Krylov Subspace Type Methods for Solving Projected Generalized Continuous-Time Lyapunov Equations

YUJIAN ZHOU, YIQIN LIN
Hunan University of Science and Engineering
Institute of Computational Mathematics
Yangzhitang Road 130, 425100 Yongzhou
China
yjzhou@yahoo.cn, yqlin@yahoo.cn

LIANG BAO
East China University of Science and Technology
Department of Mathematics
Meilong Road 130, 200237 Shanghai
China
nlbao@yahoo.cn

Abstract: In this paper we consider the numerical solution of projected generalized continuous-time Lyapunov equations with low-rank right-hand sides. The interest in this problem stems from stability analysis and control problems for descriptor systems including model reduction based on balanced truncation. Two projection methods are proposed for calculating low-rank approximate solutions. One is based on the usual Krylov subspace, while the other is based on the union of two different Krylov subspaces. The former is the Krylov subspace method and the latter is the extended Krylov subspace method. For these two methods, exact expressions for the norms of residuals are derived and results on finite termination are presented. Numerical experiments in this paper show the effectiveness of the proposed methods.

Key–Words: Projected generalized Lyapunov equation, Projection method, Krylov subspace, Alternating direction implicit method, Matrix pencil, C-stable

1 Introduction

In this paper we consider the projected generalized continuous-time Lyapunov equation of the form

$$\begin{cases} EXA^T + AXE^T + P_l BB^T P_l^T = 0, \\ X = P_r X P_r^T, \end{cases} \tag{1}$$

where $A, E \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times s}$, and $X \in \mathbb{R}^{n\times n}$ is the sought-after solution. Here, P_l and P_r are the spectral projectors onto the left and right deflating subspaces corresponding to the finite eigenvalues of the pencil $\lambda E - A$, respectively. It has been shown in [36] that if the pencil $\lambda E - A$ is c-stable, i.e., all its finite eigenvalues have negative real part, then the projected generalized continuous-time Lyapunov equation (1) has a unique, symmetric, and positive semidefinite solution.

We assume that the pencil $\lambda E - A$ is regular, i.e., $\det(\lambda E - A)$ is not identically zero. Under this assumption, the pencil $\lambda E - A$ has the Weierstrass canonical form [12]: there exist nonsingular $n \times n$ matrices W and T such that

$$E = W \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} T, \quad A = W \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} T, \tag{2}$$

where J and N are block diagonal matrices with each diagonal block being a Jordan block. The eigenvalues of J are the finite eigenvalues of the pencil $\lambda E - A$ and N corresponds to the eigenvalue at infinity. Using (2), P_l and P_r can be expressed as

$$P_l = W \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, P_r = T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T. \tag{3}$$

The projected generalized continuous-time Lyapunov equation (1) arises in stability analysis and control design problems for descriptor systems including the characterization of controllability and observability properties, balanced truncation model order reduction, determining the minimal and balanced realizations as well as computing H_2 and Hankel norms; see [1, 17, 24, 27, 37] and the references therein.

If E is nonsingular, then $P_l = P_r = I$. In this case, the projected equation (1) reduces to the generalized Lyapunov equation $EXA^T + AXE^T + BB^T = 0$. The generalized Lyapunov equation can be further reduced to the standard Lyapunov equation $\widetilde{A}X + X\widetilde{A}^T + \widetilde{B}\widetilde{B}^T = 0$, where $\widetilde{A} = E^{-1}A$ and $\widetilde{B} = E^{-1}B$.

A number of numerical solution methods have been proposed for the standard/generalized Lyapunov and Sylvester equations. Two classical direct methods are the Bartels-Stewart method [4, 13, 25, 34] and the Hammarling method [16, 21]. These methods need to compute the real Schur forms/generalized real Schur forms of the underlying matrices/matrix pencils by means of the QR/QZ algorithm [14] and require $\mathcal{O}(n^3)$ flops and $\mathcal{O}(n^2)$ memory. Besides direct methods, we mention, among several iterative methods, the Smith method [32], the alternating direction implicit iteration (ADI) method [6, 41, 23], the Smith(l) method [26], the low-rank Smith method [15, 26], the Cholesky factor-alternating direction implicit method [22], and the (generalized) matrix sign function method [5, 8, 9]. There are also several other approaches to solve large-scale Lyapunov and Sylvester equations using Krylov subspaces, see, for example, [2, 3, 18, 19, 20, 31]. The ADI methods and Krylov subspace based methods are well suited for large-scale Lyapunov and Sylvester equations with sparse coefficient matrices.

Several numerical methods have been proposed in the literature for solving the projected generalized Lyapunov equation (1). In [35], two direct methods, the generalized Bartels-Stewart method and the generalized Hammarling method, are proposed for the projected generalized Lyapunov equation of small or medium size. The generalized Hammarling method is designed to obtain the Cholesky factor of the solution. These two methods are based on the generalized real Schur form of the pencil $\lambda E - A$, and require also $\mathcal{O}(n^3)$ flops and $\mathcal{O}(n^2)$ memory.

Iterative methods to solve large sparse projected generalized Lyapunov equations have also been proposed. Stykel [40] extended the ADI method and the Smith method to the projected equation. Moreover, low-rank versions of these methods were also presented, which could be used to compute low-rank approximations to the solution. Another iterative method for the projected generalized Lyapunov equation is the modified generalized matrix sign function method [39]. Unlike the classical generalized matrix sign function method, the variant converges quadratically independent of the index of the underlying matrix pencil, see [39] for more details.

It should be mentioned that any numerical method for solving large sparse projected generalized Lyapunov equations crucially depends on the availability of expressions for the spectral projectors P_l and P_r. If such expressions are not available, the usual approach for computing these projectors via the generalized real Schur factorization of the pencil $\lambda E - A$, would be much too expensive for large-scale equations. In some applications such as computational fluid dynamics and constrained structural mechanics, some special block structure of the matrices E and A can be exploited to construct the spectral projectors P_l and P_r in explicit form, see, for example, [40].

The ADI method requires to select shift parameters. Once the shift parameters are given, the ADI method has a well-understood convergence theory. To obtain optimal shift parameters, we need to solve a rational min-max problem. This problem is only solved for standard Lyapunov equations with symmetric coefficient matrices. For the non-symmetric case, some heuristic shift selection procedures have been proposed to compute the suboptimal ADI shift parameters, see [26, 7]. However, these shift selection procedures do not work well for some applications. If some poor shift parameters are provided by the shift selection procedure, it can lead to very slow convergence in the ADI method.

