



HAL
open science

Numerical solutions to large-scale differential Lyapunov matrix equations

Mustapha Hached, Khalide Jbilou

► **To cite this version:**

Mustapha Hached, Khalide Jbilou. Numerical solutions to large-scale differential Lyapunov matrix equations. Numerical Algorithms, 2018, 79 (3), pp.741-757. 10.1007/s11075-017-0458-y. hal-04391808

HAL Id: hal-04391808

<https://hal.science/hal-04391808>

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

Numerical solutions to large-scale differential Lyapunov matrix equations

M. Hached · K. Jbilou

the date of receipt and acceptance should be inserted later

Abstract In the present paper, we consider large-scale differential Lyapunov matrix equations having a low rank constant term. We present two new approaches for the numerical resolution of such differential matrix equations. The first approach is based on the integral expression of the exact solution and an approximation 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 Lyapunov matrix equation. The latter differential matrix problem is then solved by the Backward Differentiation Formula method (BDF) and the obtained solution is used to build a low rank approximate solution of the original problem. The process is being repeated, increasing the dimension of the projection space until some prescribed accuracy is achieved. We give some new theoretical results and present numerical experiments.

Keywords Extended block Krylov · Low rank · Differential Lyapunov equations

Mathematics Subject Classification (2010) MSC 65F · MSC 15A

1 Introduction

In the present paper, we consider the differential Lyapunov matrix equation (DLE in short) of the form

$$\begin{cases} \dot{X}(t) = A(t)X(t) + X(t)A^T(t) + B(t)B(t)^T; & (DLE) \\ X(t_0) = X_0, \quad t \in [t_0, T_f], \end{cases} \quad (1)$$

where the matrix $A(t) \in \mathbb{R}^{n \times n}$ is assumed to be nonsingular and $B(t) \in \mathbb{R}^{n \times s}$ is a full rank matrix, with $s \ll n$. The initial condition $X_0 = \tilde{Z}_0 \tilde{Z}_0^T$ is assumed to be a symmetric and positive low-rank given matrix. Differential Lyapunov equations play a fundamental role in many areas such as control, filter design theory, model reduction problems, differential equations and robust control problems [1,5]. For those applications, the matrix A is generally sparse and very large. For such problems, only a few attempts have been made to solve (1).

Let us first recall the following theoretical result which gives an expression of the exact solution of (1).

M. Hached
Laboratoire Painlevé UMR 8524 (ANO-EDP), UFR Mathématiques, Université des Sciences et Technologies de Lille, IUT A
Département Chimie, Rue de la Recherche (lieu-dit "Le Recueil"), BP 179 - 59653 Villeneuve d'Ascq Cedex, France

K. Jbilou (corresponding author)
Université du Littoral, Côte d'Opale, bâtiment H. Poincaré,
50 rue F. Buisson, F-62280 Calais Cedex, France. E-mail: jbilou@lmpa.univ-littoral.fr

Theorem 1 [1] *The unique solution of the general Lyapunov differential equation*

$$\dot{X}(t) = A(t)X + XA(t)^T + M(t); \quad X(t_0) = X_0 \quad (2)$$

is defined by

$$X(t) = \Phi_A(t, t_0)X_0\Phi_A^T(t, t_0) + \int_{t_0}^t \Phi_A(t, \tau)M(\tau)\Phi_A^T(t, \tau)d\tau. \quad (3)$$

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

$$X(t) = e^{(t-t_0)A}X_0e^{(t-t_0)A^T} + \int_{t_0}^t e^{(t-\tau)A}M(\tau)e^{(t-\tau)A^T}d\tau. \quad (4)$$

We notice that the problem (1) is equivalent to the linear ordinary differential equation

$$\begin{cases} \dot{x}(t) = \mathcal{A}(t)x(t) + b(t) \\ x_0 = \text{vec}(X_0) \end{cases} \quad (5)$$

where $\mathcal{A}(t) = I \otimes A(t) + A(t) \otimes I$, $x(t) = \text{vec}(X(t))$ and $b(t) = \text{vec}(B(t)B(t)^T)$, where $\text{vec}(Z)$ is the long vector obtained by stacking the columns of the matrix Z . For moderate size problems, it is then possible to use an integration method to solve (5). However, this approach is not adapted to large problems.

In the present paper, we will consider the case for which the coefficient matrices A and B are time-independent. The general idea of the approaches that will be developed in this work is to project the original problem (1) onto extended block Krylov (or block Krylov if A is not invertible) subspaces associated to the pair (A, B) . These subspaces are defined as follows

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

for block Krylov subspaces, or

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

for extended block Krylov subspaces. Notice that the extended Krylov subspace $\mathcal{K}_m(A, B)$ is a sum of two block Krylov subspaces

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

To compute an orthonormal basis $\{V_1, \dots, V_m\}$, where V_i is of dimension $n \times s$ for the block Krylov and $n \times 2s$ 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 [7, 25]. These algorithms also generate block Hessenberg matrices $\tilde{\mathcal{T}}_m = \mathcal{V}_{m+1}^T A \mathcal{V}_m$ satisfying the following algebraic relations

$$A \mathcal{V}_m = \mathcal{V}_{m+1} \tilde{\mathcal{T}}_m, \quad (6)$$

$$= \mathcal{V}_m \mathcal{T}_m + V_{m+1} T_{m+1,m} E_m^T, \quad (7)$$

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

When the matrix A is nonsingular and when the computation of $W = A^{-1}V$ 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

an upper bound for the norm of the error and an expression of the exact residual. A second approach, presented in Section 3, for which the initial differential Lyapunov matrix equation is projected onto a block (or extended block) Krylov subspace. Then, the obtained low dimensional differential Lyapunov equation is solved by using the well known Backward Differentiation Formula (BDF). In Section 4, an application to balanced truncation method for large scale linear-time varying dynamical systems is presented. The last section is devoted to some numerical experiments.

