Approximate solutions to large Stein matrix equations

Khalide Jbilou

Abstract—In the present paper, we propose numerical methods for solving the Stein equation $AXC - X - D = 0$ where the matrix $A$ is large and sparse. Such problems appear in discrete-time control problems, filtering and image restoration. We consider the case where the matrix $D$ is of full rank and the case where $D$ is factored as a product of two matrices. The proposed methods are Krylov subspace methods based on the block Arnoldi algorithm. We give theoretical results and we report some numerical experiments.

Keywords—IIEEEtran, journal, EEEETX, paper, template.

I. INTRODUCTION

We consider the Stein matrix equation

$$AXC - X - D = 0 \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{p \times p}$, $D \in \mathbb{R}^{n \times p}$ and $X \in \mathbb{R}^{n \times p}$. The matrix equation (1) plays an important role in linear control and filtering theory for discrete-time large-scale dynamical systems and other problems; see [5], [6], [8], [17] and the references therein. They also appear in image restoration techniques [4] and in each step of Newton’s method for discrete-time algebraic Riccati equations [11]. (1) can be also referred to as discrete Sylvester equation.

Direct methods for solving the matrix equation (1) such as those proposed in [2], [3], [9] are attractive if the matrices are of small size. The matrix equation (1) can be formulated as an $np \times np$ large linear system using the Kronecker formulation

$$(A \otimes C^T - I_{np}) vec(X) = vec(D) \quad (2)$$

where $\otimes$ denotes the Kronecker product; $(F \otimes G)_{i,j} = [f_{i,j}]$. $vec(X)$ is the vector of $\mathbb{R}^{np}$ obtained by stacking the columns of the matrix $X$ and $I_{np}$ is the $np \times np$ identity matrix.

Krylov subspace methods such as the GMRES algorithm [13] could be used to solve the linear system (2). However, for large problems this approach cannot be applied directly.

The matrix Equation (1) has a unique solution if and only if $\lambda_i(A)\lambda_j(C) \neq 1$ for all $i = 1, \ldots, n$; $j = 1, \ldots, p$ where $\lambda_i(A)$ is the $i$-th eigenvalue of the matrix $A$. This will be assumed through this paper. In particular, if $\rho(A)\rho(C) < 1$ where $\rho(A)$ denotes the spectral radius of the matrix $A$, equation (1) has a unique solution.

In this work, we present Galerkin projection methods based on the block Arnoldi algorithm [14], [15]. We first consider the case where the $n \times p$ matrix $D$ is of full rank and $p \ll n$. The second part of this paper is devoted to the case where both matrices $A$ and $C$ are large and $D$ is factored as $D = EF^T$ with a low rank.

II. THE BLOCK ARNOLDI ALGORITHM

In this section, we recall the block Arnoldi process applied to the matrix $A$ and starting with the $n \times p$ orthonormal matrix $V_1$.

The block Krylov subspace $K_k(A, V_1) = span\{V_1, AV_1, \ldots, A^{k-1}V_1\}$, is the subspace generated by the columns of the matrices $V_1, AV_1, \ldots, A^{k-1}V_1$. The block Arnoldi algorithm constructs the blocks $V_1, \ldots, V_k$ whose columns form an orthonormal basis of the block Krylov subspace $K_k(A, V_1)$. The algorithm is described as follows

Algorithm 1 The block Arnoldi algorithm

1) Choose a unitary $n \times p$ matrix $V_1$.
2) For $j = 1, \ldots, k$
   - $W_j = AV_j$,
   - for $i = 1, 2, \ldots, j$
     - $H_{i,j} = V_i^T W_j$,
     - $W_j = W_j - V_i H_{i,j}$,
   - end for
   - $Q_j R_j = W_j$ ($QR$ decomposition)
   - Set $V_{j+1} = Q_j$ and $H_{j+1,j} = R_j$.
3) End

The blocks $V_1, \ldots, V_k$ constructed by Algorithm 1 have their columns mutually orthogonal provided that the upper triangular matrices $H_{j+1,j}$ are of maximum rank. If $H_{j+1,j} = 0$ then $K_j$ is invariant under $A$.

