lozenge \left(A\mathcal{K}_m(I_m \otimes B)\right) \tag{4.17}$$

Using the global QR decomposition to the matrix $\mathcal{K}_m$, the relation (4.17) can be expressed as

$$\mathcal{W}_m^T \lozenge \mathcal{W}_m = \left(A\mathcal{Q}_m(R_m \otimes I_s)(I_m \otimes B)\right)^T \lozenge \left(A\mathcal{Q}_m(R_m \otimes I_s)(I_m \otimes B)\right) \tag{4.18}$$

Since

$$(R_m \otimes I_s)(I_m \otimes B) = (I_m \otimes B)(R_m \otimes I_s)$$

(4.18) can be expressed as

$$\mathcal{W}_m^T \lozenge \mathcal{W}_m = \left(A\mathcal{Q}_m(I_m \otimes B)(R_m \otimes I_s)\right)^T \lozenge \left(A\mathcal{Q}_m(I_m \otimes B)(R_m \otimes I_s)\right)$$

Using the relation (6) of Proposition 1, we have

$$\mathcal{W}_m^T \lozenge \mathcal{W}_m = \left[\left(A\mathcal{Q}_m(I_m \otimes B)(R_m \otimes I_s)\right)^T \lozenge A\mathcal{Q}_m(I_m \otimes B)\right] R_m$$

We know that $A\mathcal{Q}_m(I_m \otimes B) = \mathcal{Q}_{m+1}(\bar{H}_m \otimes I_s)$, then

$$\mathcal{W}_m^T \lozenge \mathcal{W}_m = \left[\left(\mathcal{Q}_{m+1}(\bar{H}_m \otimes I_s)(R_m \otimes I_s)\right)^T \lozenge A\mathcal{Q}_m(I_m \otimes B)\right] R_m$$

So using the relations (4) and (6) of Proposition 1, we get

$$\mathcal{W}_m^T \lozenge \mathcal{W}_m = R_m^T \bar{H}_m^T \bar{H}_m R_m \tag{4.19}$$

Similarly, we also have

$$\mathcal{K}_m^T \lozenge \mathcal{K}_m = R_m^T R_m \tag{4.20}$$

Now, applying Theorem 4, we get

$$\frac{\|R_m^{EGl-GMRES}\|_F^2}{\|R_{m-1}^{EGl-GMRES}\|_F^2} = \frac{det(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}) det(\mathcal{W}_{m-1}^T \lozenge \mathcal{W}_{m-1})}{det(\mathcal{W}_m^T \lozenge \mathcal{W}_m) det(\mathcal{K}_m^T \lozenge \mathcal{K}_m)} \tag{4.21}$$

Hence, by substituting (4.19) and (4.20) in (4.21), we obtain

$$\frac{\|R_m^{EGl-GMRES}\|_F^2}{\|R_{m-1}^{EGl-GMRES}\|_F^2} = \frac{det(\overline{H}_{m-1}^T \overline{H}_{m-1})}{det(\overline{H}_m^T \overline{H}_m)} \frac{det(R_{m+1})^2 det(R_{m-1})^2}{det(R_m)^4} \tag{4.22}$$

Now, as $R_{m+1} = \left(\begin{bmatrix} R_m \\ 0_{1\times m} \end{bmatrix}, r_{m+1} \right)$, we get

$$\frac{\|R_m^{EGl-GMRES}\|_F^2}{\|R_{m-1}^{EGl-GMRES}\|_F^2} = \frac{det(\overline{H}_{m-1}^T \overline{H}_{m-1})}{det(\overline{H}_m^T \overline{H}_m)} \frac{r_{m+1,m+1}^2}{r_{m,m}^2}$$

Therefore, as $h_{m+1,m} = r_{m+1,m+1}/r_{m,m}$, we obtain

$$\frac{\|R_m^{EGl-GMRES}\|_F^2}{\|R_{m-1}^{EGl-GMRES}\|_F^2} = \frac{det(\overline{H}_{m-1}^T \overline{H}_{m-1})}{det(\overline{H}_m^T \overline{H}_m)} h_{m+1,m}^2 .$$

The proof of (4.16) can be carried out similarly.

Theorem 6. At the step m, let $R_m^{EGl-GMRES}$ and $R_m^{EGl-FOM}$ be produced by the EGl-GMRES and the EGl-FOM methods, respectively. Therefore, we have

$$\|R_m^{EGl-GMRES}\|_F^2 = \frac{1}{e_1^T (\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1} e_1} \tag{4.23}$$

and

$$\|R_m^{EGl-FOM}\|_F^2 = \frac{e_{m+1}^T (\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1} e_{m+1}}{(e_1^T (\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1} e_{m+1})^2} \tag{4.24}$$

Proof. The proof is similar to that of [1].