2 The first approach: using an approximation of the matrix exponential

In this section, we give a new approach for computing approximate solutions to large differential equations (1). The expression of the exact solution as

$$X(t) = e^{(t-t_0)A} X_0 e^{(t-t_0)A^T} + \int_{t_0}^t e^{(t-\tau)A} B B^T e^{(t-\tau)A^T} d\tau, \quad (8)$$

suggests the idea of computing $X(t)$ by approximating the factor $e^{(t-\tau)A} B$ and then using a quadrature method to compute the desired approximate solution.

As computing the exponential of a small matrix is straightforward, this is not the case for large scale problems, as $e^{(t-\tau)A}$ could be dense even though A is sparse. However, in our problem, the computation of $e^{(t-\tau)A}$ is not needed as we will rather consider the product $e^{(t-\tau)A} B$, for which approximations via projection methods onto block or extended block Krylov subspaces are well suited.

Krylov subspace projection methods generate a sequence of nested subspaces (Krylov or extended Krylov subspaces). Let $\mathcal{V}_m = [V_1, \dots, V_m]$ be the orthogonal matrix whose columns form an orthonormal basis of the subspace K_m . Following [21, 22, 27], an approximation to $Z = e^{(t-\tau)A} B$ can be obtained as

$$Z_m(t) = \mathcal{V}_m e^{(t-\tau)\mathcal{T}_m} \mathcal{V}_m^T B \quad (9)$$

where $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$. Therefore, the term appearing in the integral expression (8) can be approximated as

$$e^{(t-\tau)A} B B^T e^{(t-\tau)A^T} \approx Z_m(t) Z_m(t)^T. \quad (10)$$

For simplicity, we assume that $X_0 = 0$, an approximation to the solution of the differential Lyapunov equation (8) can be expressed as

$$X_m(t) = \mathcal{V}_m G_m(t) \mathcal{V}_m^T, \quad (11)$$

where

$$G_m(t) = \int_{t_0}^t \tilde{G}_m(\tau) \tilde{G}_m^T(\tau) d\tau, \quad (12)$$

and $\tilde{G}_m(\tau) = e^{(t-\tau)\mathcal{T}_m} B_m$, with $B_m = \mathcal{V}_m^T B$.

We notice that if $X_0 = \tilde{Z}_0 \tilde{Z}_0^T \neq 0$, then we can also approximate the first term $e^{(t-t_0)A} \tilde{Z}_0$ of (8) by Z_0 , using a Krylov-based or another method and then the approximate solution to (8) is given as $X_m(t) = Z_0 \tilde{Z}_0^T + \mathcal{V}_m G_m(t) \mathcal{V}_m^T$.

The next result shows that the matrix function G_m is the solution of a low-order differential Lyapunov matrix equation.

Theorem 2 *Let $G_m(t)$ be the matrix function defined by (12), then it satisfies the following low-order differential Lyapunov matrix equation*

$$\dot{G}_m(t) = \mathcal{T}_m G_m(t) + G_m(t) \mathcal{T}_m^T + B_m B_m^T, \quad t \in [t_0, T_f] \quad (13)$$

Proof The proof can be easily derived from the expression (12) and the result of Theorem 1.

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

$$\begin{aligned} \mathcal{V}_m^T R_m(t) \mathcal{V}_m &= \mathcal{V}_m^T (\dot{X}_m(t) - AX_m(t) - X_m(t)A^T - BB^T) \mathcal{V}_m \\ &= \dot{G}_m(t) - \mathcal{T}_m G_m(t) - G_m(t) \mathcal{T}_m^T - B_m B_m^T \\ &= 0, \end{aligned}$$

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

As mentioned earlier, once $\tilde{G}_m(\tau)$ is computed, we use a quadrature method to approximate the integral (12) in order to approximate $G_m(t)$.

We now briefly discuss some practical aspects of the computation of $e^{(t-\tau)\mathcal{T}_m} B_m$ where $B_m = \mathcal{V}_m^T B$, when m is small and \mathcal{T}_m is an upper block Hessenberg matrix.

In the last decade, many approximation techniques such as the use of partial fraction expansions or Padé approximation have been proposed, see for example [9,22]. However, it was remarked that a good way for evaluating the exponential of matrix times by a vector by using rational approximation to the exponential function. One of the main advantages of rational approximations as compared to polynomial approximations is the better stability of their integration schemes. Let us consider the rational function

$$F(z) = a_0 + \sum_{i=1}^p \frac{a_i}{z - \theta_i},$$

where the θ_i 's are the poles of the rational function F . Then, the approximation to $\tilde{G}_m(\tau) = e^{(t-\tau)\mathcal{T}_m} B_m$ is given by

$$\tilde{G}_m(\tau) \approx a_0 B_m + \sum_{i=1}^p a_i [(t-\tau)\mathcal{T}_m - \theta_i I]^{-1} B_m. \quad (14)$$

One of the possible choices for the rational function F is based on Chebychev approximation of the function e^x on $[0, \infty]$, see [22]. We notice that for small values of m , one can also directly compute the matrix exponential $e^{(t-\tau)\mathcal{T}_m}$ by using the well-known 'scaling and squaring method for the matrix exponential' method, [13]. This method was associated to a Padé approximation and is implemented in the `expm` Matlab routine.