Let $H_k$ denotes the $(k+1)p \times kp$ upper band-Hessenberg matrix whose nonzero entries $h_{i,j}$; $i = 1, \ldots, (k+1)p$ and $j = 1, \ldots, kp$ are defined by Algorithm 1. $H_k$ is the $kp \times kp$ matrix obtained from $H_k$ by deleting the last $p$-rows and $H_{k+1,k}$ is the $p \times p$ submatrix of the last $p$-rows and the last $p$-columns of $H_k$.

The matrix $V_k$ is defined by $V_k = [V_1, \ldots, V_k]$ where $V_i$, $i = 1, \ldots, k$ is the $i$-th block constructed by the block Arnoldi algorithm. From the block Arnoldi algorithm we can deduce the following relations

$$AV_k = V_k H_k + V_{k+1} H_{k+1,k} E_k^T \quad (3)$$

and

$$H_k = V_k^T AV_k \quad \text{and} \quad V_k^T V_k = I_k \quad (4)$$

where $E_k$ is the matrix of the last $p$ columns of the $kp \times kp$ identity matrix $I_{kp}$.

III. THE CASE WHERE D IS FULL RANK

In this section, we consider the case where the $n \times p$ right-hand side matrix $D$ of (1) is of full rank, $C$ nonsingular and we assume that $p \ll n$. 

LMPA Université du Littoral, 50 rue F. Buisson BP 699, F-62228 Calais Cedex, France E-mail: jbilou@lmpa.univ-littoral.fr.
Let $A$ be the linear operator from $\mathbb{R}^{n \times p}$ onto $\mathbb{R}^{n \times p}$ defined as follows

$$A : X \mapsto AX = AXC - X.$$  

(5)

Then the Stein equation (1) can be written as

$$A(X) = D.$$  

(6)

We will solve the problem (6) which is equivalent to the initial problem (1).

Starting from an initial guess $X_0$ and the corresponding residual $R_0 = D - AX_0C + X_0$, the block Arnoldi Stein method constructs, at step $k$, the new approximation $X_k$ such that

$$X_k(i) - X_0(i) = Z_k(i) \in K_k(A, R_0); \ i = 1, \ldots, p$$

(7)

with the orthogonality relation

$$R_k(i) \perp K_k(A, R_0); \ i = 1, \ldots, p$$

(8)

where $R_k(i)$ is the $i$th component of the residual $R_k = D - AX_kC$ and $X_k(i)$ is the $i$th component of $X_k$. We give the following result which is easy to prove [7].

**Theorem 1:** Let $A$ be the operator defined by (5) and assume that $R_0$ is of full rank. Then

$$K_k(A, R_0) = K_k(A, R_0).$$

Using this property, the relations (7) and (8) are written as

$$X_k(i) - X_0(i) = Z_k(i) \in K_k(A, R_0),$$

(9)

and

$$R_k(i) \perp K_k(A, R_0); \ i = 1, \ldots, p.$$  

(10)

Assume that $R_0$ is of rank $p$ and let $R_0 = V_1U_1$ be the QR decomposition of $R_0$ where the $n \times p$ matrix $V_1$ is orthogonal and $U_1$ is $p \times p$ upper triangular.

Now as the columns of the matrix $V_k$ (constructed by the block Arnoldi algorithm) form a basis of the block Krylov subspace $K_k(A, R_0)$, the relation (9) implies that $X_k = X_0 + V_kY_k$, where $Y_k$ is a $kp \times p$ matrix. Using the orthogonality relation (10), it follows that

$$Y_k^T (R_0 - AV_kY_kC + V_kY_k) = 0.$$  

We finally obtain the low-dimensional Stein equation

$$H_k Y_k C - Y_k = \tilde{D}$$  

(11)

with $\tilde{D} = \tilde{E}_1U_1$, where $\tilde{E}_1$ is the $kp \times p$ matrix whose upper $p \times p$ principal block is the identity matrix.

The matrix equation (11) will be solved by using a direct method such as the Hessenberg-Schur method [5]. We assume that during the iterations $\lambda_1(H_k) \lambda_1(C) < 1$ and this implies that the equation (11) has a unique solution.

