



HAL
open science

Computational Krylov-based methods for large-scale differential Sylvester matrix problems

Mustapha Hached, Khalide Jbilou

► **To cite this version:**

Mustapha Hached, Khalide Jbilou. Computational Krylov-based methods for large-scale differential Sylvester matrix problems. Numerical Linear Algebra with Applications, 2018, 25 (5), pp.1-14 / e2187. 10.1002/nla.2187 . hal-04391873

HAL Id: hal-04391873

<https://hal.science/hal-04391873v1>

Submitted on 12 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

COMPUTATIONAL KRYLOV-BASED METHODS FOR LARGE-SCALE DIFFERENTIAL SYLVESTER MATRIX PROBLEMS

M. HACHED * AND K. JBILOU†

Abstract. In the present paper, we propose Krylov-based methods for solving large-scale differential Sylvester matrix equations having a low rank constant term. We present two new approaches for solving such differential matrix equations. The first approach is based on the integral expression of the exact solution and a Krylov method for the computation of the exponential of a matrix times a block of vectors. In the second approach, we first project the initial problem onto a block (or extended block) Krylov subspace and get a low-dimensional differential Sylvester matrix equation. The latter problem is then solved by some integration numerical methods such as BDF or Rosenbrock method and the obtained solution is used to build the low rank approximate solution of the original problem. We give some new theoretical results such as a simple expression of the residual norm and upper bounds for the norm of the error. Some numerical experiments are given in order to compare the two approaches.

Key words. Extended block Krylov, Krylov subspaces, Low rank, Differential Sylvester equations.

1. Introduction. In the present paper, we consider the differential Sylvester matrix equation (DSE in short) of the form

$$\begin{cases} \dot{X}(t) = A(t)X(t) + X(t)B(t) + E(t)F(t)^T; & (DSE) \\ X(t_0) = X_0, \quad t \in [t_0, T_f], \end{cases} \quad (1.1)$$

where $A(t) \in \mathbb{R}^{n \times n}$, $B(t) \in \mathbb{R}^{p \times p}$ and $E(t) \in \mathbb{R}^{n \times s}$ and $F(t) \in \mathbb{R}^{p \times s}$ are full rank matrices, with $s \ll n, p$. The initial condition is given in a factored form as $X_0 = Z_0 \tilde{Z}_0^T$ and the matrices A and B are assumed to be large and sparse.

Differential Sylvester equations play a fundamental role in many areas such as control, filter design theory, model reduction problems, differential equations and robust control problems [1, 4]. For such differential matrix equations, only a few attempts have been made for large problems.

Let us first recall the following theoretical result which gives an expression of the exact solution of (1.1).

THEOREM 1.1. [1] *The unique solution of the general differential Sylvester equation*

$$\dot{X}(t) = A(t)X(t) + X(t)B(t) + M(t); \quad X(t_0) = X_0 \quad (1.2)$$

is defined by

$$X(t) = \Phi_A(t, 0)X_0 \Phi_{B^T}^T(t, t_0) + \int_{t_0}^t \Phi_A(t, \tau)M(\tau)\Phi_{B^T}^T(t, \tau)d\tau. \quad (1.3)$$

where the transition matrix $\Phi_A(t, t_0)$ is the unique solution to the problem

$$\dot{\Phi}_A(t, t_0) = A(t)\Phi_A(t, t_0), \quad \Phi_A(t_0, t_0) = I.$$

Futhermore, if A is assumed to be a constant matrix, then we have

$$X(t) = e^{(t-t_0)A}X_0e^{(t-t_0)B} + \int_{t_0}^t e^{(t-\tau)A}M(\tau)e^{(t-\tau)B}d\tau. \quad (1.4)$$

We notice that the problem (1.1) is equivalent to the linear ordinary differential equation

$$\begin{cases} \dot{x}(t) &= \mathcal{A}(t)x(t) + b(t) \\ x_0 &= \text{vec}(X_0) \end{cases} \quad (1.5)$$

where $\mathcal{A} = I \otimes A(t) + B^T(t) \otimes I$, $x(t) = \text{vec}(X(t))$ and $b(t) = \text{vec}(E(t)F(t)^T)$, where $\text{vec}(Z)$ is the long vector obtained by stacking the columns of the matrix Z , forming a sole column. For moderate

*Laboratoire P. Painlevé UMR 8524, UFR de Mathématiques, Université des Sciences et Technologies de Lille, IUT A, Rue de la Recherche, BP 179, 59653 Villeneuve d'Ascq Cedex, France; email: mustapha.hached@univ-lille.fr

†L.M.P.A, Université du Littoral Côte d'Opale, 50 rue F. Buisson BP. 699, F-62228 Calais Cedex, France. E-mail: jbilou@univ-littoral.fr.

size problems, it is then possible to directly apply an integration method to solve (1.5). However, this approach is not suitable for large problems. From now on, we assume that the matrices A and B are time independent.

In the present paper, we will consider projection methods onto extended block Krylov (or block Krylov) subspaces associated to the pairs (A, E) and (B^T, F) defined as follows

$$\mathbb{K}_m(A, E) = \text{range}(E, AE, \dots, A^{m-1}E)$$

for block Krylov subspaces, or

$$\mathcal{K}_m(A, E) = \text{range}(A^{-m}E, \dots, A^{-1}E, E, AE, \dots, A^{m-1}E)$$

for extended block Krylov subspaces when the matrix A is nonsingular. Notice that the extended block Krylov subspace $\mathcal{K}_m(A, E)$ is a sum of two block Krylov subspaces associated to the pairs (A, E) and $(A^{-1}, A^{-1}E)$:

$$\mathcal{K}_m(A, E) = \mathbb{K}_m(A, E) + \mathbb{K}_m(A^{-1}, A^{-1}E).$$

To compute an orthonormal basis $\{V_1, \dots, V_m\}$, where V_i is of dimension $n \times d$ where $d = s$ for the block Krylov and $d = 2s$ in the extended block Krylov case, two algorithms have been defined: the first one is the well known block Arnoldi algorithm and the second one is the extended block Arnoldi algorithm [5, 17]; see Appendix A for the description of both algorithms.

These algorithms generate the blocks V_1, V_2, \dots, V_m , $V_i \in \mathbb{R}^{n \times d}$ such that their columns form an orthonormal basis of the block Krylov subspace $\mathbb{K}_m(A, E)$ (with $d = s$) or the extended block Arnoldi $\mathcal{K}_m(A, E)$ (with $d = 2s$).

Both algorithms compute also $d \times d$ block upper Hessenberg matrices $\mathcal{T}_{m,A} = \mathcal{V}_m^T A \mathcal{V}_m$. The following algebraic relations are satisfied

$$A \mathcal{V}_m = \mathcal{V}_{m+1} \widehat{\mathcal{T}}_{m,A}, \quad (1.6)$$

$$= \mathcal{V}_m \mathcal{T}_{m,A} + \mathcal{V}_{m+1} T_{m+1,m} \widetilde{E}_m^T, \quad (1.7)$$

where $\widehat{\mathcal{T}}_{m,A} = \mathcal{V}_{m+1}^T A \mathcal{V}_m$; $T_{i,j}$ is the (i, j) block of $\widehat{\mathcal{T}}_{m,A}$ of size $d \times d$, and $\widetilde{E}_m = [O_{d \times (m-1)d}, I_d]^T$ is the matrix formed with the last d columns of the $md \times md$ identity matrix I_{md} where $d = s$ for the block Arnoldi and $d = 2s$ for the extended block Arnoldi.

When the matrix A is nonsingular and when the computation of the products $W = A^{-1}V$ is not difficult (which is the case for sparse and structured matrices), the use of the extended block Arnoldi is to be preferred.

The paper is organized as follows: In Section 2, we present a first approach based on the approximation of the exponential of a matrix times a block using a Krylov projection method. We give some theoretical results such as a simple expression of the norm of the residual and upper bounds for the norm of the error and perturbation results. In Section 3, the initial differential Sylvester matrix equation is projected onto a block (or extended block) Krylov subspace. The obtained low dimensional differential Sylvester equation is solved by using the well known Backward Differentiation Formula (BDF) and Rosenbrock methods. The last section is devoted to some numerical experiments. Throughout the paper, $\|\cdot\|$ and $\|\cdot\|_F$ will denote the 2-norm and the Frobenius norm, respectively.