From now on, we assume that the basis formed by the orthonormal columns of \mathcal{V}_m is obtained by applying the block Arnoldi or the extended block Arnoldi algorithm to the pair (A, B) .

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 $\|R_m\| < \varepsilon$ without having to compute extra products involving the matrix A . The next result shows how to compute the residual norm of $R_m(t)$ without forming the approximation $X_m(t)$ which is computed in a factored form only when convergence is achieved.

Theorem 3 *Let $X_m(t) = \mathcal{V}_m G_m(t) \mathcal{V}_m^T$ be the approximation obtained at step m by the block (or extended block) Arnoldi method. Then the residual $R_m(t)$ satisfies*

$$\|R_m(t)\| = \|T_{m+1,m} \bar{G}_m(t)\|, \quad (15)$$

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

Proof From the expression (11) of the approximation $X_m(t)$, we have

$$R_m(t) = \mathcal{V}_m \dot{G}_m(t) \mathcal{V}_m^T - A \mathcal{V}_m G_m(t) \mathcal{V}_m^T - \mathcal{V}_m G_m(t) \mathcal{V}_m^T A^T - BB^T.$$

Then using the relations (6) and the fact that $G_m(t)$ is solution of the reduced DLE (13), we get

$$R_m(t) = \mathcal{V}_{m+1} \begin{bmatrix} 0 & \bar{G}_m(t)^T T_{m+1,m}^T \\ T_{m+1,m} \bar{G}_m(t) & 0 \end{bmatrix} \mathcal{V}_{m+1}^T$$

and hence

$$\|R_m\| = \|T_{m+1,m} \bar{G}_m(t)\|.$$

The result of Theorem 3 is very important in practice, as it allows us to stop the iterations when convergence is achieved without computing the approximate solution $X_m(t)$.

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

Theorem 4 *Let $X_m(t)$ be the approximate solution given by (11). Then we have*

$$\dot{X}_m(t) = (A - F_m) X_m + X_m (A - F_m)^T + BB^T. \quad (16)$$

where $F_m = V_{m+1} T_{m+1,m} V_m^T$.

Proof Multilying the equation (13) from the left by \mathcal{V}_m and from the right by \mathcal{V}_m^T , and using the relations (6), we obtain

$$\dot{X}_m(t) = (A \mathcal{V}_m - V_{m+1} T_{m+1,m} E_m^T) G_m(t) \mathcal{V}_m^T + \mathcal{V}_m G_m(t) (\mathcal{V}_m^T A^T - E_m T_{m+1,m}^T V_{m+1}^T) + BB^T. \quad (17)$$

On the other hand, since $\mathcal{V}_m^T \mathcal{V}_m = I$, we have $E_m^T G_m(t) = E_m^T \mathcal{V}_m^T \mathcal{V}_m G_m(t) = V_{m+1}^T \mathcal{V}_m G_m(t)$. Therefore, using this last expression in (17), the result follows.

Remark 1 *The solution $X_m(t)$ can be given as a product of two low rank matrices. Consider the eigen-decomposition of the symmetric and positive matrix $md \times md$ $G_m(t) = U D U^T$ where D is the diagonal matrix of the eigenvalues of $G_m(t)$ sorted in decreasing order and $d = s$ for the block Arnoldi or $d = 2s$ for the extended block Arnoldi. Let U_l be the $md \times l$ matrix of the first l columns of U corresponding to the l eigenvalues of magnitude greater than some tolerance $dtol$. We obtain the truncated eigen-decomposition $G_m(t) \approx U_l D_l U_l^T$ where $D_l = \text{diag}[\lambda_1, \dots, \lambda_l]$. Setting $\tilde{Z}_m(t) = \mathcal{V}_m U_l D_l^{1/2}$, it follows that*

$$X_m(t) \approx \tilde{Z}_m(t) \tilde{Z}_m(t)^T. \quad (18)$$

Therefore, one has to compute and to store only the matrix $\tilde{Z}_m(t)$ which is usually the required factor in some control problems such as in the balanced truncation method for model reduction in large scale dynamical systems. This possibility is very important for storage limitations in the large scale problems.

The next result states that the error matrix $X(t) - X_m(t)$ satisfies a differential Lyapunov matrix equation.

Theorem 5 Let $X(t)$ be the exact solution of (1) and let $X_m(t)$ be the approximate solution obtained at step m . The error $E_m(t) = X(t) - X_m(t)$ satisfies the following equation

$$\dot{E}_m(t) = AE_m(t) + E_m(t)A^T - R_m(t), \quad (19)$$

and

$$E_m(t) = e^{(t-t_0)A} E_{m,0} e^{(t-t_0)A^T} + \int_{t_0}^t e^{(t-\tau)A} R_m(\tau) e^{(t-\tau)A^T} d\tau, \quad t \in [t_0, T_f]. \quad (20)$$

where $E_{m,0} = E_m(t_0)$.

Proof The result is easily obtained by subtracting the residual equation from the initial differential Lyapunov equation (1).