Let us give now an expression of the residual norm that can be used to stop the iterations in the block-Arnoldi Stein algorithm without having to compute an extra product with the matrix $A$.

**Theorem 2:** At step $k$, the norm of the residual $R_k$ is given by

$$\| R_k \|_F = \| H_{k+1} E_k^T Y_k C \|_F = \| H_{k+1} \tilde{Y}_k C \|_F,$$

where $\tilde{Y}_k$ is the $p \times p$ matrix corresponding to the last $p$ rows of the matrix $Y_k$.

**Proof:** At step $k$, the residual $R_k = D - AX_kC + X_k$, with $X_k = X_0 + V_kY_k$, is expressed as

$$R_k = R_0 - AV_kY_kC + V_kY_k$$

and from the relation $AV_k = V_k \tilde{H}_k + V_{k+1}H_{k+1}E_k^T$, it follows that

$$R_k = V_k [\tilde{D} - \tilde{H}_k Y_k C + Y_k] - V_{k+1}H_{k+1}E_k^T Y_k C.$$  

Therefore using (11) and the fact that the matrix $V_{k+1}$ is orthogonal the result follows.

The next result shows that the approximate solution $X_k$ is an exact solution of a perturbed Stein matrix equation.

**Theorem 3:** Assume that $k$ steps of the block Arnoldi Stein method have been run and let $X_k = X_0 + V_kY_k$, be the obtained approximate solution to (1) where $Y_k$ satisfies (11). Then $X_k$ is a solution of the perturbed problem

$$(A - F_k) XC - X = D - F_k X_0 C,$$

with $F_k = V_{k+1}H_{k+1}E_k^T$ and $\| F_k \|_F = \| H_{k+1} \|_F$.

**Proof:** Multiplying on the left the equation (11) by the matrix $V_k$ we get

$$V_k \tilde{H}_k Y_k C - V_kY_k = V_k \tilde{D}.$$  

Using the relation $AV_k = V_k \tilde{H}_k + V_{k+1}H_{k+1}E_k^T$ and the fact that $Y_k$ is orthogonal it follows that

$$AV_kY_k C - V_{k+1}H_{k+1}E_k^T Y_k C - V_k Y_k = V_k \tilde{D}.$$  

Then as $X_k = X_0 + V_kY_k$, $V_k E_k = V_k$ and $V_k \tilde{D} = R_0$, we get

$$(A - F_k) X_k C - X_k = D - F_k X_0 C.$$  

where $F_k = V_{k+1}H_{k+1}E_k^T$ and then $\| F_k \|_F = \| H_{k+1} \|_F$.

Note that when $H_{k+1} = 0$, $F_k = 0$ and hence $X_k$ is the exact solution of the Stein matrix equation (1). In practice, the computational requirements grow with the iteration and then the block Arnoldi algorithm will be computed in a restarted mode. The block-Arnoldi algorithm for solving (1) is summarized as follows

**Algorithm 2** The block Arnoldi algorithm for Stein equations

1. Choose a tolerance $tol$, an initial guess $X_0$ and an integer $k_{max}$.

2. Compute $R_0 = D + X_0 - AX_0 C$ and $R_0 = V_1U_1$; (QR decomposition.)

3. For $k = 1, \ldots, k_{max}$,

   - Apply Algorithm 1 to the pair $(A, V_1)$ to generate $V_1, \ldots, V_{k+1}$ and the block Hessenberg matrix $H_k$.
   - Solve by a direct method the low-order Stein equation $H_k Y_k C - Y_k = \tilde{D}$.
IV. LOW-RANK APPROXIMATE SOLUTIONS TO LARGE STEIN EQUATIONS

In this section, we consider large Stein matrix equations with low-rank right-hand sides

\[ AXC - X = EF^T \]  

(12)

where \( A \in \mathbb{R}^{n \times n}, C \in \mathbb{R}^{p \times p}, E \in \mathbb{R}^{n \times r} \) and \( F \in \mathbb{R}^{n \times r} \). We assume that \( n \) and \( p \) are large; \( r \ll n \) and \( r \ll p \). From now on, we suppose that \( \rho(A)\rho(C) < 1 \) which ensures that (12) has a unique solution.