In this paper we firstly reformulate the projected generalized Lyapunov equation (1) to the equivalent projected standard Lyapunov equation. Then we propose a Krylov subspace method for solving projected standard Lyapunov equations, which is the natural extension of the Krylov subspace method for standard Lyapunov equations. Since the coefficient matrix of the projected standard Lyapunov equation is singular, we can not generalize the iterative method proposed by Simoncini [31] directly to the projected standard Lyapunov equation. To overcome this disadvantage, we will use the $\{2\}$-inverse of the singular coefficient matrix to derive an extended Krylov subspace method, where the projection subspace is the union of two Krylov subspaces. One of these two Krylov subspaces is the usual Krylov subspace, while the other is based on the $\{2\}$-inverse of the coefficient matrix. The expressions of residuals and the results on finite termination for these two method are presented. Moreover, the performance of the newly proposed methods is compared to that of the generalized low-rank ADI method [40].

Throughout this paper, we adopt the following notations. We denote by I_s the $s \times s$ identity matrix. If the dimension of I_s is apparent from the context, we drop the index and simply use I. The zero vector or zero matrix is denoted by 0. The dimensions of these vectors and matrices, if not specified, are deduced by the context. The space of $m \times n$ real matrices is denoted by $\mathbb{R}^{m\times n}$. The Frobenius matrix norm is denoted by $\|\cdot\|_F$. The superscript " $\cdot^T$ " denotes the transpose of a vector or a matrix. The notation $\mathrm{span}\{v_1, v_2, \cdots, v_m\}$ denotes the space spanned by the sequence $v_1, v_2, \cdots, v_m$.

The remainder of the paper is organized as follows. In Section 2, we first briefly review the definition of the Krylov subspace and the Arnoldi process for establishing an orthonormal basis of this subspace. Then, we present a Krylov subspace method for solving the projected Lyapunov equation. In Section 3, we generalize the definition of the extended Krylov subspace to the singular matrix, and propose an extended Krylov subspace method for the projected Lyapunov equation. Section 4 is devoted to some numerical tests. Finally, conclusions are given in Section 5.

2 Krylov subspace method

In the remainder of this paper, for the sake of the simplicity of presentation, we assume $s = 1$, that is, B is a vector, although the whole discussion and algorithms could be stated for B being a matrix by using some extra notation and technical treatment.

2.1 Krylov subspace and the Arnoldi process

In this subsection we recall the definition of a Krylov subspace and the Arnoldi process for constructing an orthonormal basis of the Krylov subspace.

Let F be an $n \times n$ real matrix and v an $n-$dimensional real vector. The Krylov subspace $\mathcal{K}_m(F, v)$ is defined by

$$\mathcal{K}_m(F, v) = \text{span}\{v, Fv, F^2v, \cdots, F^{m-1}v\}.$$

The Arnoldi process [29] can be used to establish an orthonormal basis of the Krylov subspace $\mathcal{K}_m(F, v)$. *The Arnoldi process based on the modified Gram-Schmidt procedure is presented in the following algorithm:*

Algorithm 1 Arnoldi process

1. *Let* $v_1 = v/\|v\|$.
2. *For* $j = 1, 2, \cdots, m$
3. $\quad w = Fv_j$.
4. $\quad$ *For* $i = 1, 2, \cdots, j$
5. $\qquad h_{ij} = v_i^T w$.
6. $\qquad w = w - v_i h_{ij}$.
7. $\quad$ *End For*
8. $\quad h_{j+1,j} = \|w\|$.
9. $\quad v_{j+1} = w/h_{j+1,j}$.
10. *End For*

Algorithm 1 is known as an implementation of the Arnoldi process using the modified Gram-Schmidt orthogonalization [10] for generating an orthonormal basis of $\mathcal{K}_m(F, v)$. It is well known that in the presence of finite precision arithmetic, a loss of orthogonality can occur when the orthogonalization algorithm progresses, see [10, 14, 28]. A remedy is the so-called reorthogonalization, where the current vectors have to be orthogonalized against previously created vectors. One can choose between a selective reorthogonalization or a full reorthogonalization.

Define

$$V_m = [v_1, v_2, \cdots, v_m]$$

and the $(m + 1) \times m$ upper Hessenberg matrix

$$\widetilde{H}_m = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ & h_{32} & \ddots & \vdots \\ & & \ddots & h_{mm} \\ & & & h_{m+1,m} \end{bmatrix}.$$

The columns of the matrix V_m form an orthonormal basis of the Krylov subspace $\mathcal{K}_m(F, v)$. Moreover, from the Arnoldi algorithm we can deduce the following Arnoldi relations

$$\begin{aligned} FV_m &= V_{m+1}\widetilde{H}_m, \\ FV_m &= V_m H_m + h_{m+1,m} v_{m+1} e_m^T, \\ H_m &= V_m^T F V_m, \\ V_m^T V_m &= I_m, \end{aligned}$$

where H_m is the $m \times m$ matrix obtained from $\widetilde{H}_m$ by deleting the last row and e_m is the last column of the $m \times m$ identity matrix I_m.

2.2 Krylov subspace method

We always assume that the pencil $\lambda E - A$ is c-stable, i.e., all its finite eigenvalues have negative real parts. It follows that A is nonsingular. In this case, by making use of the expressions of A, P_l, and P_r in (2) and (3), we have $A^{-1}P_l = P_r A^{-1}$. Hence the projected generalized Lyapunov equation (1) is equivalent to the projected standard Lyapunov equation of the form

$$\begin{cases} (A^{-1}E)X + X(A^{-1}E)^T = -P_r A^{-1} B B^T A^{-T} P_r^T, \\ X = P_r X P_r^T. \end{cases} \tag{4}$$

We now apply the framework of a projection technique to derive a method for solving the projected equation (4). The projection subspace used here is

$$\begin{aligned} &\mathcal{K}_m(A^{-1}E, B_r) \\ &= \text{span}\{B_r, A^{-1}EB_r, \cdots, (A^{-1}E)^{m-1}B_r\}, \end{aligned}$$

where $B_r = P_r A^{-1} B$.