Next, we give an upper bound for the norm of the error.

Theorem 6 Assume that $X(t_0) = X_m(t_0)$, then we have the following upper bound

$$\|E_m(t)\| \leq \|T_{m+1,m}\| \|\bar{G}_m\|_\infty \frac{e^{2(t-t_0)\mu_2(A)} - 1}{2\mu_2(A)}, \quad (21)$$

where $\mu_2(A) = \frac{1}{2} \lambda_{\max}(A + A^T)$ is the 2-logarithmic norm and $\|\bar{G}_m\|_\infty = \max_{\tau \in [t_0, t]} \|\bar{G}_m(\tau)\|$. The matrix \bar{G}_m is the $d \times md$ matrix corresponding to the last d rows of G_m where $d = s$ when using the block Arnoldi and $d = 2s$ for the extended block Arnoldi.

Proof We first remind that the logarithmic norm provides the following bound $\|e^{tA}\| \leq e^{\mu_2(A)t}$. Therefore, using the expression (20), we obtain the following relation

$$\|E_m(t)\| \leq \int_{t_0}^t \|e^{(t-\tau)A}\|^2 \|R_m(\tau)\| d\tau.$$

Therefore, using (15) and the fact that $\|e^{(t-\tau)A}\| \leq e^{(t-\tau)\mu_2(A)}$, we get

$$\begin{aligned} \|E_m(t)\| &\leq \|T_{m+1,m}\| \|\bar{G}_m\|_\infty \int_{t_0}^t e^{2(t-\tau)\mu_2(A)} d\tau \\ &\leq \|T_{m+1,m}\| \|\bar{G}_m\|_\infty e^{2t\mu_2(A)} \int_{t_0}^t e^{-2\tau\mu_2(A)} d\tau \\ &\leq \|T_{m+1,m}\| \|\bar{G}_m\|_\infty e^{2t\mu_2(A)} \times \frac{e^{-2\mu_2(A)t} - e^{-2\mu_2(A)t_0}}{-2\mu_2(A)} \\ &= \|T_{m+1,m}\| \|\bar{G}_m\|_\infty \frac{e^{2(t-t_0)\mu_2(A)} - 1}{2\mu_2(A)}, \end{aligned}$$

which gives the desired result.

Notice that if $\|T_{m+1,m}\|$ is close to zero, which is the case when m is close to the degree of the minimal polynomial of A for B , then Theorem 6 shows that the error $E_m(t)$ tends to zero.

Next, we give another error bound for the norm of the error.

Theorem 7 Let $X(t)$ be the exact solution to (1) and let $X_m(t)$ be the approximate solution obtained at step m . Then we have

$$\|X(t) - X_m(t)\| \leq e^{t\mu_2(A)} (\|B\| + \|B_m\|) \int_{t_0}^t e^{-\tau\mu_2(A)} \|e^{(t-\tau)A} B - \mathcal{V}_m e^{(t-\tau)\mathcal{T}_m} B_m\| d\tau$$

where $\mu_2(A) = \lambda_{\max}((A + A^T)/2)$, $Z(\tau) = e^{(t-\tau)A} B$ and $Z_m(\tau) = \mathcal{V}_m e^{(t-\tau)\mathcal{T}_m} B_m$ with $B_m = \mathcal{V}_m^T B$.

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

$$\begin{aligned} \|X(t) - X_m(t)\| &= \left\| \int_{t_0}^t (Z(\tau)Z(\tau)^T - Z_m(\tau)Z_m(\tau)^T) d\tau \right\| \\ &= \left\| \int_{t_0}^t [Z(\tau)(Z(\tau) - Z_m(\tau))^T + (Z(\tau) - Z_m(\tau))Z_m(\tau)^T] d\tau \right\| \\ &\leq \int_{t_0}^t (\|Z(\tau)\| + \|Z_m(\tau)\|) \|Z(\tau) - Z_m(\tau)\| d\tau, \end{aligned}$$

Therefore, using the fact that $\mu_2(\mathcal{T}_m) = \lambda_{\max}((\mathcal{T}_m + \mathcal{T}_m^T)/2) \leq \lambda_{\max}((A + A^T)/2) = \mu_2(A)$, where $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$, it follows that

$$\begin{aligned} \|X(t) - X_m(t)\| &\leq e^{t\mu_2(A)} (\|B\| + \|B_m\|) \int_{t_0}^t e^{-\tau\mu_2(A)} \|Z(\tau) - Z_m(\tau)\| d\tau \\ &\leq e^{t\mu_2(A)} (\|B\| + \|B_m\|) \int_{t_0}^t e^{-\tau\mu_2(A)} \|e^{(t-\tau)A} B - \mathcal{V}_m e^{(t-\tau)\mathcal{T}_m} B_m\| d\tau, \end{aligned}$$

When using a block Krylov subspace method such as the block Arnoldi method, then one can generalize to the block case the results already stated in many papers; see [7, 9, 12, 22]. In particular, we can easily generalize the result given in [22] for the case $s = 1$ to the case $s > 1$. In this case, we have the following upper bound.

$$\|e^A B - \mathcal{V}_m e^{\mathcal{T}_m} B_m\| \leq 2 \|B\| \frac{\rho^m e^\rho}{m!}, \quad (22)$$

where $\rho = \|A\|$

The upper bound (22) could be used in Theorem 7 to obtain a new upper bound for the norm of the error. In that case, we obtain the following upper bound

$$\|X(t) - X_m(t)\| \leq 2 \|B\| \frac{\rho^m}{m!} e^{t(\mu_2(A)+\rho)} (\|B\| + \|B_m\|) \int_{t_0}^t e^{-\tau(\mu_2(A)+\rho)} (t-\tau)^m d\tau, \quad (23)$$