Equations of the form (12) arise in many application such as control theory and model reduction in large scale discrete-time dynamical systems [17]. This is the case for example when one has to compute the controllability matrix \( X_c \) and observability matrix \( X_o \) Gramians by solving two symmetric Stein equations

\[ AX_cA^T - X_c + E F^T = 0 \]  

and

\[ A^T X_o A - X_o + F F^T = 0 \]  

(13)

The Gramians of linear time-invariant systems play a fundamental role in many analysis and design problems such as computing the Hankel singular values, the \( H_2 \) norm of dynamical systems and model reduction techniques [8], [17].

Next, we will show how to extract low-rank approximate solutions to (12) via the block Arnoldi algorithm. At step \( k \), let \( K_k(A, E) \) and \( K_k(C, F) \) be the block Krylov subspaces associated with \( (A, E) \) and \( (C^T, F) \), respectively. Consider the QR decompositions \( E = V_l A U_l, F = V_l C U_l \) and apply the block Arnoldi process to the pairs \( (A, E) \) and \( (C^T, F) \) starting with \( V_1 \) and \( V_1 \) respectively. We obtain two block orthonormal bases \( \{V_1, A, \ldots, V_k, A\} \) and \( \{V_1, C, \ldots, V_k, C\} \) of the Krylov subspaces \( K_k(A, E) \) and \( K_k(C, F) \) respectively. We denote by \( \mathcal{H}_{k,A} \) and \( \mathcal{H}_{k,C} \) the block upper Hessenberg matrices given by

\[ \mathcal{H}_{k,A} = V_l^T A V_{k,A} \quad \text{and} \quad \mathcal{H}_{k,C} = V_l^T C F^T V_{k,C} \]

where \( V_{k,A} = [V_1, A, \ldots, V_k, A], V_{k,C} = [V_1, C, \ldots, V_k, C] \) and \( \mathcal{H}_{k,A} = [H_{ij}]_{i,j=1,\ldots,p} \). We also have the following relations

\[ A V_{k,A} = V_{k,A} \mathcal{H}_{k,A} + V_{k+1, A} H_{k+1, k} E_k^T \]

(13)

and

\[ C^T V_{k,C} = V_{k,C} \mathcal{H}_{k,C} + V_{k+1, C} H_{k+1, k} E_k^T \]

(14)

where \( E_k \) is the matrix of the last \( r \) columns of the \( k r \times k r \) identity matrix \( I_{k r} \).

The following result gives the exact solution of (12) in terms of the two block Arnoldi bases.

Theorem 4: Let \( q \) and \( l \) be the degrees of the minimal polynomials of \( A \) for \( E \) and \( C^T \) for \( F \) respectively. Then the exact solution of the Stein equation (12) is given by

\[ X = V_l A Z V_{l,C}^T \]

(15)

where \( Z \) solves the problem

\[ \mathcal{H}_{l,C} Z \mathcal{H}_{l,C}^T - Z = \bar{E} F^T \]

(16)

with \( \bar{E} = \bar{E}_1 U_1, \bar{F} = \bar{E}_1 U_2 \) and \( \bar{E}_1 \) is the \( kp \times p \) matrix whose upper \( p \times p \) principal block is the identity matrix \( I_p \).

Proof: Since \( q \) and \( l \) are the degrees of the minimal polynomials of \( A \) for \( E \) and \( C^T \) for \( F \), respectively, it follows that

\[ A V_{l,A} = V_{l,A} \mathcal{H}_{l,A} \quad \text{and} \quad C^T V_{l,C} = V_{l,C} \mathcal{H}_{l,C}. \]

Multiplying on the left the two sides of (16) by \( V_{l,A} \) and \( V_{l,C}^T \) respectively, we get

\[ V_{l,A} A H_{l,A} Z V_{l,C}^T - V_{l,A} Z V_{l,C}^T = V_{l,A} \bar{E} F^T V_{l,C}. \]

Using (17) and the fact that \( V_{l,A} \bar{E} = E \) and \( V_{l,C} \bar{F} = F \), equation (18) is written as

\[ A V_{l,A} Z V_{l,C}^T - A V_{l,A} Z V_{l,C}^T = E F^T. \]

This shows that \( X = V_{l,A} Z V_{l,C}^T \) is the solution (unique) of (12).