2. Solutions via of the matrix exponential approximation . In this section, we give a new approach for computing approximate solutions to large differential Sylvester equations (1.1).

We recall that the exact solution to (1.1) can be expressed as follows

$$X(t) = e^{(t-t_0)A} X_0 e^{(t-t_0)B} + \int_{t_0}^t e^{(t-\tau)A} E F^T e^{(t-\tau)B} d\tau. \quad (2.1)$$

For our first approach, we use this expression of $X(t)$ to obtain low rank approximate solutions. We first approximate the factors $e^{(t-\tau)A} E$ and $e^{(t-\tau)B^T} F$ and then, use a quadrature method to compute the desired approximate solution. As the matrices $e^{(t-\tau)A}$ and $e^{(t-\tau)B^T}$ are large and could be dense even though A and B are sparse, computing the exponential is not recommended. However, in our problem, the computation of $e^{(t-\tau)A}$ and $e^{(t-\tau)B^T}$ are not needed explicitly as we will rather consider the products $e^{(t-\tau)A} E$ and $e^{(t-\tau)B^T} F$ for which approximations via projection methods onto block or extended block Krylov subspaces are well suited.

In what follows, we consider projections onto extended block Krylov (or just block Krylov) subspaces. Let $\mathcal{V}_m = [V_1, \dots, V_m]$ and $\mathcal{W}_m = [W_1, \dots, W_m]$ be the orthogonal matrices whose columns form an

orthonormal basis of the subspace $\mathcal{K}_m(A, E)$ and $\mathcal{K}_m(B^T, F)$, respectively. Following [15, 16, 19], an approximation to $Z_A = e^{(t-\tau)A} E$ can be obtained as

$$Z_{m,A}(\tau) = \mathcal{V}_m e^{(t-\tau)\mathcal{T}_{m,A}} \mathcal{V}_m^T E \quad (2.2)$$

where $\mathcal{T}_{m,A} = \mathcal{V}_m^T A \mathcal{V}_m$. In the same way, an approximation to $e^{(t-\tau)B^T} F$, is given by

$$Z_{m,B}(\tau) = \mathcal{W}_m e^{(t-\tau)\mathcal{T}_{m,B}} \mathcal{W}_m^T F, \quad (2.3)$$

where $\mathcal{T}_{m,B} = \mathcal{W}_m^T B^T \mathcal{W}_m$. Therefore, the integrand in the expression (2.1) can be approximated as

$$e^{(t-\tau)A} E F^T e^{(t-\tau)B} \approx Z_{m,A}(\tau) Z_{m,B}(\tau)^T. \quad (2.4)$$

If for simplicity, we assume that $X(0) = 0$, an approximation to the solution of the differential Sylvester equation (1.1) can be expressed as the product

$$X_m(t) = \mathcal{V}_m G_m(t) \mathcal{W}_m^T, \quad t \in [t_0, T_f], \quad (2.5)$$

where

$$G_m(t) = \int_{t_0}^t Z_{m,A}(\tau) Z_{m,B}^T(\tau) d\tau, \quad (2.6)$$

with $E_m = \mathcal{V}_m^T E$ and $F_m = \mathcal{W}_m^T F$.

The next result shows that the $md \times md$ matrix function G_m is solution of a low-order differential Sylvester matrix equation.

PROPOSITION 2.1. *Let $G_m(t)$ be the matrix function defined by (2.6), then it satisfies the following low-order differential Sylvester matrix equation*

$$\dot{G}_m(t) = \mathcal{T}_{m,A} G_m(t) + G_m(t) \mathcal{T}_{m,B}^T + E_m F_m^T, \quad t \in [t_0, T_f]. \quad (2.7)$$

Proof. The proof can be easily derived from the expression (2.6) and the result of Theorem 1.1.

As a consequence, introducing the residual $R_m(t) = \dot{X}_m(t) - A X_m(t) - X_m(t) B - E F^T$ associated to the approximation $X_m(t)$, we have the following relation

$$\begin{aligned} \mathcal{V}_m^T R_m(t) \mathcal{W}_m &= \mathcal{V}_m^T (\dot{X}_m(t) - A X_m(t) - X_m(t) B - E F^T) \mathcal{W}_m \\ &= \dot{G}_m(t) - \mathcal{T}_{m,A} G_m(t) - G_m(t) \mathcal{T}_{m,B}^T - E_m F_m^T \\ &= 0, \end{aligned}$$

which shows that the residual satisfies a Petrov-Galerkin condition.

As mentioned earlier and for our first exponential-based approach, once $Z_{m,A}(\tau)$ and $Z_{m,B}(\tau)$ are computed, we use a quadrature method to approximate the integral (2.6) in order to get an approximation of $G_m(t)$ and hence to compute $X_m(t)$ from (2.5).

The computation of $X_m(t)$ (and of $R_m(t)$) becomes expensive as m increases. So, in order to stop the iterations, one has to test if $\|R_m(t)\| < \epsilon$ without having to compute extra products involving the matrices A and B . The next result shows how to compute the norm of $R_m(t)$ without forming the approximation $X_m(t)$ which is computed in a factored form only when convergence is achieved.

PROPOSITION 2.2. *Let $X_m(t) = \mathcal{V}_m G_m(t) \mathcal{W}_m^T$ be the approximation obtained at step m by the block (or extended block) Arnoldi method. Then the residual $R_m(t)$ satisfies the relation*

$$\|R_m(t)\|_F^2 = \|T_{m+1,m}^A \bar{G}_m(t)\|_F^2 + \|T_{m+1,m}^B \bar{G}_m(t)\|_F^2, \quad (2.8)$$

and for the 2-norm, we have

$$\|R_m(t)\| = \max\{\|T_{m+1,m}^A \bar{G}_m(t)\|, \|T_{m+1,m}^B \bar{G}_m(t)\|\}, \quad (2.9)$$

where \bar{G}_m is the $d \times md$ matrix corresponding to the last d rows of G_m where $d = 2s$ when using the extended block Arnoldi algorithm and $d = s$ with the block Arnoldi algorithm. **Proof.** . The

proof comes from the fact that the residual $R_m(t)$ can be expressed as

$$R_m(t) = \mathcal{V}_{m+1} \begin{pmatrix} \dot{G}_m(t) - \mathcal{T}_{m,A} G_m(t) - G_m(t) \mathcal{T}_{m,B}^T - E_m F_m^T & -T_{m+1,m}^B \bar{G}_m(t) \\ T_{m+1,m}^A G_m(t) & 0 \end{pmatrix} \mathcal{W}_{m+1}^T, \quad (2.10)$$

where $G_m(t)$ solves the low dimensional problem (2.7). Therefore, we get

$$\begin{aligned}\|R_m(t)\|_F^2 &= \left\| \begin{pmatrix} 0 & -T_{m+1,m}^B \bar{G}_m(t) \\ T_{m+1,m}^A \bar{G}_m(t) & 0 \end{pmatrix} \right\|_F^2 \\ &= \|T_{m+1,m}^A \bar{G}_m(t)\|_F^2 + \|T_{m+1,m}^B \bar{G}_m(t)\|_F^2.\end{aligned}$$

To prove the expression (2.9) with the 2-norm, let us first remark that if

$$M = \begin{pmatrix} 0 & M_1 \\ M_2 & 0 \end{pmatrix}, \text{ then } M^T M = \begin{pmatrix} M_1^T M_1 & 0 \\ 0 & M_2^T M_2 \end{pmatrix},$$

which shows that the singular values of M are the sum of the singular values of M_1 and those of M_2 which implies that

$$\|M\| = \sigma_{max}(M) = \max\{\sigma_{max}(M_1), \sigma_{max}(M_2)\} = \max\{\|M_1\|, \|M_2\|\}.$$

Therefore, using this remark and the fact that

$$\|R_m(t)\| = \left\| \begin{pmatrix} 0 & -T_{m+1,m}^B \bar{G}_m(t) \\ T_{m+1,m}^A \bar{G}_m(t) & 0 \end{pmatrix} \right\|,$$

the result follows.

The approximate solution $X_m(t)$ is computed only when convergence is achieved and in a factored form which is very important for storage requirements in large-scale problems. This procedure is described as follows.