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 DLE'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 compute an orthonormal basis $\mathcal{V}_m = [V_1, \dots, V_m]$ of $\mathcal{K}_m(A, B) = \text{Range}[B, A^{-1}B, \dots, A^{-m}B, A^{m-1}B]$ and the upper block Hessenberg matrix \mathcal{T}_m .
 - Set $B_m = \mathcal{V}_m^T B$ and compute $\tilde{G}_m(\tau) = e^{(t-\tau)\mathcal{T}_m} B_m$ using the matlab function `expm`.
 - Use a quadrature method to compute the integral (12) and get an approximation of $G_m(t)$ for each $t \in [t_0, T_f]$.
 - If $\|R_m(t)\| = \|T_{m+1,m}\tilde{G}_m(t)\| < tol$ stop and compute the approximate solution $X_m(t)$ in the factored form given by the relation (18).
 - End
-

3 A second approach: Projecting and solving with BDF**3.1 Low-rank approximate solutions via BDF**

In this section, we show how to obtain low rank approximate solutions to the differential Lyapunov equation (1) by projecting directly the initial problem onto small block Krylov or extended block Krylov subspaces.

We first apply the block Arnoldi algorithm (or the extended block Arnoldi) to the pair (A, B) to get the matrices \mathcal{V}_m and $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$. Let $X_m(t)$ be the desired low rank approximate solution given as

$$X_m(t) = \mathcal{V}_m Y_m(t) \mathcal{V}_m^T, \quad (24)$$

satisfying the Petrov-Galerkin orthogonality condition

$$\mathcal{V}_m^T R_m(t) \mathcal{V}_m = 0, \quad t \in [t_0, T_f], \quad (25)$$

where $R_m(t)$ is the residual $R_m(t) = \dot{X}_m(t) - A X_m(t) - X_m(t) A^T - B B^T$. Then, from (24) and (25), we obtain the low dimensional differential Lyapunov equation

$$\dot{Y}_m(t) - \mathcal{T}_m Y_m(t) - Y_m(t) \mathcal{T}_m^T - B_m B_m^T = 0, \quad (26)$$

with $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$ and $B_m = \mathcal{V}_m^T B$. The obtained low dimensional differential Lyapunov equation (26) is the same as the one given by (13). For this second approach, we have to solve the latter low dimensional differential Lyapunov equation by some integration method such as the well known Backward Differentiation Formula (BDF).

Notice that we can also compute the norm of the residual without computing the approximation $X_m(t)$ which is also given, when convergence is achieved, in a factored form as in (18). The norm of the residual is given as

$$\|R_m(t)\| = \|T_{m+1,m} \tilde{Y}_m(t)\|, \quad (27)$$

where \tilde{Y}_m is the $d \times md$ matrix corresponding to the last d rows of Y_m where $d = s$ when using the block Arnoldi and $d = 2s$ for the extended block Arnoldi.

3.2 BDF for solving the low order differential Lyapunov equation (26)

In this subsection, we will apply the Backward Differentiation Formula (BDF) method for solving, at each step m of the block (or extended block) Arnoldi process, the low dimensional differential Lyapunov matrix equation (26). We notice that BDF is especially used for the solution of stiff differential equations.

At each time t_k , let $Y_{m,k}$ be the approximation of $Y_m(t_k)$, where Y_m is a solution of (26). 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

$$Y_{m,k+1} = \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i} + h_k \beta \mathcal{F}(Y_{m,k+1}), \quad (28)$$

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 1 and $\mathcal{F}(X)$ is given by

$$\mathcal{F}(Y) = \mathcal{T}_m Y + Y \mathcal{T}_m^T + B_m B_m^T.$$

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

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

$$-Y_{m,k+1} + h_k \beta (\mathcal{T}_m Y_{m,k+1} + Y_{m,k+1} \mathcal{T}_m^T + B_m B_m^T) + \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i} = 0,$$

which can be written as the following Lyapunov matrix equation

$$\mathbb{T}_m Y_{m,k+1} + Y_{m,k+1} \mathbb{T}_m^T + \mathbb{B}_{m,k} \mathbb{B}_{m,k}^T = 0. \quad (29)$$

We assume that at each time t_k , the approximation $Y_{m,k}$ is factorized as a low rank product $Y_{m,k} \approx Z_{m,k} Z_{m,k}^T$, where $Z_{m,k} \in \mathbb{R}^{n \times m_k}$, with $m_k \ll n$. In that case, the coefficient matrices appearing in (29) are given by

$$\mathbb{T}_m = h_k \beta \mathcal{T}_m - \frac{1}{2} I \text{ and } \mathbb{B}_{m,k+1} = [\sqrt{h_k \beta} B, \sqrt{\alpha_0} Z_{m,k}, \dots, \sqrt{\alpha_{p-1}} Z_{m,k+1-p}].$$

The Lyapunov matrix equation (29) can be solved by applying direct methods based on Schur decomposition such as the Bartels-Stewart algorithm [3, 11]. We notice that for large problems, many Krylov subspace type methods have been proposed to solve (29); [8, 14–17, 25, 21].