By applying the Arnoldi process to the Krylov subspace $\mathcal{K}_m(A^{-1}E, B_r)$, we generate the matrix V_m, whose columns form an orthonormal basis of $\mathcal{K}_m(A^{-1}E, B_r)$. With H_m defined by the Arnoldi process, we obtain the following Arnoldi relation

$$A^{-1}EV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T = V_{m+1}\widetilde{H}_m. \tag{5}$$

The approximate solution to X is constructed as

$$X_m = V_m Y_m V_m^T.$$

Let $\beta = \|B_r\|_F$. Since $B_r = \beta v_1 = \beta V_m e_1$ with e_1 being the first column of the identity I_m, the residual matrix R_m can be then expressed as

$$\begin{aligned} R_m &= A^{-1}EX_m + X_m(A^{-1}E)^T \\ &\quad + P_r A^{-1}BB^T A^{-T} P_r^T \\ &= A^{-1}EV_m Y_m V_m^T \\ &\quad + V_m Y_m V_m^T (A^{-1}E)^T \\ &\quad + \beta^2 V_m e_1 e_1^T V_m^T. \end{aligned} \tag{6}$$

According to the Galerkin condition, we want to find an approximate solution $X_m = V_m Y_m V_m^T$ such that

$$V_m^T R_m V_m = 0. \tag{7}$$

Since $V_m^T A^{-1} E V_m = H_m$, it follows from (7) and (6) that Y_m satisfies

$$H_m Y_m + Y_m H_m^T + \beta^2 e_1 e_1^T = 0. \tag{8}$$

The following theorem is one main result of this section.

Theorem 2 *Suppose that m steps of the Arnoldi process have been taken for $\mathcal{K}_m(A^{-1}E, B_r)$. Let $X_m = V_m Y_m V_m^T$ with Y_m satisfying (8) be the approximate solution of the projected Lyapunov equation (4). Then,*

(a) the approximate solution $X_m = V_m Y_m V_m^T$ satisfies the second equation of (4) exactly, i.e.,

$$X_m = P_r X_m P_r^T;$$

(b) the norm of the residual matrix R_m can be computed by

$$\|R_m\|_F = \sqrt{2}\|h_{m+1,m} e_m^T Y_m\|_F. \tag{9}$$

Proof: By exploiting (2) and (3), we can easily obtain

$$(A^{-1}E)^i B_r = P_r (A^{-1}E)^i B_r, \quad i = 1, 2, \cdots.$$

Thus we have

$$\begin{aligned} &\mathcal{K}_m(A^{-1}E, B_r) \\ &= \mathrm{span}\{P_r B_r, P_r A^{-1} E B_r, \cdots, P_r (A^{-1}E)^{m-1} B_r\}, \end{aligned}$$

which together with $\mathcal{K}_m(A^{-1}E, B_r) = \mathrm{span}\{V_m\}$ shows that

$$P_r V_m = V_m.$$

The result $X_m = P_r X_m P_r^T$ follows immediately.

By substituting (5) into (6), we have

$$\begin{aligned} R_m &= (V_m H_m + h_{m+1,m} v_{m+1} e_m^T) Y_m V_m^T \\ &\quad + V_m Y_m (V_m H_m + h_{m+1,m} v_{m+1} e_m^T)^T \\ &\quad + \beta^2 V_m e_1 e_1^T V_m^T \\ &= V_m (H_m Y_m + Y_m H_m^T + \beta^2 e_1 e_1^T) V_m^T \\ &\quad + h_{m+1,m} v_{m+1} e_m^T Y_m V_m^T \\ &\quad + h_{m+1,m} V_m Y_m e_m v_{m+1}^T \\ &= V_{m+1} \begin{bmatrix} 0 & h_{m+1,m} Y_m e_m \\ h_{m+1,m} e_m^T Y_m & 0 \end{bmatrix} V_{m+1}^T. \end{aligned}$$

From the above expression for R_m, we obtain (9). □

The expression for the norm of the residual R_m given by (9) can be used to stop the iterations in the Krylov subspace method. The approximate solution X_m is computed only when convergence is achieved. This reduces the cost of this method.

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

Theorem 3 *Suppose that m steps of the Arnoldi process have been taken for $\mathcal{K}_m(A^{-1}E, B_r)$. Let $X_m = V_m Y_m V_m^T$ be the low-rank approximate solution of (4), where Y_m satisfies (8). Then*

$$\left\{ \begin{aligned} &(A^{-1}E - \Delta_m) X_m + X_m (A^{-1}E - \Delta_m)^T \\ &\qquad + P_r A^{-1} B B^T A^{-T} P_r^T = 0, \\ &X_m = P_r X_m P_r^T, \end{aligned} \right. \tag{10}$$

where $\Delta_m = h_{m+1,m} v_{m+1} v_m^T$ and $\|\Delta_m\|_F = |h_{m+1,m}|$.

Proof: We have

$$\begin{aligned} &A^{-1}EX_m + X_m(A^{-1}E)^T \\ &\quad + P_r A^{-1}BB^T A^{-T} P_r^T \\ &= A^{-1}EV_m Y_m V_m^T + V_m Y_m V_m^T (A^{-1}E)^T \\ &\quad + \beta^2 V_m e_1 e_1^T V_m^T \\ &= h_{m+1,m} v_{m+1} e_m^T Y_m V_m^T \\ &\quad + h_{m+1,m} V_m Y_m e_m v_{m+1}^T. \end{aligned} \tag{11}$$

The first equation of (10) follows by rearranging (11) and noting that $e_m^T = v_m^T V_m$. The expression for $\|\Delta_m\|_F$ follows from the fact that v_m and v_{m+1} are vectors with unit length. The second equation of (10) follows from Theorem 2. □

The Krylov subspace method for solving the projected generalized Lyapunov equation (1) is summarized as follows.

Algorithm 4 Krylov subspace method

1. *Choose a tolerance $\epsilon > 0$ and a positive integer k_1. Set $m = k_1$.*
2. *Construct an orthonormal basis $v_1, v_2 \cdots, v_m$ of the subspace $\mathcal{K}_m(A^{-1}E, B_r)$ by Algorithm 1.*
3. *Solve the low-dimensional problem $H_m Y_m + Y_m H_m^T + \beta^2 e_1 e_1^T = 0$ with $\beta = \|B_r\|_F$ by a direct method.*
4. *Compute the residual norm: $\|R_m\|_F = \sqrt{2}\|h_{m+1,m} e_m^T Y_m\|$. If $\|R_m\|_F < \epsilon$, form the approximate solution $X_m = V_m Y_m V_m^T$, and then stop.*
5. *Augment the orthonormal basis $v_1, v_2 \cdots, v_m$ of the subspace $\mathcal{K}_m(A^{-1}E, B_r)$ into an orthonormal basis $v_1, v_2 \cdots, v_{m+k_1}$ of the subspace $\mathcal{K}_{m+k_1}(A^{-1}E, B_r)$.*
6. *Set $m = m + k_1$ and go to step 3.*

In exact arithmetic, the Arnoldi process applied to the Krylov subspace $\mathcal{K}_m(A^{-1}E, B_r)$ will break down as $h_{m+1,m} = 0$. In this case, as shown in the following theorem, the exact solution of (4) is obtained.