Consider the singular value decomposition of the matrix $G_m(t) = U \Sigma V$ where Σ is the diagonal matrix of the singular values of $G_m(t)$ sorted in decreasing order. Let U_l be the $md \times l$ matrix of the first l columns of U corresponding to the l singular values of magnitude greater than some tolerance $dtol$. We obtain the truncated singular value decomposition $G_m(t) \approx U_l \Sigma_l V_l^T$ where $\Sigma_l = \text{diag}[\lambda_1, \dots, \lambda_l]$. Setting $\tilde{Z}_{m,A}(t) = \mathcal{V}_m U_l \Sigma_l^{1/2}$ and $\tilde{Z}_{m,B}(t) = \mathcal{W}_m V_l \Sigma_l^{1/2}$, it follows that

$$X_m(t) \approx \tilde{Z}_{m,A}(t) \tilde{Z}_{m,B}(t)^T. \quad (2.11)$$

Therefore, only the matrices $\tilde{Z}_{m,A}(t)$ and $\tilde{Z}_{m,B}(t)$ are needed.

The following result shows that the approximation X_m is an exact solution of a perturbed differential Sylvester equation.

PROPOSITION 2.3. *Let $X_m(t)$ be the approximate solution given by (2.5). Then we have*

$$\dot{X}_m(t) = (A - F_{m,A}) X_m(t) + X_m(t) (B - F_{m,B}) + EF^T. \quad (2.12)$$

where $F_{m,A} = V_{m+1} T_{m+1,m}^A V_m^T$ and $F_{m,B} = W_m (T_{m+1,m}^B)^T W_{m+1}^T$

Proof. . As $X_m(t) = \mathcal{V}_m G_m(t) \mathcal{W}_m^T$, we have

$$\dot{X}_m(t) - (AX_m(t) + X_m(t)B + EF^T) = \mathcal{V}_m \dot{G}_m(t) \mathcal{W}_m^T - (A \mathcal{V}_m G_m(t) \mathcal{W}_m^T + B \mathcal{V}_m G_m(t) \mathcal{W}_m^T + EF^T). \quad (2.13)$$

Now, using the fact that

$$A \mathcal{V}_m = \mathcal{V}_m \mathcal{T}_{m,A} + V_{m+1} T_{m+1,m}^A \tilde{E}_m^T, \text{ and } B^T \mathcal{W}_m = \mathcal{W}_m \mathcal{T}_{m,B} + W_{m+1} T_{m+1,m}^B \tilde{E}_m^T,$$

equation (2.13) becomes

$$\begin{aligned}\dot{X}_m(t) - (AX_m(t) + X_m(t)B + EF^T) &= \mathcal{V}_m \dot{G}_m(t) \mathcal{W}_m^T - ([\mathcal{V}_m \mathcal{T}_{m,A} + V_{m+1} T_{m+1,m} \tilde{E}_m^T] G_m(t) \mathcal{W}_m^T \\ &\quad + \mathcal{V}_m G_m(t) [\mathcal{W}_m \mathcal{T}_{m,B} + W_{m+1} T_{m+1,m}^B \tilde{E}_m^T]^T + EF^T).\end{aligned}$$

Therefore

$$\begin{aligned}\dot{X}_m(t) - (AX_m(t) + X_m(t)B + EF^T) &= \mathcal{V}_m [\dot{G}_m(t) - \mathcal{T}_{m,A} G_m(t) - G_m(t) \mathcal{T}_{m,B}^T - EF^T] \mathcal{W}_m \\ &\quad - (V_{m+1} T_{m+1,m}^A \tilde{E}_m^T G_m(t) \mathcal{W}_m + \mathcal{V}_m G_m(t) \tilde{E}_m (T_{m+1,m}^B)^T W_{m+1}^T).\end{aligned}$$

On the other hand we have $\mathcal{V}_m G_m(t) = X_m(t) \mathcal{W}_m$, $G_m(t) \mathcal{W}_m^T = \mathcal{V}_m^T X_m(t)$, $\mathcal{V}_m \tilde{E}_m = V_m$, $\mathcal{W}_m \tilde{E}_m = W_m$ and $EF^T = \mathcal{V}_m E_m F_m^T \mathcal{W}_m^T$. So using these relations and the fact that G_m solves the low dimensional differential Sylvester equation (2.7), we obtain the desired result.

The next result states that the error $\mathcal{E}_m(t) = X(t) - X_m(t)$ satisfies also a differential Sylvester matrix equation.

PROPOSITION 2.4. *Let $X(t)$ be the exact solution of (1.1) and let $X_m(t)$ be the approximate solution obtained at step m . The error $\mathcal{E}_m(t) = X(t) - X_m(t)$ satisfies the following equation*

$$\dot{\mathcal{E}}_m(t) - A\mathcal{E}_m(t) - \mathcal{E}_m(t)B = F_{m,A}X_m(t) + X_m(t)F_{m,B} = -R_m(t), \quad (2.14)$$

where $F_{m,A}$ and $F_{m,B}$ are defined in Proposition 2.3 and $R_m(t) = \dot{X}_m(t) - AX_m(t) - X_m(t)B - EF^T$. **Proof.** The result is easily obtained by subtracting the equation (2.12) from the initial differential Sylvester equation (1.1).

Notice that from Proposition 2.4, the error $\mathcal{E}_m(t)$ can be expressed in the integral form as follows

$$\mathcal{E}_m(t) = e^{(t-t_0)A}\mathcal{E}_{m,0}e^{(t-t_0)B} - \int_{t_0}^t e^{(t-\tau)A}R_m(\tau)e^{(t-\tau)B}d\tau, \quad t \in [t_0, T_f]. \quad (2.15)$$

where $\mathcal{E}_{m,0} = \mathcal{E}_m(0)$.

Next, we give an upper bound for the norm of the error by using the 2-logarithmic norm defined by $\mu_2(A) = \lim_{h \rightarrow 0^+} \frac{\|I + hA\|_2 - 1}{h} = \frac{1}{2}\lambda_{\max}(A + A^T)$.

PROPOSITION 2.5. *Assume that the matrices A and B are such that $\mu_2(A) + \mu_2(B) \neq 0$. Then at step m of the extended block Arnoldi (or block Arnoldi) process, we have the following upper bound for the norm of the error $\mathcal{E}_m(t) = X(t) - X_m(t)$,*

$$\|\mathcal{E}_m(t)\| \leq \|\mathcal{E}_{m,0}\|e^{(t-t_0)(\mu_2(A)+\mu_2(B))} + \alpha_m \frac{e^{(t-t_0)(\mu_2(A)+\mu_2(B))} - 1}{\mu_2(A) + \mu_2(B)}, \quad (2.16)$$

where α_m is given by $\alpha_m = \max_{\tau \in [t_0, t]} \left(\max\{\|T_{m+1,m}^A \bar{G}_m(\tau)\|_2, \|T_{m+1,m}^B \bar{G}_m(\tau)\|_2\} \right)$. The matrix \bar{G}_m is the $d \times md$ matrix corresponding to the last d rows of G_m . **Proof.** We first point out that $\|e^{tA}\| \leq e^{\mu_2(A)t}$. Using the expression (2.15) of $\mathcal{E}_m(t)$, we obtain the following relation

$$\|\mathcal{E}_m(t)\| \leq \|e^{(t-t_0)A}\mathcal{E}_{m,0}e^{(t-t_0)B}\| + \int_{t_0}^t \|e^{(t-\tau)A}R_m(\tau)e^{(t-\tau)B}\|d\tau.$$

Therefore, using (2.15) and the fact that $\|e^{(t-\tau)A}\| \leq e^{(t-\tau)\mu_2(A)}$, we get

$$\begin{aligned} \|\mathcal{E}_m(t)\| &\leq \|\mathcal{E}_{m,0}\|e^{(t-t_0)(\mu_2(A)+\mu_2(B))} + \max_{\tau \in [t_0, t]} \|R_m(\tau)\| \int_{t_0}^t e^{(t-\tau)\mu_2(A)} e^{(t-\tau)\mu_2(B)} d\tau \\ &= \|\mathcal{E}_{m,0}\|e^{(t-t_0)(\mu_2(A)+\mu_2(B))} + \max_{\tau \in [t_0, t]} \|R_m(\tau)\| e^{t(\mu_2(A)+\mu_2(B))} \int_{t_0}^t e^{-\tau(\mu_2(A)+\mu_2(B))} d\tau. \end{aligned}$$