Following the result of Theorem 4.1, we consider low-rank approximations of the form

\[ X_k = V_{k,A} Z_k V_{k,C}^T \]

(19)

where \( Z_k \in \mathbb{R}^{k p \times k p} \) is the solution of the low order Stein equation

\[ \mathcal{H}_{k,A} Z_k \mathcal{H}_{k,C}^T = Z_k = \bar{E} F^T \]

(20)

The low-dimensional discrete Stein equation (20) will be solved by a direct method such as the Hessenberg-Schur method [2]. We assume that during the iterations, \( \lambda_i(\mathcal{H}_{k,C}) \lambda_i(\mathcal{H}_{k,A}) < 1 \) which ensures that (20) has a unique solution.

In the following, we give some theoretical results. The next theorem shows that the low-order approximate solution \( X_k \) is a solution of a perturbed Stein equation.

Theorem 5: At step \( k \), let \( X_k \) be the low-rank approximate solution given by (19) and (20). Then

\[ (A - A_k) X_k (C - C_k) - X_k = E F^T \]

(21)

where \( A_k = V_{k+1, A} H_{k+1, A} V_{k,A}^T \) and \( C_k = V_{k+1, C} H_{k+1, C} V_{k,C}^T \).

Proof: Multiplying the low order Stein equation (20) on the left by \( V_{k,A} \) and on the right by \( V_{k,C}^T \), using the relations (13) and (14) and the fact that the two matrices \( V_{k,A}, V_{k,C} \) are orthogonal, we get

\[ A X_k C - X_k = A X_k C_k - A_k X_k C + A_k X_k C_k = E F^T \]

(22)

with \( A_k C_c = V_{k+1, A} H_{k+1, A} V_{k,A}^T \) and \( C_k = V_{k+1, C} H_{k+1, C} V_{k,C}^T \). This shows the result.

The computation of the approximation \( X_k \) given by (19) needs the product of three matrices and this becomes very expensive as \( k \) increases. In the next theorem, we show how to compute the residual norms used to stop the iterations without computing the approximation \( X_k \). When convergence is achieved, \( X_k \) is given in a factored form and not formed explicitly.

Theorem 6: Let \( X_k = V_{k,A} Z_k V_{k,C}^T \) be the approximate solution obtained, at step \( k \), by the block Arnoldi Stein
method where \( Z_k \) is the solution of (20) and let \( R_k \) be the corresponding residual. Then
\[
\| R_k \|_2^2 = \| R_{k,1} \|^2 + \| R_{k,2} \|^2 + \| R_{k,3} \|^2
\]  
(23)
with
\[
R_{k,1} = H_{k,A} Z_k E_k (H_{k+1,k}^T)^T, \quad R_{k,2} = H_{k+1,k}^T E_k^T Z_k H_{k,C}^T \quad \text{and} \quad R_{k,3} = H_{k+1,k}^T E_k Z_k E_k H_{k+1,k}^T
\]
where \( E_k \) is the \( kp \times p \) matrix of the last \( p \) columns of the identity matrix \( I_{kp} \).

Proof: The residual is given by \( R_k = EF^T - AV_{k,A} Z_k V_{k,C}^T + V_{k,A} Z_k V_{k,C}^T \).
Then using the relations \( AV_{k,A} = V_{k+1,A} H_{k,A} \), \( CT V_{k,C} = V_{k+1,C} H_{k,C} \) and the expressions \( \tilde{H}_{k,A} = \left( H_{k+1,A}^T I_k \right)^T \) and \( \tilde{H}_{k,C} = \left( H_{k+1,C}^T I_k \right)^T \), the residual \( R_k \) can be expressed in a matrix form
\[
R_k = V_{k+1,A} Z_k V_{k+1,C}^T
\]  
(24)
with
\[
Z_k = \left( \begin{array}{cc}
0 & H_{k+1,k}^T E_k^T Z_k H_{k,C}^T \\
H_{k+1,k}^T E_k^T Z_k H_{k,C}^T & H_{k+1,k}^T E_k Z_k H_{k+1,k}^T
\end{array} \right)
\]
where \( Z_k \) solves (20). Therefore taking the norm of (24) and using the fact that \( V_{k+1,A} = [V_{k,A}, V_{k+1,A}] \) and \( V_{k+1,C} = [V_{k,C}, V_{k+1,C}] \) are orthonormal matrices, the result (23) follows.