Theorem 5 *Suppose that the Arnoldi process applied to $\mathcal{K}_m(A^{-1}E, B_r)$ breaks down at step m. Then we find the exact solution of (4).*

Proof: The result follows directly from the expression (9) for the norm of the residual matrix R_m. □

3 Extended Krylov subspace method

In this section, we will introduce a class of new subspaces, which will be employed to construct the projecting subspaces for solving the projected Lyapunov equation (4).

Suppose that the matrix $F \in \mathbb{R}^{n\times n}$ is invertible and $v \in \mathbb{R}^n$. The extended Krylov subspace $\mathbf{K}_m(F, v)$ is defined by

$$\mathbf{K}_m(F, v) \\ = \operatorname{span}\{v, F^{-1}v, Fv, F^{-2}v, \cdots, F^{m-1}v, F^{-m}v\}.$$

Note that the extended subspace $\mathbf{K}_m(F, v)$ contains information on both F and F^{-1}. Clearly, the extended Krylov subspace $\mathbf{K}_m(F, v)$ is the union of the Krylov subspace $\mathcal{K}_m(F, v)$ and $\mathcal{K}_m(F^{-1}, F^{-1}v)$, that is,

$$\mathbf{K}_m(F, v) = \mathcal{K}_m(F, v) \cup \mathcal{K}_m(F^{-1}, F^{-1}v).$$

Simoncini [31] proposed a project method based on the extended Krylov subspace $\mathbf{K}_m(A, B)$ for solving the standard Lyapunov equation

$$AX + XA^T = BB^T$$

with A being dissipative. This class of subspaces was also used by Druskin and Knizhnerman [11] for approximating matrix functions.

Due to the singularity of E, the inverse of $A^{-1}E$ does not exist. Therefore we can not apply the extended Krylov subspace method directly to the projected Lyapunov equation (4).

To overcome this disadvantage, we will use the $\{2\}$-inverse P of E to establish a class of new projection subspaces, which are also named as extended Krylov subspaces in this paper. Note that the $\{2\}$-inverse P of E can be formulated as

$$\begin{aligned} P &= P_r(EP_r + A(I - P_r))^{-1} \\ &= (P_l E + (I - P_l)A)^{-1} P_l \\ &= T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, \end{aligned}$$

see, for example, [33]. For the $\{2\}$-inverse of a singular matrix, the interesting reader is referred to [42].

With this preparation, we now define the extended Krylov subspace based on $A^{-1}E$ and B_r by

$$\mathbf{K}_m(A^{-1}E, B_r) = \mathcal{K}_m(A^{-1}E, B_r) \cup \mathcal{K}_m(PA, PAB_r).$$

In the following algorithm, an Arnoldi-like process is presented for establishing an orthonormal basis of the subspace $\mathbf{K}_m(A^{-1}E, B_r)$. We point out that this algorithm is a direct extension of the one proposed in [31] for constructing an orthonormal basis of an extended Krylov subspace based on an invertible matrix.

Algorithm 6 Arnoldi-like process

1. *Let $\widehat{V}_1 = [B_r, PAB_r]$.*
2. *Compute V_1 by the QR decomposition: $V_1 R = \widehat{V}_1$.*
3. *For $j = 1, 2, \cdots, m$*
4. *Set $V_j^{(1)} = V_j(:, 1)$ and $V_j^{(2)} = V_j(:, 2)$.*
5. $\widehat{V}_{j+1} = [A^{-1}EV_j^{(1)}, PAV_j^{(2)}]$.
6. *For $i = 1, 2, \cdots, j$*
7. $H_{ij} = V_i^T \widehat{V}_{j+1}$.
8. $\widehat{V}_{j+1} = \widehat{V}_{j+1} - V_i H_{ij}$.
9. *End For*
10. *Compute the QR decomposition $V_{j+1} H_{j+1,j} = \widehat{V}_{j+1}$.*
11. *End For*

The columns of the matrix $\mathcal{V}_m = [V_1, V_2, \cdots, V_m]$ with $V_j \in \mathbb{R}^{n\times 2}$ are an orthonormal basis of the subspace $\mathbf{K}_m(A^{-1}E, B_r)$.

Let $\mathcal{H}_n \in \mathbb{R}^{2m\times 2m}$ be the block upper Hessenberg matrix with each block 2 by 2, whose nonzero blocks are generated by Algorithm 6. It is easy to verify that for $j = 1, 2, \cdots, m$,

$$\widehat{V}_{j+1} = [A^{-1}EV_j^{(1)}, PAV_j^{(2)}] - \mathcal{V}_j\mathcal{H}_jE_j, \quad (12)$$

$$V_{j+1}H_{j+1,j} = \widehat{V}_{j+1}, \quad (13)$$

where $E_j^T = [0, 0, \cdots, I_2] \in \mathbb{R}^{2\times 2j}$.

The following theorem shows the relation between the subspace $A^{-1}E\mathbf{K}_m(A^{-1}E, B_r)$ and the extended Krylov subspace $\mathbf{K}_{m+1}(A^{-1}E, B_r)$.

Theorem 7 *For any $m \geq 1$, the space $\mathbf{K}_m(A^{-1}E, B_r)$ satisfies*

$$A^{-1}E\mathbf{K}_m(A^{-1}E, B_r) \subseteq \mathbf{K}_{m+1}(A^{-1}E, B_r).$$

Proof: Define $\widehat{B} = TA^{-1}B$, and partition $\widehat{B}$ appropriately as

$$\widehat{B} = \begin{bmatrix} \widehat{B}_1 \\ \widehat{B}_2 \end{bmatrix}.$$

By using (2) and (3), it is easy to obtain

$$B_r = T^{-1}\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} TA^{-1}B = T^{-1}\begin{bmatrix} \widehat{B}_1 \\ 0 \end{bmatrix},$$

$$(A^{-1}E)^iB_r = T^{-1}\begin{bmatrix} J^{-i}\widehat{B}_1 \\ 0 \end{bmatrix}, \quad i = 1, 2, \cdots,$$

$$(PA)^iB_r = T^{-1}\begin{bmatrix} J^{i}\widehat{B}_1 \\ 0 \end{bmatrix}, \quad i = 1, 2, \cdots.$$