Hence

$$\|\mathcal{E}_m(t)\| \leq \|\mathcal{E}_{m,0}\|e^{(t-t_0)(\mu_2(A)+\mu_2(B))} + \max_{\tau \in [t_0, t]} \|R_m(\tau)\| \frac{e^{(t-t_0)(\mu_2(A)+\mu_2(B))} - 1}{\mu_2(A) + \mu_2(B)}. \quad (2.17)$$

Using the result of Proposition 2.2, we obtain $\max_{\tau \in [t_0, t]} \|R_m(\tau)\| = \alpha_m$ and then

$$\|\mathcal{E}_m(t)\| \leq \|\mathcal{E}_{m,0}\|e^{(t-t_0)(\mu_2(A)+\mu_2(B))} + \alpha_m \frac{e^{(t-t_0)(\mu_2(A)+\mu_2(B))} - 1}{\mu_2(A) + \mu_2(B)}.$$

Notice that if the matrices A and B are stable (ie all the eigenvalues are in the open half plane) then $\mu_2(A) < 0$ and $\mu_2(B) < 0$ which ensures the condition of Proposition 2.5 is satisfied. Notice also that since $R_m(\tau) = -F_{m,A}X_m(\tau) - X_m(\tau)F_{m,B}$, where $F_{m,A} = V_{m+1}^A T_{m+1,m}^A V_m^T$ and $F_{m,B} = W_m(T_{m+1,m}^B)^T W_{m+1}^T$, we get

$$\|R_m(\tau)\| \leq \max_{\tau \in [t_0, t]} \|\bar{G}_m(\tau)\| \left(\|T_{m+1,m}^A\| + \|T_{m+1,m}^B\| \right).$$

Hence, replacing in (2.17), we get the new upper bound

$$\|\mathcal{E}_m(t)\| \leq \|\mathcal{E}_{m,0}\|e^{(t-t_0)(\mu_2(A)+\mu_2(B))} + \beta_m \frac{e^{(t-t_0)(\mu_2(A)+\mu_2(B))} - 1}{\mu_2(A) + \mu_2(B)}, \quad (2.18)$$

where

$$\beta_m = \max_{\tau \in [t_0, t]} \|\tilde{G}_m(\tau)\| \left(\|T_{m+1,m}^A\| + \|T_{m+1,m}^B\| \right).$$

In Figure 2.1, we compare the computed error to the two error upper bounds given by Formulae (2.16) and (2.18) for A and B being two 100×100 matrices obtained by the finite differences discretization of linear differential operators on the unit square $[0, 1] \times [0, 1]$ with homogeneous Dirichlet boundary conditions. Matrices E and F were chosen as rank 2 matrices which entries are randomly generated over the interval $[0, 1]$. In order to compute the error, we took the approximate solution given by the integral form of the solution as a reference.

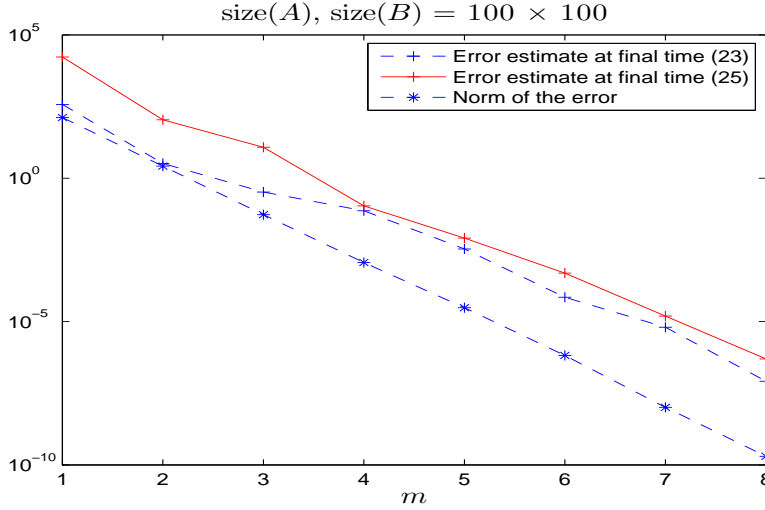


FIG. 2.1. Norm of the error vs number of Arnoldi iterations m

We observe that the bound (2.16) stated in Proposition 2.5 is slightly better in this example. Next, we give another upper bound for the norm of the error $\mathcal{E}_m(t)$.

PROPOSITION 2.6. *Let $X(t)$ be the exact solution to (1.1) and let $X_m(t)$ be the approximate solution obtained at step m . Then we have*

$$\|\mathcal{E}_m(t)\| \leq \|F\| e^{t\mu_2(B)} \Gamma_{1,m}(t) + \|E_m\| e^{t\mu_2(A)} \Gamma_{2,m}(t), \quad (2.19)$$

where

$$\Gamma_{1,m}(t) = \int_{t_0}^t e^{-\tau\mu_2(B)} \|Z_A(\tau) - Z_{m,A}(\tau)\| d\tau, \quad \Gamma_{2,m}(t) = \int_{t_0}^t e^{-\tau\mu_2(A)} \|Z_B(\tau) - Z_{m,B}(\tau)\| d\tau.$$

Proof. From the expressions of $X(t)$ and $X_m(t)$, we have

$$\|\mathcal{E}_m(t)\| = \left\| \int_{t_0}^t \left(Z_A(\tau) Z_B(\tau)^T - Z_{m,A}(\tau) Z_{m,B}(\tau)^T \right) d\tau \right\|, \quad (2.20)$$

where $Z_{m,A} = \mathcal{V}_m e^{(t-\tau)\mathcal{T}_{m,A}} E_m$, $Z_{m,B}(\tau) = \mathcal{W}_m e^{(t-\tau)\mathcal{T}_{m,B}} F_m$, $Z_A(\tau) = e^{(t-\tau)A} E$ and $Z_B(\tau) = e^{(t-\tau)B} F$. Then, using the relation

$$Z_A(\tau) Z_B(\tau)^T - Z_{m,A}(\tau) Z_{m,B}(\tau)^T = (Z_A(\tau) - Z_{m,A}(\tau)) Z_B^T + Z_{m,A}(\tau) (Z_B(\tau) - Z_{m,B}(\tau))^T,$$

we obtain

$$\begin{aligned} \|Z_A(\tau) Z_B(\tau)^T - Z_{m,A}(\tau) Z_{m,B}(\tau)^T\| &\leq \|Z_B(\tau)\| \|Z_A(\tau) - Z_{m,A}(\tau)\| \\ &\quad + \|Z_{m,A}(\tau)\| \|Z_B(\tau) - Z_{m,B}(\tau)\|. \end{aligned}$$

Now as $\|Z_B(\tau)\| \leq e^{(t-\tau)\mu_2(B)}\|F\|$ and since $\mu_2(\mathcal{T}_{m,A}) \leq \mu_2(A)$, we also have $\|Z_{m,A}(\tau)\| \leq e^{(t-\tau)\mu_2(\mathcal{T}_{m,A})}\|E_m\| \leq e^{(t-\tau)\mu_2(A)}\|E_m\|$. Using all these relations in (2.20), we get

$$\begin{aligned} \|\mathcal{E}_m(t)\| &\leq \int_{t_0}^t \left[e^{(t-\tau)\mu_2(B)}\|F\|\|Z_A(\tau) - Z_{m,A}(\tau)\| + e^{(t-\tau)\mu_2(A)}\|E_m\|\|Z_B(\tau) - Z_{m,B}(\tau)\| \right] d\tau \\ &\leq \|F\|e^{t\mu_2(B)} \int_{t_0}^t e^{-\tau\mu_2(B)}\|Z_A(\tau) - Z_{m,A}(\tau)\|d\tau \\ &\quad + \|E_m\|e^{t\mu_2(A)} \int_{t_0}^t e^{-\tau\mu_2(A)}\|Z_B(\tau) - Z_{m,B}(\tau)\|d\tau, \end{aligned}$$

which ends the proof.

One can use some known results [12, 16] to derive upper bounds for $\|Z_A(\tau) - Z_{m,A}(\tau)\|$ and $\|Z_B(\tau) - Z_{m,B}(\tau)\|$, when using Krylov or block Krylov subspaces. For general matrices A and B , we can use the following result to get upper bounds for $\|Z_A(\tau) - Z_{m,A}(\tau)\|$ and $\|Z_B(\tau) - Z_{m,B}(\tau)\|$.