Remark 2 *The main difference between Approach 1 and Approach 2 is the fact that in the first case, we compute an approximation of an integral using a quadrature formula, while in the second case, we have to solve a low dimensional differential Lyapunov equation using the BDF method. Mathematically, the two approaches are equivalent and they mainly differ in the way of numerically computing the low-order approximations: G_m in the first approach and Y_m in the second one.*

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

Algorithm 2 The extended block Arnoldi (EBA-BDF) method for DLE'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 compute an orthonormal basis $\mathcal{V}_m = [V_1, \dots, V_m]$ of $\mathcal{K}_m(A, B) = \text{Range}[B, A^{-1}B, \dots, A^{-m}B, A^{m-1}B]$ and the upper block Hessenberg matrix \mathcal{T}_m .
 - Set $B_m = \mathcal{V}_m^T B$ and use the BDF method to solve the low dimensional differential Lyapunov equation

$$\dot{Y}_m(t) - \mathcal{T}_m Y_m(t) - Y_m(t) \mathcal{T}_m^T - B_m B_m^T = 0, t \in [t_0, T_f]$$

- If $\|R_m(t)\| = \|T_{m+1,m} \tilde{Y}_m(t)\| < tol$ stop and compute the approximate solution $X_m(t)$ in the factored form given by the relation (18).
 - End
-

4 Numerical examples

In this section, we compare the two 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) applying a quadrature method to compute the projected exponential form solution (12). We used a scaling and squaring strategy, implemented in the MATLAB `expm` function; see [12, 18] for more details. The second method (Algorithm 2) is based on the BDF integration method applied to the projected differential Lyapunov equation (26). The basis of the projection subspaces were generated by the extended block Arnoldi algorithm for both methods. All the experiments were performed on a laptop with an Intel Core i7 processor and 16GB of RAM. The algorithms were coded in Matlab R2014b.

Example 1. The matrix A was obtained from the 5-point discretization of the operators

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

on the unit square $[0, 1] \times [0, 1]$ with homogeneous Dirichlet boundary conditions. The number of inner grid points in each direction is $n_0 =$ and the dimension of the matrix A was $n = n_0^2$. Here we set $f_1(x, y) = 10xy$, $f_2(x, y) = e^{x^2y}$, $f_3(x, y) = 100y$, $f_4(x, y) = x^2y$, $g_1(x, y) = 20y$ and $g_2(x, y) = xy$. The time interval considered was $[0, 2]$ and the initial condition $X_0 = X(0)$ was chosen as the low rank product $X_0 = Z_0 Z_0^T$, where $Z_0 = 0_{n \times 2}$. For both 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 method, we used a 2-step BDF scheme with a constant timestep h . The entries of the matrix B were random values uniformly distributed on the interval $[0, 1]$ and the number of the columns in B was $s = 2$.

In order to check if our approaches produce reliable results, we began comparing our results to the one given by Matlab's `ode23s` solver which is designed for stiff differential equations. This was done by vectorizing our DLE, stacking the columns of X one on top of each other. This method, based on Rosenbrock integration scheme, 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 matrix A .

In Figure 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 $\text{size}(A) = 100 \times 100$ and a constant timestep $h = 10^{-3}$. We observe that all the considered methods give similar results in terms of accuracy. The relative error norms

$$\frac{\|X_{EBA-exp}(t_f) - X_{ode23s}(t_f)\|}{\|X_{ode23s}(t_f)\|} \quad \text{and} \quad \frac{\|X_{EBA-BDF(2)}(t_f) - X_{ode23s}(t_f)\|}{\|X_{ode23s}(t_f)\|} \quad \text{at final}$$

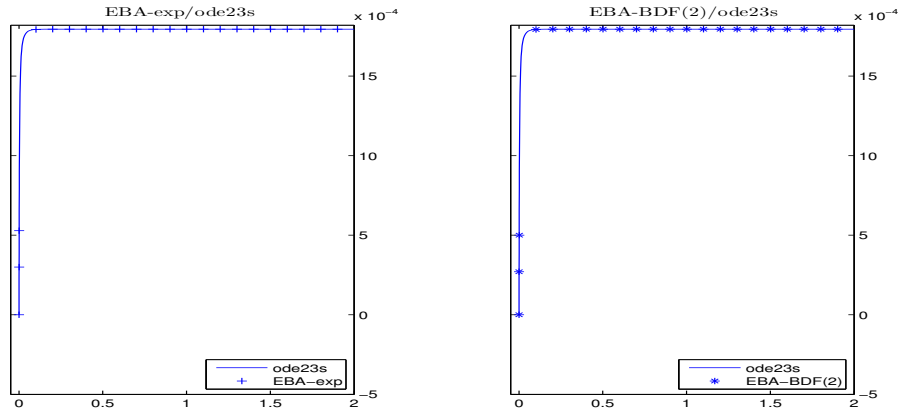


Fig. 1 Values of $X_{11}(t)$ for $t \in [0, 2]$

time $t_f = 2$ were equal to 1.8×10^{-10} and 9.1×10^{-11} respectively. The runtimes were respectively 0.59s, 5.1s for the EBA-exp and EBA-BDF(2) methods and 1001s for the ode23s routine.