Therefore, for $i = 1, 2, \cdots$, we have

$$\mathcal{K}_i(A^{-1}E, B_r) = T^{-1}\mathrm{span}\left\{\begin{bmatrix} \widehat{B}_1 \\ 0 \end{bmatrix}, \begin{bmatrix} J^{-1}\widehat{B}_1 \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} J^{-(i-1)}\widehat{B}_1 \\ 0 \end{bmatrix}\right\},$$

$$\mathcal{K}_i(PA, PAB_r) = T^{-1}\mathrm{span}\left\{\begin{bmatrix} J\widehat{B}_1 \\ 0 \end{bmatrix}, \begin{bmatrix} J^{2}\widehat{B}_1 \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} J^{i}\widehat{B}_1 \\ 0 \end{bmatrix}\right\},$$

$$A^{-1}E\mathcal{K}_i(A^{-1}E, B_r) = T^{-1}\mathrm{span}\left\{\begin{bmatrix} J^{-1}\widehat{B}_1 \\ 0 \end{bmatrix}, \begin{bmatrix} J^{-2}\widehat{B}_1 \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} J^{-i}\widehat{B}_1 \\ 0 \end{bmatrix}\right\},$$

$$A^{-1}E\mathcal{K}_i(PA, PAB_r) = T^{-1}\mathrm{span}\left\{\begin{bmatrix} \widehat{B}_1 \\ 0 \end{bmatrix}, \begin{bmatrix} J\widehat{B}_1 \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} J^{i-1}\widehat{B}_1 \\ 0 \end{bmatrix}\right\}.$$

From the definition of $\mathbf{K}_m(A^{-1}E, B_r)$ and the above equalities, it follows that

$$A^{-1}E\mathbf{K}_m(A^{-1}E, B_r) \subseteq \mathbf{K}_{m+1}(A^{-1}E, B_r).$$

□

Define $\widetilde{\mathcal{T}}_m = \mathcal{V}_{m+1}^TA^{-1}E\mathcal{V}_m$ and let $\mathcal{T}_m$ be the $2m \times 2m$ matrix obtained from $\widetilde{\mathcal{T}}_m$ by deleting the last 2 rows. We observe that the results of Theorem 7 ensure that $\widetilde{\mathcal{T}}_m$ is a block upper Hessenberg matrix, since $T_{ij} = V_i^TA^{-1}EV_j = 0$ for $i > j+1, j = 1, 2, \cdots, m$, i.e., $\widetilde{\mathcal{T}}_m$ has the form

$$\widetilde{\mathcal{T}}_m = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1m} \\ T_{21} & T_{22} & \cdots & T_{2m} \\ & T_{32} & \ddots & \vdots \\ & & \ddots & T_{mm} \\ & & & T_{m+1,m} \end{bmatrix},$$

where $T_{ij} = V_i^TA^{-1}EV_j \in \mathbb{R}^{2\times 2}$. So, we obtain the following relation

$$A^{-1}E\mathcal{V}_m = \mathcal{V}_{m+1}\widetilde{\mathcal{T}}_m = \mathcal{V}_m\mathcal{T}_m + V_{m+1}T_{m+1,m}E_m^T, \quad (14)$$

where $E_m^T = [0, 0, \cdots, I_2] \in \mathbb{R}^{2\times 2m}$.

Similar to the result in [31], there is a relation between $\widetilde{\mathcal{T}}_m$ and $\widetilde{\mathcal{H}}_m$, by which we can compute $\widetilde{\mathcal{T}}_m$ without additional matrix-vector products with $A^{-1}E$ and extra inner products of long vectors. This important relation is given in the following theorem, which is the same as Proposition 3.2 in [31].

Theorem 8 *Let $l^{(j)} = (l_{ik})$ be the 2×2 matrix such that $V_j = \widehat{V}_jl^{(j)}$, $j = 1, 2, \cdots, m$. Let*

$$\widetilde{\mathcal{T}}_m = (t_{ik})_{i=1,\cdots,2m+2, k=1,\cdots,2m},$$
$$\mathcal{H}_m = (h_{ik})_{i=1,\cdots,2m, j=1,\cdots,2m}.$$

Then (odd columns)

$$t_{:,2j-1} = h_{:,2j-1}, \quad j = 1, \cdots, m,$$

while (even columns)

$$(j=1) \quad \begin{aligned} & t_{:,2} = l_{11}^{(1)}(h_{:,1}l_{12}^{(1)} + e_1l_{22}^{(1)}), \\ & t_{:,4} = (e_2 - \widetilde{\mathcal{T}}_1h_{1:2,2})l_{22}^{(2)}, \\ & \rho^{(2)} = l_{12}^{(2)}l_{11}^{(2)}, \end{aligned}$$

$$(1 < j \leq n) \quad \begin{aligned} & t_{:,2j} = t_{:,2j} + t_{:,2j-1}\rho^{(j)}, \\ & t_{:,2j+2} = (e_{2j} - \widetilde{\mathcal{T}}_jh_{1:2j,2j})l_{22}^{(j+1)}, \\ & \rho^{(j+1)} = l_{12}^{(j+1)}l_{11}^{(j+1)}. \end{aligned}$$

Proof: For the odd columns of $\widetilde{\mathcal{T}}_m$, the proof is the same as that of Proposition 3.2 in [31].

We now consider the even columns of $\widetilde{\mathcal{T}}_m$. For $j \geq 1$, it follows from (12) and (13) that

$$PAV_j^{(2)} = \widehat{V}_{j+1}^{(2)} + \mathcal{V}_j \mathcal{H}_j e_{2j} = V_{j+1} H_{j+1,j} e_2 + \mathcal{V}_j \mathcal{H}_j e_{2j}. \tag{15}$$

By using

$$PA = T^{-1} \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} T, A^{-1}E = T^{-1} \begin{bmatrix} J^{-1} & 0 \\ 0 & N \end{bmatrix} T, \tag{16}$$

the relation (15) can be written as

$$T^{-1} \begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} TV_j^{(2)} = \widehat{V}_{j+1}^{(2)} + \mathcal{V}_j \mathcal{H}_j e_{2j},$$

which is equivalent to

$$\begin{bmatrix} J & 0 \\ 0 & 0 \end{bmatrix} TV_j^{(2)} = T\widehat{V}_{j+1}^{(2)} + T\mathcal{V}_j \mathcal{H}_j e_{2j}. \tag{17}$$

By using (2) and (3), it is easy to verify that for $j = 1, 2, \cdots, m$,

$$P_r \mathcal{V}_j = \mathcal{V}_j,$$

i.e.,

$$T^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T\mathcal{V}_j = \mathcal{V}_j. \tag{18}$$

Let the order of J be n_1. From (18), it follows that all the last $n - n_1$ elements of $TV_{j+1}^{(1)}, T\widehat{V}_{j+1}^{(2)}, T\widehat{V}_{j+1}^{(1)}$ and $T\widehat{V}_{j+1}^{(2)}$ are zeros for $j = 1, 2, \cdots, m$. Thus, (17) can be reformulated as

$$\begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} TV_j^{(2)} = T\widehat{V}_{j+1}^{(2)} + T\mathcal{V}_j \mathcal{H}_j e_{2j},$$

i.e.,

$$\begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}^{-1} T\widehat{V}_{j+1}^{(2)} = TV_j^{(2)} - \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}^{-1} T\mathcal{V}_j \mathcal{H}_j e_{2j}. \tag{19}$$