PROPOSITION 2.7. *When using the extended block Arnoldi (or the block Arnoldi), we get the following upper bound for the exponential approximation error $e_{m,A}(\tau) = Z_A(\tau) - Z_{m,A}(\tau)$:*

$$\|e_{m,A}(\tau)\| \leq \|T_{m+1,m}^A\| \int_0^\tau e^{(u-\tau)\nu(A)}\|L_{m,A}(u)\|du, \quad (2.21)$$

where $L_{m,A}(u) = \tilde{E}_m e^{(t-u)\mathcal{T}_{m,A}} E_m$ and $\nu(A) = \lambda_{\min}\left(\frac{A+A^T}{2}\right)$. **Proof.** We have

$$Z_A(\tau) = e^{(t-\tau)A}E, \quad \text{and} \quad Z_{m,A}(\tau) = \mathcal{V}_m e^{(t-\tau)\mathcal{T}_{m,A}} E_m.$$

Then $Z'_A(\tau) = -Ae^{(t-\tau)A}E = -AZ_A(\tau)$, and

$$Z'_{m,A}(\tau) = -\mathcal{V}_m \mathcal{T}_{m,A} e^{(t-\tau)\mathcal{T}_{m,A}} E_m = -[A\mathcal{V}_m - V_{m+1}T_{m+1,m}^A \tilde{E}_m] e^{(t-\tau)\mathcal{T}_{m,A}} E_m.$$

Hence,

$$Z'_{m,A}(\tau) = -AZ_{m,A}(\tau) + V_{m+1}T_{m+1,m}^A L_{m,A}(\tau), \quad (2.22)$$

where $L_{m,A}(\tau) = \tilde{E}_m e^{(t-\tau)\mathcal{T}_{m,A}} E_m$.

Therefore, the error $e_{m,A}(\tau) = Z_A(\tau) - Z_{m,A}(\tau)$ is such that

$$e'_{m,A}(\tau) = -Ae_m(\tau) - V_{m+1}T_{m+1,m}^A L_{m,A}(\tau),$$

which gives the following expression of e_m :

$$e_{m,A}(\tau) = - \int_0^\tau e^{(u-\tau)A} V_{m+1} T_{m+1,m}^A L_{m,A}(u) du. \quad (2.23)$$

On the other hand, since $\tau - u > 0$, it follows that

$$\|e^{(u-\tau)A}\| \leq e^{(\tau-u)\mu_2(-A)} = e^{(u-\tau)\nu(A)}.$$

Then, we get

$$\|e_{m,A}(\tau)\| \leq \|T_{m+1,m}^A\| \int_0^\tau e^{(u-\tau)\nu(A)}\|L_{m,A}(u)\|du.$$

Notice that if $\nu(A)$ is not known but $\nu(A) \geq 0$ (which is the case for positive semidefinite matrices) then we get the upper bound

$$\|e_{m,A}(\tau)\| \leq \|T_{m+1,m}^A\| \int_0^\tau \|L_{m,A}(u)\|du. \quad (2.24)$$

To define a new upper bound for the norm of the global error $\mathcal{E}_m(t)$, we can use the upper bounds for the errors $e_{m,A}$ and $e_{m,B}$ in the expression (2.19) stated in Proposition 2.6 to get

$$\begin{aligned} \|\mathcal{E}_m(t)\| &\leq \|F\|e^{t\mu_2(B)} \int_{t_0}^t e^{-\tau\mu_2(B)}\|e_{m,A}(\tau)\|d\tau \\ &\quad + \|E_m\|e^{t\mu_2(A)} \int_{t_0}^t e^{-\tau\mu_2(A)}\|e_{m,B}(\tau)\|d\tau, \end{aligned}$$

and then we obtain

$$\|\mathcal{E}_m(t)\| \leq \|F\| e^{t\mu_2(B)} \|T_{m+1,m}^A\| \int_{t_0}^t e^{-\tau\mu_2(B)} S_{m,A}(\tau) d\tau \quad (2.25)$$

$$+ \|E_m\| e^{t\mu_2(A)} \|T_{m+1,m}^B\| \int_{t_0}^t e^{-\tau\mu_2(A)} S_{m,B}(\tau) d\tau, \quad (2.26)$$

where $S_{m,A}(\tau) = \int_0^\tau e^{(u-\tau)\nu(A)} \|L_{m,A}(u)\| du$ and $S_{m,B}(\tau) = \int_0^\tau e^{(u-\tau)\nu(B)} \|L_{m,B}(u)\| du$.

As m is generally very small as compared to n and p , the factors $L_{m,A}$ and $L_{m,B}$ can be computed using Matlab functions such as `expm` and the integral appearing in the right sides of (2.21) and (2.25), can be approximated via a quadrature formulae.

We summarize the steps of our proposed first approach (using the extended block Arnoldi) in the following algorithm

Algorithm 1 The extended block Arnoldi (EBA-exp) method for DSE's

- Input $X_0 = X(t_0)$, a tolerance $tol > 0$, an integer m_{max} .
 - For $m = 1, \dots, m_{max}$
 - Apply the extended block Arnoldi algorithm to (A, E) and (B^T, F) to get the orthonormal matrices $\mathcal{V}_m = [V_1, \dots, V_m]$ and $\mathcal{W}_m = [W_1, \dots, W_m]$ and the upper block Hessenberg matrices $\mathcal{T}_{m,A}$ and $\mathcal{T}_{m,B}$.
 - Set $E_m = \mathcal{V}_m^T E$, $F_m = \mathcal{W}_m^T F$ and compute $Z_{m,A}(\tau) = e^{(t-\tau)\mathcal{T}_{m,A}} E_m$ and $Z_{m,B}(\tau) = e^{(t-\tau)\mathcal{T}_{m,B}} F_m$ using the matlab function `expm`.
 - Use a quadrature method to compute the integral (2.6) and get an approximation of $G_m(t)$ for each $t \in [t_0, T_f]$.
 - If $\|R_m(t)\| = \max\{\|T_{m+1,m}^A \tilde{G}_m(t)\|, \|T_{m+1,m}^B \tilde{G}_m(t)\|\} < tol$ stop and compute the approximate solution $X_m(t)$ in the factored form given by the relation (2.11).
 - End
-

3. Projecting and solving the low dimensional problem.

3.1. Low-rank approximate solutions. In this section, we show how to obtain low rank approximate solutions to the differential Sylvester equation (1.1) by first projecting directly the initial problem onto block (or extended block) Krylov subspaces and then solve the obtained low dimensional differential problem. We first apply the block Arnoldi algorithm (or the extended block Arnoldi) to the pairs (A, E) and (B^T, F) to get the orthonormal matrices \mathcal{V}_m and \mathcal{W}_m , whose columns form orthonormal bases of the extended block Krylov subspaces $\mathcal{K}_m(A, E)$ and $\mathcal{K}_m(B^T, F)$, respectively. We also get the upper block Hessenberg matrices $\mathcal{T}_{m,A} = \mathcal{V}_m^T A \mathcal{V}_m$ and $\mathcal{T}_{m,B} = \mathcal{W}_m^T B^T \mathcal{W}_m$. Let $X_m(t)$ be the desired low rank approximate solution given as

$$X_m(t) = \mathcal{V}_m Y_m(t) \mathcal{W}_m^T, \quad (3.1)$$

satisfying the Petrov-Galerkin orthogonality condition

$$\mathcal{V}_m^T R_m(t) \mathcal{W}_m = 0, \quad t \in [t_0, T_f], \quad (3.2)$$

where $R_m(t)$ is the residual $R_m(t) = \dot{X}_m(t) - A X_m(t) - X_m(t) B - E F^T$. Then, from (3.1) and (3.2), we obtain the low dimensional differential Sylvester equation

$$\dot{Y}_m(t) - \mathcal{T}_{m,A} Y_m(t) - Y_m(t) \mathcal{T}_{m,B}^T - E_m F_m^T = 0, \quad (3.3)$$

where $E_m = \mathcal{V}_m^T E$ and $F_m = \mathcal{W}_m^T F$. The obtained low dimensional differential Sylvester equation (3.3) is the same as the one given by (2.7). We have now to solve the latter differential equation by some integration method such as the well known Backward Differentiation Method (BDF) [3] or the Rosenbrock method [3, 18].