The block-Arnoldi algorithm for solving (12) is summarized as follows

Algorithm 3 The block Arnoldi algorithm for Stein equations

1) Choose a tolerance tol and an integer \( knax \).
2) Compute \( E = V_{1,A} U_1 \) and \( F = V_{1,C} U_2 \) (QR)
3) For \( k = 1, \ldots, knax \):
   a) Apply Algorithm 1 to \( (A, V_1) \) and \( (CT, V_1) \) to generate \( V_{1,A}, \ldots, V_{k,A}; V_{1,C}, \ldots, V_{k,C} \) and the block Hessenberg matrices \( \tilde{H}_{k,A} \) and \( \tilde{H}_{k,C} \).
   b) Solve by a direct method the low-order Stein equation: \( H_{k,A} Z_k \tilde{H}_{k,C} - Z_k = \tilde{B} F^T \).
   c) If \( \| R_k \|_F < tol \), stop
4) End.

V. THE SYMMETRIC STEIN EQUATION

In this section, we consider symmetric Stein equations
\[
AX^T = X + BB^T = 0
\]  
(25)
where \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{n \times p} \) with \( p \ll n \).

If \( \rho(A) < 1 \) where \( \rho(A) \) denotes the spectral radius of \( A \), the symmetric Stein equation (25) (called also Schur stable) has a unique solution given by (see [11])
\[
X = \sum_{i=0}^{\infty} A^i BB^T A^T.
\]  
(26)
As in [10], we apply the block Arnoldi algorithm to the pair \( (A, B) \) and get the matrices \( V_k \) and \( \tilde{H}_k \). We then consider approximations of the form \( X_k = V_k Z_k V_k^T \) where \( Z_k \) solves the low-order symmetric Stein equation
\[
\tilde{H}_k Z_k \tilde{H}_k^T - Z_k = \tilde{B} \tilde{B}^T.
\]  
(27)
with \( \tilde{B} = V_k^T B \).

Using Theorem 6 and Theorem 7, we get the following results

Theorem 7: Let \( X_k \) be the low-rank approximate solution obtained at step \( k \) and let \( X \) be the exact solution of the symmetric Stein equation (25). Then
\[
(A - A_k) X_k (A - A_k)^T - X_k + BB^T = 0
\]  
(28)
and
\[
\| R_k \|_2^2 = \| H_k Y_k E_k H_k^T \|^2_F + H_k E_k Y_k E_k H_k^T \|^2_F
\]  
(29)
where \( E_k \) is the \( kp \times p \) matrix of the last \( p \) columns of the identity matrix \( I_{kp} \) and \( A_k = V_{k+1} H_{k+1,k} V_{k+1}^T \).

In the following theorem, we give an upper bound of the norm of the error \( X - X_k \) where \( X \) is the exact solution of the problem (25) and \( X_k \) is the low-rank approximate solution of (25) obtained at step \( k \) by applying the block Arnoldi algorithm.

Theorem 8: Assume that \( k \) steps of the block Arnoldi Stein algorithm have been run and let \( X_k \) be the obtained low-rank approximation. Then if \( \| A \|_2 \leq 1 \), we have
\[
\| X - X_k \|_F \leq 2 \sqrt{p} \| A \|_2 \| H_{k+1,k} \|_F \| F \|_F - 1 = \| A \|_2^2.
\]

Proof: Subtracting (28) from (25), it follows that the error \( X - X_k \) is the unique solution of the symmetric Stein equation
\[
A(X - X_k) A^T - (X - X_k) = -AX_k A^T_k - A_k X_k A^T.
\]  
(30)
Now since \( \rho(A) < 1 \), the unique solution of (30) is written as
\[
X - X_k = \sum_{i=0}^{\infty} A^i [AX_k A^T_k + A_k X_k A^T] A^T.
\]
Therefore
\[
\| X - X_k \|_2 \leq 2 \| AX_k A^T_k \|_2 \sum_{i=0}^{\infty} \| A \|_2^2.
\]  
(31)
On the other hand if \( G \in \mathbb{R}^{n \times p} \) we have \( \| G \|_2 \leq \| G \|_F \leq \sqrt{p} \| G \|_2 \).