It follows from $P_r \mathcal{V}_j = \mathcal{V}_j$ that $P_r \widehat{V}_j^{(2)} = \widehat{V}_j^{(2)}$. Hence, by making use of (19), we have

$$\begin{aligned}
& \mathcal{V}_{m+1}^T A^{-1} E \widehat{V}_{j+1}^{(2)} \\
= \;& \mathcal{V}_{m+1}^T A^{-1} E P_r \widehat{V}_{j+1}^{(2)} \\
= \;& \mathcal{V}_{m+1}^T T^{-1} \begin{bmatrix} J^{-1} & 0 \\ 0 & 0 \end{bmatrix} T\widehat{V}_{j+1}^{(2)} \\
= \;& \mathcal{V}_{m+1}^T T^{-1} \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}^{-1} T\widehat{V}_{j+1}^{(2)} \\
= \;& \mathcal{V}_{m+1}^T T^{-1} \left(TV_j^{(2)} - \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}^{-1} T\mathcal{V}_j \mathcal{H}_j e_{2j} \right) \\
= \;& \mathcal{V}_{m+1}^T T^{-1} \left(TV_j^{(2)} - \begin{bmatrix} J^{-1} & 0 \\ 0 & 0 \end{bmatrix} T\mathcal{V}_j \mathcal{H}_j e_{2j} \right) \\
= \;& \mathcal{V}_{m+1}^T V_j^{(2)} - \mathcal{V}_{m+1}^T T^{-1} \begin{bmatrix} J^{-1} & 0 \\ 0 & 0 \end{bmatrix} T\mathcal{V}_j \mathcal{H}_j e_{2j} \\
= \;& e_{2j} - \mathcal{V}_{m+1}^T A^{-1} E \mathcal{V}_j \mathcal{H}_j e_{2j} \\
= \;& e_{2j} - \begin{bmatrix} \widetilde{\mathcal{T}}_j \\ 0 \end{bmatrix} \mathcal{H}_j e_{2j}.
\end{aligned}$$

Then, the result for the even columns of $\widetilde{\mathcal{T}}_m$ can be proved by the same argument as used for proving the even column case in [31, Proposition 3.2]. □

To solve the projected Lyapunov equation (4), we first apply Algorithm 6 to the extended Krylov subspace $\mathbf{K}_{m+1}(A^{-1}E, B_r)$ to obtain an orthonormal basis $\mathcal{V}_m$. Then, by solving the reduced equation

$$\mathcal{T}_m Y_m + Y_m \mathcal{T}_m^T + E_1 B_1 B_1^T E_1^T = 0, \tag{20}$$

we obtain an approximate solution $X_m = \mathcal{V}_m Y_m \mathcal{V}_m^T$. Here, $B_1 = V_1^T B_r$ and E_1 being the first 2 columns of the identity I_{2m}.

Concerning the approximate solution X_m and the residual matrix R_m generated by applying the extended Krylov subspace method to the projected Lyapunov equation (4), we have the following theorem, whose proof is similar to that of Theorem 2.

Theorem 9 *Suppose that m steps of the Arnoldi-like process have been taken for $\mathbf{K}_m(A^{-1}E, B_r)$. Let $X_m = \mathcal{V}_m Y_m \mathcal{V}_m^T$ with Y_m satisfying (20) be the approximate solution of the projected Lyapunov equation (4). Then,*

(a) the approximate solution $X_m = \mathcal{V}_m Y_m \mathcal{V}_m^T$ satisfies the second equation of (4) exactly, i.e.,

$$X_m = P_r X_m P_r^T;$$

(b) the norm of the residual matrix R_m can be formulated as

$$\|R_m\|_F = \sqrt{2} \|T_{m+1,n} E_m^T Y_m\|_F, \tag{21}$$

where E_m is the last 2 columns of the identity I_{2m}.

The following result shows that the approximate solution X_m generated by the extended Krylov subspace method is also an exact solution of a perturbed projected Lyapunov equation.

Theorem 10 *Suppose that m steps of the Arnoldi-like process have been taken for the extended Krylov subspace $\mathbf{K}_m(A^{-1}E, B_r)$. Let $X_m = V_m Y_m V_m^T$ be the low-rank approximate solution of (4), where Y_m satisfies (20). Then*

$$\left\{\begin{array}{l}(A^{-1}E-\Delta_m)X_m + X_m(A^{-1}E-\Delta_m)^T \\ \qquad +P_rA^{-1}BB^TA^{-T}P_r^T = 0, \\ X_m = P_rX_mP_r^T,\end{array}\right. \tag{22}$$

where $\Delta_m = V_{m+1}T_{m+1,m}V_m^T$ and $\|\Delta_m\|_F = \|T_{m+1,m}\|_F$.

Proof: We have

$$\begin{aligned} & A^{-1}EX_m + X_m(A^{-1}E)^T \\ & +P_rA^{-1}BB^TA^{-T}P_r^T \\ = \; & A^{-1}E\mathcal{V}_mY_m\mathcal{V}_m^T + \mathcal{V}_mY_m\mathcal{V}_m^T(A^{-1}E)^T \\ & +\mathcal{V}_mE_1B_1B_1^TE_1^T\mathcal{V}_m^T \\ = \; & V_{m+1}T_{m+1,m}E_m^TY_m\mathcal{V}_m^T \\ & +\mathcal{V}_mY_mE_mT_{m+1,m}^TV_{m+1}^T. \end{aligned} \tag{23}$$

The first equation of (22) follows by rearranging (23) and noting that $E_m^T = V_m^T\mathcal{V}_m$. The expression for $\|\Delta_m\|_F$ follows from the fact that V_m and V_{m+1} are matrices with orthonormal columns. The second equation of (22) follows from Theorem 9. □

The extended Krylov subspace method for solving the projected generalized Lyapunov equation (1) is summarized in the following algorithm.