Notice that all the properties and results such as the expressions of the residual norms or the upper bounds for the norm of the error given in the last section are still valid with this second approach. The two approaches only differ in the way the projected low dimensional differential Sylvester matrix equations are numerically solved.

3.2. BDF for solving the low order differential Sylvester equation (3.3).

We use the Backward Differentiation Formula (BDF) method for solving, at each step m of the extended block Arnoldi (or block Arnoldi) process, the low dimensional differential Sylvester matrix equation (3.3). We notice that BDF is especially used for the solution of stiff differential equations. At each time t_k , let $Y_{m,k}$ of the approximation of $Y_m(t_k)$, where Y_m is a solution of (3.3). Then, the new approximation $Y_{m,k+1}$ of $Y_m(t_{k+1})$ obtained at step $k+1$ by BDF is defined by the implicit relation

$$Y_{m,k+1} = \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i} + h_k \beta \mathcal{F}(Y_{m,k+1}), \quad (3.4)$$

where $h_k = t_{k+1} - t_k$ is the step size, α_i and β_i are the coefficients of the BDF method as listed in Table 3.1 and $\mathcal{F}(Y)$ is given by

$$\mathcal{F}(Y) = \mathcal{T}_{m,A} Y + Y \mathcal{T}_{m,B}^T + E_m F_m^T.$$

p	β	α_0	α_1	α_2
1	1	1		
2	2/3	4/3	-1/3	
3	6/11	18/11	-9/11	2/11

TABLE 3.1

Coefficients of the p -step BDF method with $p \leq 3$.

The approximate $Y_{m,k+1}$ solves the following matrix equation

$$-Y_{m,k+1} + h_k \beta (\mathcal{T}_{m,A} Y_{m,k+1} + Y_{m,k+1} \mathcal{T}_{m,B}^T + E F^T) + \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i} = 0,$$

which can be written as the following Sylvester matrix equation

$$\mathbb{T}_{m,A} Y_{m,k+1} + Y_{m,k+1} \mathbb{T}_{m,B}^T + \mathbb{E}_{m,k} \mathbb{F}_{m,k}^T = 0. \quad (3.5)$$

We assume that at each time t_k , the approximation $Y_{m,k}$ is factorized as a low rank product $Y_{m,k} \approx \tilde{U}_{m,k} \tilde{V}_{m,k}^T$, where $\tilde{U}_{m,k} \in \mathbb{R}^{n \times m_k}$ and $\tilde{V}_{m,k} \in \mathbb{R}^{p \times m_k}$, with $m_k \ll n, p$. In that case, the coefficient matrices appearing in (3.5) are given by

$$\mathbb{T}_{m,A} = h_k \beta \mathcal{T}_{m,A} - \frac{1}{2} I; \quad \mathbb{T}_{m,B} = h_k \beta \mathcal{T}_{m,B} - \frac{1}{2} I,$$

$$\mathbb{E}_{m,k+1} = [\sqrt{h_k \beta} E^T, \sqrt{\alpha_0} \tilde{U}_{m,k}^T, \dots, \sqrt{\alpha_{p-1}} \tilde{U}_{m,k+1-p}^T]^T$$

and

$$\mathbb{F}_{m,k+1} = [\sqrt{h_k \beta} F^T, \sqrt{\alpha_0} \tilde{V}_{m,k}^T, \dots, \sqrt{\alpha_{p-1}} \tilde{V}_{m,k+1-p}^T]^T.$$

The Sylvester matrix equation (3.5) can be solved by applying direct methods based on Schur decomposition such as the Bartels-Stewart algorithm [2, 9].

Notice that we can also use the BDF method applied directly to the original problem (1.1) and then at each iteration, one has to solve large Sylvester matrix equations which can be done by using Krylov-based methods as developed in [6, 13].

3.3. Solving the low dimensional problem with the Rosenbrock method.

Applying Rosenbrock method [3, 18] to the low dimensional differential Sylvester matrix equation (3.3), the new approximation $Y_{m,k+1}$ of $Y_m(t_{k+1})$ obtained at step $k+1$ is defined, in the ROS(2) particular case by the relations

$$Y_{m,k+1} = Y_{m,k} + \frac{3}{2} K_1 + \frac{1}{2} K_2, \quad (3.6)$$

where K_1 and K_2 solve the following Sylvester equations

$$\tilde{\mathbb{T}}_{m,A} K_1 + K_1 \tilde{\mathbb{T}}_{m,B} = -\mathcal{F}(t_k, Y_{m,k}), \quad (3.7)$$

and

$$\tilde{\mathbb{T}}_{m,A}K_2 + K_2\tilde{\mathbb{T}}_{m,B} = -\mathcal{F}(t_{k+1}, Y_{m,k} + K_1) + \frac{2}{h}K_1, \quad (3.8)$$

where

$$\tilde{\mathbb{T}}_{m,A} = \gamma\mathcal{T}_{m,B} - \frac{1}{2h}I \quad \text{and} \quad \tilde{\mathbb{T}}_{m,B} = \gamma\mathcal{T}_{m,B}^T - \frac{1}{2h}I,$$

and

$$\mathcal{F}(Y) = \mathcal{T}_{m,A}Y + Y\mathcal{T}_{m,B}^T + E_mF_m^T.$$

We summarize the steps of the second approach (using the extended block Arnoldi) in the following algorithm

Algorithm 2 The extended block Arnoldi (EBA) method for DSE's

- Input $X_0 = X(t_0)$, a tolerance $tol > 0$, an integer m_{max} .
- For $m = 1, \dots, m_{max}$
 - Apply the extended block Arnoldi algorithm to the pairs (A, E) and (B^T, F) to compute the orthonormal bases $\mathcal{V}_m = [V_1, \dots, V_m]$ and $\mathcal{W}_m = [W_1, \dots, W_m]$ and also the the upper block Hessenberg matrices $\mathcal{T}_{m,A}$ and $\mathcal{T}_{m,B}$.
 - Use the BDF or the Rosenbrock method to solve the low dimensional differential Sylvester equation

$$\dot{Y}_m(t) - \mathcal{T}_{m,A}Y_m(t) - Y_m(t)\mathcal{T}_{m,B}^T - E_mF_m^T = 0, \quad t \in [t_0, T_f]$$

- If $\|R_m(t)\| < tol$ stop and compute the approximate solution $X_m(t)$ in the factored form given by the relation (2.11).
 - End
-

4. Numerical examples. In this section, we compare the approaches presented in this paper. The exponential approach (EBA-exp) summarized in Algorithm 1, which is based on the approximation of the solution to (1.1) applying a quadrature method to compute the projected exponential form solution (2.6). We used a scaling and squaring strategy, implemented in the MATLAB `expm` function; see [11, 14] for more details. The second method (Algorithm 2) is based on the BDF integration method applied to the projected Sylvester equation as described in Section (3.2). Finally, we considered the EBA-ROS(2) method as described in Section (3.3). The basis of the projection subspaces were generated by the extended block Arnoldi algorithm for all methods. All the experiments were performed on a laptop with an Intel Core i7 processor and 8GB of RAM. The algorithms were coded in Matlab R2014b.

Example 1. For this example, the matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{p \times p}$ were obtained from the 5-point discretization of the operators

$$L_A = \Delta u - f_1(x, y) \frac{\partial u}{\partial x} + f_2(x, y) \frac{\partial u}{\partial y} + g_1(x, y),$$

and

$$L_B = \Delta u - f_3(x, y) \frac{\partial u}{\partial x} + f_4(x, y) \frac{\partial u}{\partial y} + g_2(x, y),$$

on the unit square $[0, 1] \times [0, 1]$ with homogeneous Dirichlet boundary conditions. The number of inner grid points in each direction are n_0 for A and p_0 for B and the dimension of the matrices A and B are $n = n_0^2$ and $p = p_0^2$ respectively. Here we set $f_1(x, y) = x + 10y^2$, $f_2(x, y) = \sqrt{2x^2 + y^2}$, $f_3(x, y) = x + 2y$, $f_4(x, y) = \exp(y - x)$, $g_1(x, y) = x^2 - y^2$ and $g_2(x, y) = y^2 - x^2$. The time interval considered was $[0, 2]$ and the initial condition $X_0 = X(0)$ was $X_0 = Z_0 Z_0^T$, where $Z_0 = 0_{n \times 2}$.