Theorem 7.

$$1 \geq \frac{\|R_m^{EGl-GMRES}\|_F}{\|R_0^{EGl-GMRES}\|_F} \geq 2 \frac{\sqrt{\kappa(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})}}{\left(1 + \kappa(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})\right)} \tag{4.25}$$

where $\mathcal{K}_{m+1}$ is the Extended global Krylov matrix and $\kappa(Z)$ denotes the condition number of the matrix Z ·

Proof. We have

$\|R_m^{EGl-GMRES}\|_F^2 = (R_m^{EGl-GMRES})^T \lozenge R_m^{EGl-GMRES}$

Using (3.13), relation (6) of Proposition 1 and (3.14), we obtain

$$\|R_m^{EGl-GMRES}\|_F^2 = R_0{}^T \lozenge R_0 - \left(\mathcal{W}_m{}^T \lozenge R_0\right)^T \left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)^{-1} \left(\mathcal{W}_m{}^T \lozenge R_0\right). \qquad (4.26)$$

We know that $\left\|R_0^{EGl-GMRES}\right\|_F^2 = R_0{}^T \lozenge R_0$, then (4.26) can be expressed as

$$\|R_m^{EGl-GMRES}\|_F^2 = \left\|R_0^{EGl-GMRES}\right\|_F^2 - \left(\mathcal{W}_m{}^T \lozenge R_0\right)^T \left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)^{-1} \left(\mathcal{W}_m{}^T \lozenge R_0\right) \qquad (4.27)$$

As each matrix $\left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)^{-1}$ is symmetric positive definite, then

$$\left(\mathcal{W}_m{}^T \lozenge R_0\right)^T \left(\mathcal{W}_m{}^T \lozenge \mathcal{W}_m\right)^{-1} \left(\mathcal{W}_m{}^T \lozenge R_0\right) \geq 0 \qquad (4.28)$$

Therefore, from (4.27) and (4.28), we get

$$\frac{\|R_m^{EGl-GMRES}\|_F}{\left\|R_0^{EGl-GMRES}\right\|_F} \leq 1 .$$

Since each matrix $\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}$ is symmetric positive definite, then using the Kantorovich inequality, we reveal the second inequality.

By substituting e_1 in the Kantorovich inequality and using the fact that

$$\lambda_{max}(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}) = \|\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}\|_2$$

and

$$\lambda_{min}(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}) = \frac{1}{\|(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1}\|_2}$$

we have

$$\frac{(e_1, e_1)^2}{\left((\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})e_1, e_1\right)\left((\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1}e_1, e_1\right)} \geq \frac{4\|\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}\|_2 \|(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1}\|_2}{(\|\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}\|_2 \|(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1}\|_2 + 1)^2}$$

Therefore, using the fact that

$\kappa(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}) = \|\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}\|_2 \|(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1}\|_2$, $\left((\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})e_1, e_1\right) = \left\|R_0^{EGl-GMRES}\right\|_F^2$ and $((\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})^{-1}e_1, e_1) = \frac{1}{\|R_m^{EGl-GMRES}\|_F^2}$, we get

$$\frac{\|R_m^{EGl-GMRES}\|_F}{\|R_0^{EGl-GMRES}\|_F} \geq 2\frac{\sqrt{\kappa(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})}}{\left(1+\kappa(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1})\right)}.$$

The result of the preceding theorem shows that there is no convergence as long as the extended block Krylov matrix is well-conditioned.

Example. We consider the matrix equation $A_n XB = F$, where

$$A_n = \begin{pmatrix} 0 & \dots & \dots & 0 & 1 \\ 1 & \ddots & \ddots & \vdots & 0 \\ 0 & \ddots & \ddots & \ddots \vdots & \vdots \\ \vdots \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \qquad F = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}.$$

For this example, $\kappa(A_n) = \kappa(B) = 1$. Now, if $X_0 = 0_{n\times 2}$ then for $m = 1, \dots, n-1$, we have

$\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1} = 2I_{m+1}$, $\kappa(\mathcal{K}_{m+1}^T \lozenge \mathcal{K}_{m+1}) = 1$ and $\|R_0^{EGl-GMRES}\|_F^2 = 2$

Using (4.23), we also obtain $\|R_m^{EGl-GMRES}\|_F^2 = 2$.

Hence applying Theorem 7, it follows that there is no convergence.

5. Conclusion

In this work, we present some convergence results of two extended block Krylov subspace methods when applied to matrix equation $AXB = F$ without referring to any algorithm. We also derive new expressions of the approximations and the corresponding residual norms.

References

[1] R. Bouyouli, 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** (2006), 498-511.

[2] F. Ding, T.Chen, Iterative least-squares solutions of coupled Sylvester matrix equations, *Systems & Control Letters* **54** (2005), 95-107.

[3] F. Ding, T. Chen, On iterative solutions of general coupled matrix equations, *SIAM Journal on Control and Optimization* **44** (2006), 2269-2284.

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

[5] J. Ding, Y. Liu, F. Ting, Iterative solutions to matrix equations of the form $A_iXB_i=F_i$, *Computers and Mathematics with Applications* **59** (2010), 3500-3507.

[6] G. X. Haung, F. Yin, K. Guo, An iterative method for the skew-symmetric solution and the optimal approximate solution of the matrix equation $AXB=C$, *Journal of Computational and Applied Mathematics* **212** (2008), 231-244.

[7] R. Lancaster, Theory of Matrix, Academic Press, New York, (1969).

[8] A. Messaoudi, Recursive interpolation algorithm: a formalism for solving systems of linear equations—I. Direct methods, *Journal of Computational and Applied Mathematics* **76** (1996), 13-30.

[9] F. Panjeh Ali Beik, Note to the global GMRES for solving the matrix equation $AXB=F$, *International Journal of Engineering and Natural Sciences* **5** (2011), 101-105.