Algorithm 11 Extended Krylov subspace method

1. *Choose a tolerance $\epsilon > 0$ and a positive integer k_1. Set $m = k_1$.*
2. *Construct an orthonormal basis $V_1, V_2 \cdots, V_m$ of the subspace $\mathbf{K}_m(A^{-1}E, B_r)$ by Algorithm 6.*
3. *Compute $\widetilde{\mathcal{T}}_m$ according to Theorem 8.*
4. *Solve the low-dimensional problem $\mathcal{T}_mY_m + Y_m\mathcal{T}_m^T + E_1B_1B_1^TE_1^T = 0$ by a direct method.*
5. *Compute the residual norm: $\|R_m\|_F = \sqrt{2}\|T_{m+1,n}E_m^TY_m\|_F$. If $\|R_m\|_F < \epsilon$, form the approximate solution $X_m = \mathcal{V}_mY_m\mathcal{V}_m^T$, and then stop.*
6. *Augment the orthonormal basis $V_1, V_2 \cdots, V_m$ of the subspace $\mathbf{K}_m(A^{-1}E, B_r)$ into an orthonormal basis $V_1, V_2 \cdots, V_{m+k_1}$ of the subspace $\mathbf{K}_{m+k_1}(A^{-1}E, B_r)$.*
7. *Set $m = m + k_1$ and go to step 3.*

The following result is useful for proving the property of finite termination of Algorithm 11 for solving the projected Lyapunov equation (4).

Lemma 12 *Suppose that $m-1$ steps of the Arnoldi-like process have been taken for $\mathbf{K}_m(A^{-1}E, B_r)$. At the mth step, assume that $\widehat{V}_{m+1}$ has rank less than two. Then we have*

$$A^{-1}E\mathcal{V}_m = \mathcal{V}_m\mathcal{T}_m,$$

or

$$A^{-1}E[\mathcal{V}_m, V_{m+1}^{(1)}] = [\mathcal{V}_m, V_{m+1}^{(1)}]\widehat{\mathcal{T}}_m,$$

where $\widehat{\mathcal{T}}_m$ is the restriction of $\mathcal{T}_{m+1}$ to the first $(2m+1)$ columns and rows.

Proof: By using the expressions for P_r and PA, $A^{-1}E$ in (3) and (16), and following the proof of Proposition 3.4 in [31], we can prove this lemma. □

The following theorem provides the result concerning the finite termination of Algorithm 11 for solving the projected Lyapunov equation (4).

Theorem 13 *Suppose that $m-1$ steps of the Arnoldi-like process have been taken for $\mathbf{K}_m(A^{-1}E, B_r)$. At the mth step, assume that $\widehat{V}_{m+1}$ has rank less than two. Then we can find the exact solution of (4).*

Proof: From Lemma 12, we have

$$A^{-1}E\mathcal{V}_m = \mathcal{V}_m\mathcal{T}_m,$$

or

$$A^{-1}E[\mathcal{V}_m, V_{m+1}^{(1)}] = [\mathcal{V}_m, V_{m+1}^{(1)}]\widehat{\mathcal{T}}_m,$$

where the columns of $\mathcal{V}_m$ and $[\mathcal{V}_m, V_{m+1}^{(1)}]$ are orthonormal.

For the first case $A^{-1}E\mathcal{V}_m = \mathcal{V}_m\mathcal{T}_m$, let

$$X_m = \mathcal{V}_mY_m\mathcal{V}_m^T,$$

where Y_m is the solution of $\mathcal{T}_mY_m + Y_m\mathcal{T}_m^T + E_1B_1B_1^TE_1^T = 0$. Then, we obtain

$$\begin{aligned} & A^{-1}EX_m + X_m(A^{-1}E)^T + B_rB^TA^{-T}P_r^T \\ = \; & A^{-1}E\mathcal{V}_mY_m\mathcal{V}_m^T + \mathcal{V}_mY_m\mathcal{V}_m^T(A^{-1}E)^T \\ & +\mathcal{V}_mE_1B_1B_1^TE_1^T\mathcal{V}_m^T \\ = \; & \mathcal{V}_m(\mathcal{T}_mY_m + Y_m\mathcal{T}_m^T + E_1B_1B_1^TE_1^T)\mathcal{V}_m^T \\ = \; & 0. \end{aligned}$$

For the case $A^{-1}E[\mathcal{V}_m, V_{m+1}^{(1)}] = [\mathcal{V}_m, V_{m+1}^{(1)}]\widehat{\mathcal{T}}_m$, the proof is similar to the first case. This completes the proof. □

4 Numerical experiments

In this section, we present three numerical examples to illustrate the performance of the Krylov subspace method (Algorithm 4) and the extended Krylov subspace method (Algorithm 11) for the projected generalized Lyapunov equation (1). Algorithm 4 and Algorithm 11 are denoted by KS and EKS, respectively. For the purpose of comparison, we also present the test results obtained by the generalized low-rank alternating direction implicit method (denoted by LR-ADI) proposed in [40].

In the following examples, we compare the numerical behavior of these three methods with respect to the dimension of computed subspace (DIM), CPU time (in seconds) and the relative residual (RES). Here the relative residual is defined by

$$\mathrm{RES} = \frac{\|EX_mA^T + AX_mE^T + P_lBB^TP_l^T\|_F}{\|P_lBB^TP_l^T\|_F},$$

where X_m denotes the approximate solution obtained by KS, EKS, or LR-ADI.

We express the approximate solution X_m in the low-rank form, i.e.,

$$X_m = Z_mZ_m^T.$$

For the KS method, $Z_m = V_mL_m$, where $L_m \in \mathbb{R}^{m\times m}$ is the Cholesky factor of the solution of the reduced Lyapunov equation (8), while for the EKS method, $Z_m = \mathcal{V}_mL_m$, where $L_m \in \mathbb{R}^{2m\times 2m}$ is the Cholesky factor of the solution of the reduced Lyapunov equation (20). The existence of the Cholesky factors for the solutions of (8) and (20) requires H_m in (8) and $\mathcal{T}_m$ in (20) to be stable, respectively. Note that this does not hold for general cases. However, the following numerical experiments shows that it holds true for our examples. For the LR-ADI method, Z_m is the low-rank factor generated by m steps of LR-ADI.

For the KS method and the EKS method, we need to solve linear systems with the coefficient matrix A, while for the LR-ADI method, we require to solve linear systems with $E - \mu A$, where μ is one of the ADI shift parameters. Note that for different iteration steps, the ADI shift parameter may be different. In our tests, we employ the restarted GMRES [29, 30] to solve the corresponding linear systems. The preconditioner for GMRES is constructed by the incomplete LU factorization of the coefficient matrix with threshold 0.01. In all the tests, we use GMRES(20) with tolerance 10^{-12}.

We use the heuristic algorithm proposed by Penzl [26] to compute the suboptimal shift parameters for the LR-ADI method. This algorithm is based on the Arnoldi iterations [29] applied to the matrices $A^{-1}E$ and PA, see [40] for the details. In the following tests, we use 15 shift parameters for the LR-ADI method. If the number of shift parameters is smaller than the number of iterations required to obtain a prescribed tolerance, then we reuse these parameters in a cyclic manner.