For all projection-based methods, we used projections onto the Extended Block Krylov subspaces $\mathcal{K}_k(A, B) = \text{Range}(B, AB, \dots, A^{m-1}B, A^{-1}B, \dots, (A^{-1})^m B)$ and the tolerance was set to 10^{-10} for the stop test on the residual. For the EBA-BDF and Rosenbrock methods, we used a constant timestep h . The entries of the matrices E and F were random values uniformly distributed on the interval $[0, 1]$ and their rank were set to $s = 2$.

To the authors' knowledge, there are no available exact solutions of large scale matrix Sylvester differential equations in the literature. In order to check if our approaches produce reliable results, we first compared our results to the one given by Matlab's `ode23s` solver which is designed for stiff differential equations. This was done by vectorizing our DSE, stacking the columns of X one on top of each other. This method is not suited to large-scale problems. Due to the memory limitation of our computer when running the `ode23s` routine, we chose a size of 100×100 for the matrices A and B .

In Figure 4.1, we compared the component X_{11} of the solution obtained by the methods tested in this section, to the solution provided by the `ode23s` method from Matlab, on the time interval $[0, 2]$, for $\text{size}(A), \text{size}(B) = 100 \times 100$ and a constant timestep $h = 10^{-2}$. We observe that all the considered methods give good results in terms of accuracy. The relative error norms at final time $T_f = 2$ were of order $\mathcal{O}(10^{-10})$ for the EBA-exp method and $\mathcal{O}(10^{-12})$ for the others. The runtimes were respectively 0.6s for EBA-exp, 7.3s for EBA-BDF(1), 20.8s for EBA-BDF(2) and 29.2s for EBA-ROS(2). The `ode23s` routine required 978s. In Table 4.1, we give the obtained runtimes in seconds, the number of Arnoldi iterations and the Frobenius residual norm at final time, for the resolution of Equation (1.1) for $t \in [0, 2]$, with a timestep $h = 0.01$. The results in Table 4.1 show

	EBA-exp	EBA-BDF(1)	EBA-BDF(2)	EBA-ROS(2)
$n, p = 2500, 2500$	3.8s ($m = 16$)	6.1s ($m = 18$)	13.6s ($m = 18$)	28.8s ($m = 23$)
$\ R_m(T_f)\ _F$	1.04×10^{-8}	2.45×10^{-10}	2.45×10^{-10}	3.05×10^{-10}
$n, p = 10000, 10000$	35.2s ($m = 22$)	38.4s ($m = 25$)	80.3s ($m = 25$)	104.7s ($m = 33$)
$\ R_m(T_f)\ _F$	4.4×10^{-9}	4.1×10^{-11}	4.2×10^{-11}	5.8×10^{-11}
$n, p = 22500, 10000$	137.3s ($m = 22$)	166.5s ($m = 30$)	342.3s ($m = 30$)	246s ($m = 35$)
$\ R_m(T_f)\ _F$	4.2×10^{-8}	3.7×10^{-11}	3.6×10^{-11}	1.78×10^{-9}

TABLE 4.1
Runtimes in seconds and the residual norms

that the EBA-exp method is outperformed by the other approaches in terms of accuracy, although

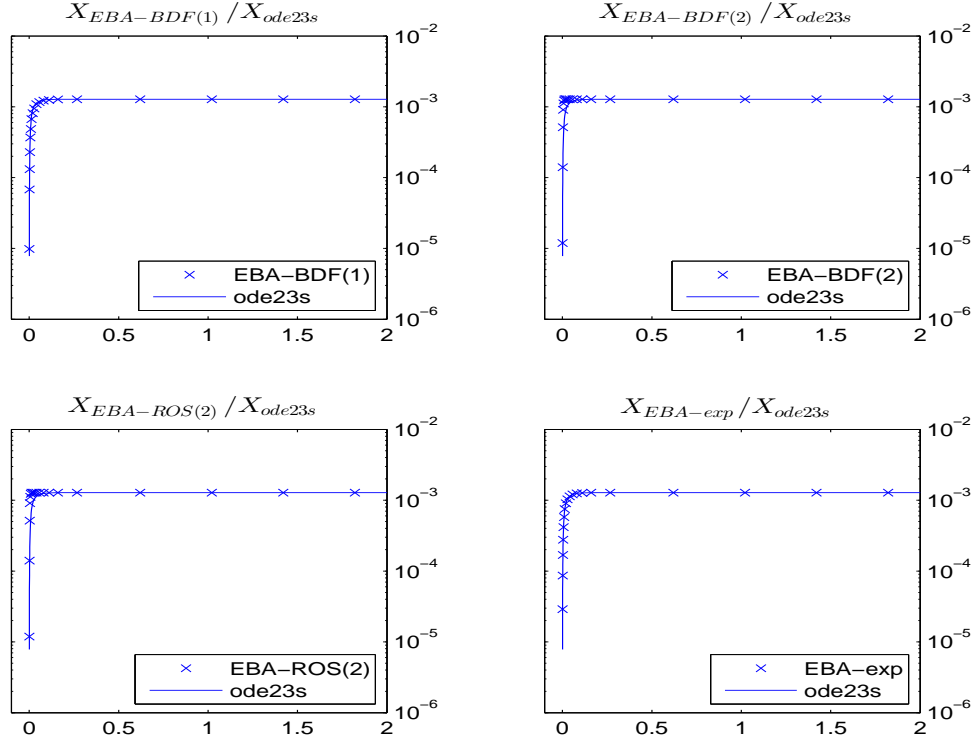


FIG. 4.1. Values of $X_{11}(t)$ for $t \in [0, 2]$

it allows to obtain an acceptable approximation more quickly. The EBA-BDF(1) appears to be the better option in terms of time and accuracy.

Example 2 In this second example, we considered the particular case

$$\begin{cases} \dot{X}(t) = A(t)X(t) + X(t)A(t) - E(t)F(t)^T; & (DSE) \\ X(t_0) = X_0, & t \in [t_0, T_f], \end{cases} \quad (4.1)$$

where the matrix $A = \text{Rail1357}$ was extracted from the IMTEK collection Optimal Cooling of Steel Profiles¹. We compared the EBA-BDF(1) method to the EBA-exp and EBA-ROS(2) methods for the problem size $n = 1357$ on the time interval $[0, 2]$. The initial value X_0 was chosen as $X_0 = 0$ and the timestep was set to $h = 0.001$. The tolerance for EBA stop test was set to 10^{-7} for all methods and the projected low dimensional Sylvester equations were numerically solved by the solver (`lyap` from Matlab at each iteration of the extended block Arnoldi algorithm for the EBA-BDF(1), EBA-BDF(2) and EBA-ROS(2) methods. As the size of the coefficient matrices allowed it, we also computed an approximate solution of (4.1) applying a quadrature method to the integral form of the exact solution given by Formula(1.4) and took it as a reference solution. In Table 4.2, we reported the runtimes, in seconds, the number m of Arnoldi iterations and the Frobenius norm $\|\mathcal{E}(T_f)_m\|_F$ of the error at final time. As can be seen from the reported results in Table 4.2, the EBA-exp method clearly outperforms all the other listed options.

In Figure (4.2), we plotted the Frobenius residual norm $\|R_m(t_f)\|_F$ at final time T_f in function of the number m of Arnoldi iterations for the EBA-exp method.

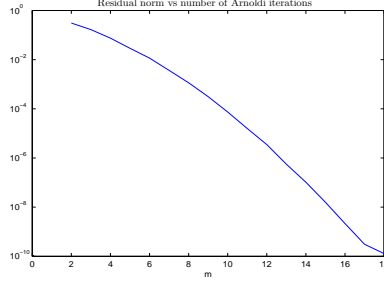
5. Appendix A. Here we recall the extended block Arnoldi (EBA) and block Arnoldi (BA) algorithms, when applied to the pair (A, E) . EBA is described in Algorithm 3 as follows

¹<https://portal.uni-freiburg.de/imteksimulation/downloads/benchmark>

	EBA-exp	EBA-BDF(1)	EBA-BDF(2)	EBA-ROS(2)
Runtime (s)	48.4 s ($m = 18$)	471.9 s ($m = 18$)	1549.2s ($m = 23$)	1827s ($m = 21$)
$\ \mathcal{E}_m(T_f)\ _F$	1.28×10^{-10}	5×10^{-5}	1.48×10^{-4}	4.9×10^{-5} .