In Table 2, we give the obtained runtimes in seconds, for the resolution of Equation (1) for $t \in [0, 2]$, with a timestep $h = 0.001$ and the Frobenius norm of the relative residual at the final time.

size(A)	EBA-exp	EBA-BDF(2)	Residual norm
2500×2500	2.7 s	25.2 s	$\mathcal{O}(10^{-9})$ ($m = 16$)
6400×6400	4.3 s	41.7 s	$\mathcal{O}(10^{-9})$ ($m = 19$)
10000×10000	4.9 s	43.3 s	$\mathcal{O}(10^{-9})$ ($m = 19$)
22500×22500	10.0 s	77.9 s	$\mathcal{O}(10^{-9})$ ($m = 24$)
160000×160000	68.7 s	283.4 s	$\mathcal{O}(10^{-8})$ ($m = 33$)

Table 2 runtimes and relative residual norms for EBA-exp and EBA+BDF(2)

The results in Table 2 illustrate that the EBA-exp method clearly outperforms the EBA-BDF(2) method in terms of computation time even though both methods are equally accurate. In Figure 2, we featured the norm of the residual at final time $t = 2$ for both EBA-exp and EBA-BDF(2) methods for $\text{size}(A) = 6400 \times 6400$ in function of the number m of extended Arnoldi iterations. We observe that the plots coincide for both methods.

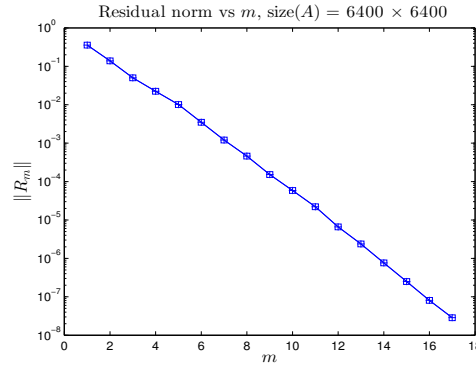


Fig. 2 Residual norms vs the number of extended Arnoldi iterations m

Example 2. This example comes from the autonomous linear-quadratic optimal control problem of one dimensional heat flow

$$\begin{aligned} \frac{\partial}{\partial t} x(t, \eta) &= \frac{\partial^2}{\partial \eta^2} x(t, \eta) + b(\eta)u(t) \\ x(t, 0) &= x(t, 1) = 0, t > 0 \\ x(0, \eta) &= x_0(\eta), \eta \in [0, 1] \\ y(x) &= \int_0^1 c(\eta)x(t, \eta)d\eta, x > 0. \end{aligned}$$

Using a standard finite element approach based on the first order B-splines, we obtain the following ordinary differential equation

$$M\dot{\mathbf{x}}(t) = K\mathbf{x}(t) + Fu(t) \quad (30)$$

$$y(t) = C\mathbf{x}(t), \quad (31)$$

where the matrices M and K are given by:

$$M = \frac{1}{6n} \begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{pmatrix}, \quad K = -\alpha n \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

Using the semi-implicit Euler method, we get the following continuous-time dynamical system

$$(M - \Delta t K)\dot{\mathbf{x}}(t) = M\mathbf{x}(t) + \Delta t Fu_k.$$

We set $A = (M - \Delta t K)^{-1}M$ and $B = \Delta t (M - \Delta t K)^{-1}F$. The entries of the $n \times s$ matrix F and the $s \times n$ matrix C were random values uniformly distributed on $[0, 1]$. In our experiments we used $\Delta t = 0.01$, $\alpha = 0.05$, $s = 2$ for different values of n .

In Table 3, we give the obtained runtimes in seconds, for the resolution of Equation (1) for $t \in [0, 2]$, with a timestep $h = 0.001$ and the Frobenius norm of the relative residual at the final time. The figures in Table 3 illustrate the gain of speed provided by the EBA-exp method. Again, both methods performed similarly in terms

size(A)	EBA-exp	EBA-BDF(2)	Residual norms
2500×2500	0.9 s	8.0 s	$\mathcal{O}(10^{-12})$ ($m = 12$)
6400×6400	4.9 s	14.4 s	$\mathcal{O}(10^{-12})$ ($m = 12$)
10000×10000	10.9 s	23.2 s	$\mathcal{O}(10^{-13})$ ($m = 12$)
20000×20000	62.0 s	112.4 s	$\mathcal{O}(10^{-12})$ ($m = 13$)

Table 3 runtimes and residual norms for EBA-exp and EBA-BDF(2)

of accuracy. In Figure 3, we considered the case $\text{size}(A) = 100 \times 100$ and plotted the upper bound of the error norms as stated in Formula (21) at the final time T_f against the computed norm of the errors, taking the solution given by the integral formula (8) as a reference, in function of the number m of Arnoldi iterations for the EBA-exp method.