Invoking the expression of \( A_k \) used in Theorem 5.1 and the fact that \( X_k = V_k Z_k V_k^T \), we obtain
\[
\| X_k A^T_k \|_F \leq \| H_{k+1,k} \|_F \| Y_k \|_F.
\]  
(32)
Then using (31) and (32), we obtain the desired result.

VI. NUMERICAL EXAMPLES

The tests reported in this section were run on SUN Microsystems workstations using Matlab. In all our experiments, we divided the matrices \( A \) and \( C \) by \( \| A \|_1 \) and \( \| C \|_1 \) respectively.

We considered the Stein equation
\[
AXC - X = EF^T
\]
where the matrices \( A, C, E \) and \( F \) are of dimension \( n \times n \), \( p \times p \), \( n \times r \) and \( r \times r \) respectively with \( r \ll n, p \).

For all our experiments, the tests where stopped when \( R_k \| F \| \leq 10^{-8} \).
Example 1 For this experiment, we used matrices from Harwell-Boeing Collection: Sherman4 (n = 1104 and nnz(A) = 3786), PDE2961 (n = 2961 and nnz(A) =), Sherman5 (n = 3312 and nnz(A) = 20793) and Fidap009 (n = 3363 and nnz(A) = 99397) where nnz(A) denotes the number of nonzero coefficients in A.

The entries of the matrices E and F were random values uniformly distributed on [0, 1] and we used r = 4.

In Table I, we listed the results obtained with different matrices. A maximum number of itemax = 50 iterations was allowed to the block-Arnoldi Stein algorithm (Algorithm 3). The expression given in Theorem 7 was used to compute the norm of the residual $R_k$ without computing the approximation $X_k$ which is given in a factored form when convergence is achieved.

Example 2 In this experiment, we applied the block-Arnoldi Stein algorithm (Algorithm 3) with matrices A and B defined as follows. The matrix A is generated from the 5-point discretization of the operator

$$L_1(u) = \Delta u - f_1(x,y) \frac{\partial u}{\partial x} - f_2(x,y) \frac{\partial u}{\partial y} - f_3(x,y) u$$

on the unit square $[0, 1] \times [0, 1]$ with homogeneous Dirichlet boundary conditions. We set $f_1(x,y) = e^{x^2+y}$, $f_2(x,y) = 2xy$ and $f_3(x,y) = \cos(xy)$.

The matrix C is also generated from the 5-point discretization of the operator

$$L_2(u) = -\Delta u + g_1(x,y) \frac{\partial u}{\partial x} + g_2(x,y) \frac{\partial u}{\partial y} + g_3(x,y) u$$

on the unit square $[0, 1] \times [0, 1]$ with homogeneous Dirichlet boundary conditions. We set $g_1(x,y) = \sin(x + 2y)$, $g_2(x,y) = e^{xy}$ and $g_3(x,y) = xy$.

The entries of the matrices E and F were random values uniformly distributed on $[0, 1]$. The dimensions of the matrices A and C are $n = n_0^2$ and $p = p_0^2$ respectively, where $n_0$ and $p_0$ are the number of inner grid points in each direction.

For this experiment we used $n = 40,000$, $p = 10,000$, which corresponds to a very large linear system of dimension $4 \times 10^8 \times 4 \times 10^8$. We used different values of r ($r = 5$, $r = 10$, $r = 20$ and $r = 30$). The obtained results are reported in Table II.

VII. CONCLUSION

We proposed in this paper block Krylov subspace methods for solving large and sparse Stein matrix equations. We first considered the case when the right hand side is of full rank. In the second part of the paper, we showed how to apply the block Arnoldi algorithm to derive low-rank approximate solutions to Stein matrix equations with factored right-hand sides. In the two cases, we gave some theoretical results. The numerical examples show that the proposed methods are attractive and could be used for large problems.

REFERENCES