TABLE 4.2

Optimal Cooling of Steel Profiles: runtimes, number of Arnoldi iterations and error norms

FIG. 4.2. Residual norm vs number m of Arnoldi iterations

Algorithm 3 The extended block Arnoldi algorithm (EBA)

- Inputs: A an $n \times n$ matrix, E an $n \times s$ matrix and m an integer.
 - Compute the QR decomposition of $[E, A^{-1}E]$, i.e., $[E, A^{-1}E] = V_1 \Lambda$;
Set $\mathcal{V}_0 = []$;
 - For $j = 1, \dots, m$
 - Set $V_j^{(1)}$: first s columns of V_j and $V_j^{(2)}$: second s columns of V_j
 - $\mathcal{V}_j = [\mathcal{V}_{j-1}, V_j]$; $\hat{V}_{j+1} = [AV_j^{(1)}, A^{-1}V_j^{(2)}]$.
 - Orthogonalize \hat{V}_{j+1} w.r.t \mathcal{V}_j to get V_{j+1} , i.e.,
For $i = 1, 2, \dots, j$
 $H_{i,j} = V_i^T \hat{V}_{j+1}$;
 $\hat{V}_{j+1} = \hat{V}_{j+1} - V_i H_{i,j}$;
Endfor i
 - Compute the QR decomposition of \hat{V}_{j+1} , i.e., $\hat{V}_{j+1} = V_{j+1} H_{j+1,j}$.
 - Endfor j .
-

The block Arnoldi algorithm is summarized in Algorithm 4 as follows

Since the above algorithms implicitly involve a Gram-Schmidt process, the obtained blocks $\mathcal{V}_m = [V_1, V_2, \dots, V_m]$ ($V_i \in \mathbb{R}^{n \times d}$), where $d = s$ for the block Arnoldi and $d = 2s$ for the extended block Arnoldi, have their columns mutually orthogonal provided none of the upper triangular matrices $H_{j+1,j}$ are rank deficient. Hence, after m steps, Algorithm 3 and Algorithm 4 build orthonormal bases \mathcal{V}_m of the Krylov subspaces $\mathcal{K}_m(A, E) = \text{Range}(E, AE, \dots, A^{m-1}E, A^{-1}E, \dots, (A^{-1})^m E)$ or $\mathbb{K}_m(A, E) = \text{Range}(E, AE, \dots, A^{m-1}E)$, respectively and a block upper Hessenberg matrix \mathcal{H}_m whose nonzero sub-blocks are the $H_{i,j}$. Note that each submatrix $H_{i,j}$ ($1 \leq i \leq j \leq m$) is of order d .

Let $\mathcal{T}_m \in \mathbb{R}^{d \times d}$ be the restriction of the matrix A to the extended Krylov subspace $\mathcal{K}_m(A, E)$ (or to the block Krylov subspace $\mathbb{K}_m(A, E)$), i.e., $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$. Then it can be shown that matrix \mathcal{T}_m is also block upper Hessenberg with $d \times d$ blocks, see [10, 17]. For the block Arnoldi algorithm, $\mathcal{T}_m = \mathcal{H}_m$ while for the extended block Arnoldi algorithm, a recursion can be derived to compute \mathcal{T}_m from \mathcal{H}_m without requiring matrix-vector products with A , see [17]. We notice that for large

Algorithm 4 The block Arnoldi algorithm (BA)

- Inputs: A an $n \times n$ matrix, E an $n \times s$ matrix and m an integer.
 - Compute the QR decomposition of E , i.e., $E = V_1 R_1$.
 - For $j = 1, \dots, m$
 1. $W_j = AV_j$,
 2. for $i = 1, 2, \dots, j$
 - $H_{i,j} = V_i^T W_j$,
 - $W_j = W_j - V_j H_{i,j}$,
 3. endfor
 4. $Q_j R_j = W_j$ (QR decomposition)
 5. $V_{j+1} = Q_j$, and $H_{j+1,j} = R_j$.
 - EndFor j
-

and non structured problems, the inverse of the matrix A is not computed explicitly and in this case we can use iterative solvers with preconditioners to solve linear systems with A .

6. Conclusion. We presented in the present paper two new approaches for computing approximate solutions to large scale differential Sylvester matrix equations. The first one comes naturally from the exponential expression of the exact solution and the use of approximation techniques of the exponential of a matrix times a block of vectors. The second approach is obtained by first projecting the initial problem onto a block Krylov (or extended Krylov) subspace, obtain a low dimensional differential Sylvester equation which is solved by using the well known BDF or Rosenbrock integration method. We gave some theoretical results such as the exact expression of the residual norm and also upper bounds for the norm of the error. Numerical experiments show that both approaches are promising for large-scale problems, with a clear advantage for the EBA-exp method in terms of computation time although the EBA-BDF(1) method shows to offer a good balance between the execution time and the accuracy in some cases.

REFERENCES

- [1] H. ABOU-KANDIL, G. FREILING, V. IONESCU, G. JANK, Matrix Riccati Equations in Control and Systems Theory, in Systems & Control Foundations & Applications, Birkhauser, (2003).
- [2] R.H. BARTELS, G.W. STEWART, Algorithm 432: Solution of the matrix equation $AX+XB=C$, Circ. Syst. Signal Proc., 13 (1972), 820–826.
- [3] J.C. BUTCHER, Numerical Methods for Ordinary Differential Equations, John Wiley & Sons, 2008
- [4] M. J. CORLESS AND A. E. FRAZHO, Linear systems and control - An operator perspective, Pure and Applied Mathematics. Marcel Dekker, New York-Basel, 2003.
- [5] V. DRUSKIN, L. KNIZHNERMAN, Extended Krylov subspaces: approximation of the matrix square root and related functions, SIAM J. Matrix Anal. Appl., 19(3)(1998), 755–771.
- [6] A. EL GUENNOUNI, K. JBILOU AND A.J. RIQUET, Block Krylov subspace methods for solving large Sylvester equations, Numer. Alg., 29 (2002), 75–96.
- [7] E. GALLOPOULOS AND Y. SAAD, Efficient solution of parabolic equations by Krylov approximation methods, SIAM J. Sci. Statist. Comput., 13 (1992), 1236–1264.
- [8] K. GLOVER, All optimal Hankel-norm approximations of linear multivariable systems and their L-infinity error bounds. International Journal of Control, 39(1984) 1115–1193.
- [9] G.H. GOLUB, S. NASH AND C. VAN LOAN, A Hessenberg Schur method for the problem $AX + XB = C$, IEEE Trans. Automat. Contr., 24 (1979), 909–913.
- [10] M. HEYOUNI, K. JBILOU, An Extended Block Arnoldi algorithm for Large-Scale Solutions of the Continuous-Time Algebraic Riccati Equation, Electronic Transactions on Numerical Analysis, 33, 53–62, (2009).
- [11] N. J. HIGHAM, The scaling and squaring method for the matrix exponential revisited. SIAM J. Matrix Anal. Appl., 26(4) (2005), 1179–1193.
- [12] M. Hochbruck and C. Lubich, On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal., 34 (1997), 1911–1925.
- [13] K. JBILOU, Low-rank approximate solution to large Sylvester matrix equations, App. Math. Comput., 177 (2006), 365–376.

- [14] C.B. MOLER, C.F. VAN LOAN, Nineteen Dubious Ways to Compute the Exponential of a Matrix, SIAM Review 20, 1978, pp. 801–836. Reprinted and updated as "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later," SIAM Review 45(2003), 3–49.
- [15] Y. SAAD, Numerical solution of large Lyapunov equations, in Signal Processing, Scattering, Operator Theory and Numerical Methods. Proceedings of the international symposium MTNS-89, vol. 3, M.A. Kaashoek, J.H. van Schuppen and A.C. Ran, eds., Boston, 1990, Birkhauser, pp. 503–511.
- [16] Y. SAAD, Analysis of some Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal., 29 (1992), 209–228.
- [17] V. SIMONCINI, A new iterative method for solving large-scale Lyapunov matrix equations, SIAM J. Sci. Comp., 29(3) (2007), 1268–1288.
- [18] ROSENBROCK H. H. Some general implicit processes for the numerical solution of differential equations, J. Comput. , 5(1963), 329?330.
- [19] H. VAN DER VORST, Iterative Krylov Methods for Large Linear Systems, Cambridge University Press, Cambridge, 2003.