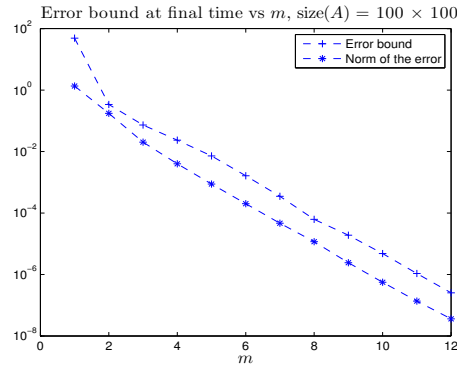


Fig. 3 Upper bounds of the error norms and computed error norms vs the number of iterations

Example 3 In this last example, we considered the well-known problem: Optimal Cooling of Steel Profiles. The matrices were extracted from the IMTEK collection¹. We compared the EBA-BDF(2) method to the EBA-exp method for all the available problem sizes $n = 1357, 5177, 20209, 79841$, on the time interval $[0, 1000]$. The initial value X_0 was chosen as $X_0 = 0$ and the timestep was set to $h = 1$. The tolerance for the extended-Arnoldi stop test was set to 10^{-7} for both methods and the projected low dimensional Lyapunov equations were numerically solved by the solver (`lyap` from Matlab) at each iteration of the extended block Arnoldi algorithm for the EBA-BDF(2) method.

size(A)	EBA-exp	EBA-BDF(2)	Residual norms
1357×1357	7.6 s	45.6 s	$\mathcal{O}(10^{-8})$ ($m = 27$)
5177×5177	20.4 s	106.7 s	$\mathcal{O}(10^{-7})$ ($m = 36$)
20209×20209	57.8 s	231 s	$\mathcal{O}(10^{-7})$ ($m = 47$)
79841×79841	448 s	722 s	$\mathcal{O}(10^{-7})$ ($m = 63$)

Table 4 Optimal Cooling of Steel Profiles: runtimes and residual norms for EBA-exp and EBA-BDF(2)

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

In Table 4, we listed the obtained runtimes which again showed the advantage of the EBA-exp method in terms of execution time and similar accuracy for both methods.

5 Conclusion

We presented in the present paper two new approaches for computing approximate solutions to large scale differential Lyapunov 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 by projection methods. The second approach is obtained by first projecting the initial problem onto a block Krylov (or extended Krylov) subspace, obtain a low dimensional differential Lyapunov equation which is solved by using the well known BDF 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 errors. Numerical experiments show that both methods are promising for large-scale problems, with a clear advantage for the EBA-exp method in terms of computation time.

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. B.D.O. ANDERSON, J.B. MOORE, *Linear Optimal Control*, Prentice-Hall, Englewood Cliffs, NJ, (1971).
3. R.H. BARTELS, G.W. STEWART, *Algorithm 432: Solution of the matrix equation $AX+XB=C$* , Circ. Syst. Signal Proc., 13 (1972), 820–826.
4. C. BREZINSKI, *Computational Aspects of Linear Control*, Kluwer, Dordrecht, 2002.
5. M. J. CORLESS AND A. E. FRAZHO, *Linear systems and control - An operator perspective*, Pure and Applied Mathematics. Marcel Dekker, New York-Basel, 2003.
6. B.N. DATTA, *Numerical Methods for Linear Control Systems Design and Analysis*, Elsevier Academic Press, (2003).
7. 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.
8. A. EL GUENNOUNI, K. JBILOU AND A.J. RIQUET, *Block Krylov subspace methods for solving large Sylvester equations*, Numer. Alg., 29 (2002), 75–96.
9. E. GALLOPOULOS AND Y. SAAD, *Efficient solution of parabolic equations by Krylov approximation methods*, SIAM J. Sci. Statist. Comput., 13 (1992), 1236–1264.
10. 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.
11. 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.
12. N. J. HIGHAM, *The scaling and squaring method for the matrix exponential revisited*. SIAM J. Matrix Anal. Appl., 26(4) (2005), 1179–1193.
13. N. J. HIGHAM AND A. H AL-MOHY, *A new scaling and squaring algorithm for the matrix exponential*. SIAM J. Matrix Anal. Appl., 31(3) (2009), 970–989.
14. D.Y. HU, L. REICHEL, *Krylov-subspace methods for the Sylvester equation*, Lin. Alg. Appl., 172 (1992), 283–313.
15. I.M. JAÏMOUKHA AND E.M. KASENALLY, *Krylov subspace methods for solving large Lyapunov equations*, SIAM J. Numer. Anal., 31 (1994), 227–251.
16. K. JBILOU, *Low-rank approximate solution to large Sylvester matrix equations*, App. Math. Comput., 177 (2006), 365–376.
17. K. JBILOU AND A. J. RIQUET, *Projection methods for large Lyapunov matrix equations*, Lin. Alg. Appl., 415 (2006), 344–358.
18. 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.
19. B.C. MOORE, *Principal component analysis in linear systems: controllability, observability and model reduction*, IEEE Trans. Automatic Contr., AC-26(1981), 17–32.
20. C. T. MULLIS AND R. A. ROBERTS, *Synthesis of minimum roundoff noise fixed point digital filters*, IEEE Trans. Acoust. Speech Signal Process, 24, 1976.
21. 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.

22. Y. SAAD, *Analysis of some Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 29 (1992), 209–228.
23. H. SANDBERG, *Linear Time-Varying Systems: Modeling and Reduction*, Ph.D. thesis, Department of Automatic Control, Lund Institute of Technology, Lund, 2002.
24. S. SHOKOOHI, L. SILVERMAN, AND P. VAN DOOREN, *Linear time-variable systems: balancing and model reduction*, IEEE Trans. Automat. Control, 28 (1983), 810–822.
25. V. SIMONCINI, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comp., 29(3) (2007), 1268–1288.
26. E. I. VERRIEST AND T. KAILATH, *On generalized balanced realizations*, IEEE Trans. Automat. Control, 28 (1983), 833–844.
27. H. VAN DER VORST, *Iterative Krylov Methods for Large Linear Systems*, Cambridge University Press, Cambridge, 2003.