We will use

$$\|R_m\|_F = \sqrt{2}\|h_{m+1,m}e_m^TY_m\|_F < 10^{-10}$$

as the stopping criterion for the KS method, and

$$\|R_m\|_F = \sqrt{2}\|T_{m+1,m}E_m^TY_m\|_F < 10^{-10}$$

as the stopping criterion for the EKS method. Since the residual for the LR-ADI method does not admit such a simple expression as the KS method or the EKS method, we use

$$\frac{\|z_m\|}{\|Z_{m-1}\|_F} < 10^{-5}$$

as the stopping criterion for the LR-ADI method. Here, z_m is the update generated at step m in the LR-ADI method. For the details of the LR-ADI method, the interesting reader is referred to [40].

For these three methods, the low-rank expression $X_m = Z_mZ_m^T$ can be employed to calculate the relative residual RES, see, for example, [26].

All numerical experiments are performed on an Intel Pentium Dual E2160 with CPU 1.80GHz and RAM 2GB under the Window XP operating system and the usual double precision, where the floating point relative accuracy is 2.22×10^{-16}.

4.1 Example 1

For the first experiment, we consider the 2D instationary Stokes equation that describes the flow of an incompressible fluid

$$\begin{aligned} \frac{\partial x}{\partial t} &= \Delta x - \nabla\rho + f, \quad (\xi,t) \in \Omega \times (0,t_e), \\ \mathrm{div}x &= 0, \quad (\xi,t) \in \Omega \times (0,t_e) \end{aligned}$$

with appropriate initial and boundary conditions. Here $x(\xi,t) \in \mathbb{R}^2$ is the velocity vector, $\rho(\xi,t) \in \mathbb{R}$ is the pressure, $f(\xi,t) \in \mathbb{R}^2$ is the vector of external forces, $\Omega \subset \mathbb{R}^2$ is a bounded open domain, and $t_e > 0$ is the endpoint of the considered time interval.

The spatial discretization of this equation by the finite difference method on a uniform staggered grid leads to the descriptor system

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t). \end{aligned} \qquad (24)$$

This example for the projected generalized Lyapunov equation was presented by Stykel, see [38, 39, 40] and the references therein. The coefficient matrices in (24) are given by

$$E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}, A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} \in \mathbb{R}^{n\times n},$$

where $A_{11} = A_{11}^T$ and $A_{12} = A_{21}^T$. Since A is symmetric and E is positive semidefinite, the finite eigenvalues of $\lambda E - A$ are real. These matrices are sparse and have special block structure. Using this structure, the projectors P_l and P_r onto the left and right deflating subspaces of the pencil $\lambda E - A$ can be expressed as

$$P_l = \begin{bmatrix} \Pi & -\Pi A_{11}A_{12}(A_{21}A_{12})^{-1} \\ 0 & 0 \end{bmatrix},$$

$$P_r = \begin{bmatrix} \Pi & 0 \\ -(A_{21}A_{12})^{-1}A_{21}A_{11}\Pi & 0 \end{bmatrix},$$

where $\Pi = I - A_{12}(A_{21}A_{12})^{-1}A_{21}$ is a projector onto the kernel of A_{21} along the image of A_{12}. The product PA in this case is given by

$$PA = \begin{bmatrix} \Pi A_{11}\Pi & 0 \\ -(A_{21}A_{12})^{-1}A_{21}A_{11}\Pi A_{11}\Pi & 0 \end{bmatrix}.$$

We use the spatial discretization of the Stokes equation on the square domain $[0, 1] \times [0, 1]$ to lead to problems of order $n = 7700, 14559, 30400$. The matrix $B \in \mathbb{R}^n$ is chosen at random.

In the case $n = 14559$, $A_{21} = A_{12}^T \in \mathbb{R}^{4899\times 9660}$, $A_{11} \in \mathbb{R}^{9660\times 9660}$, the matrix A has 67336 nonzero elements, and $A_{21}A_{12}$ has 24215 nonzero elements. Figure 1 and 2 shows the sparsity structures of the matrices A and $A_{21}A_{12}$, respectively.

The numerical results for Example 1 are reported in Table 1. Table 1 shows that for this example, the LR-ADI method needs the least dimensional subspace for reaching the the stopping criterion while in terms of the CPU time, the EKS method with the inner iterative method GMRES(20) has the best performance. It clearly indicates that the EKS method is more efficient than the LR-ADI method for this example. The EKS method requires less subspace dimension than the KS method for this example.

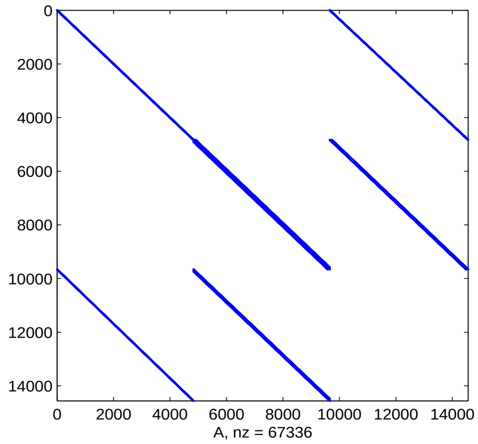

Figure 1: Example 5.1. Sparsity structures of the matrices A.

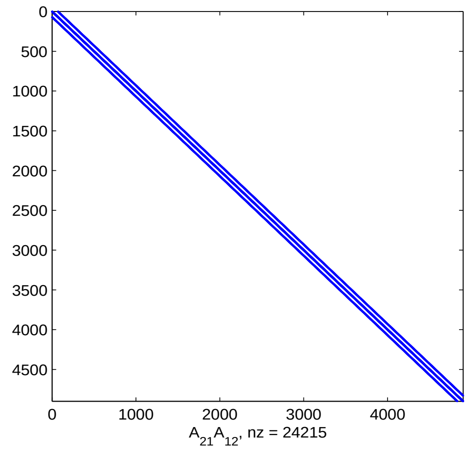

Figure 2: Example 5.1. Sparsity structures of the matrices $A_{21}A_{12}$.

4.2 Example 2

For the second experiment, we consider a holonomically constrained damped mass-spring system with g masses as in [39]. The ith mass m_i is connected to the $(i + 1)$th mass m_{i+1} by a spring and a damper with constants k_i and d_i, and also to the ground by another spring and damper with constants κ_i and δ_i. Moreover, the first mass is connected to the last one by a rigid bar and it can be influenced by a control. The vibration of this system is described by the descriptor system (24) with the matrices

$$E = \begin{bmatrix} I & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & I & 0 \\ K & D & -N^T \\ N & 0 & 0 \end{bmatrix},$$

where $M = \mathrm{diag}(m_1, m_2, \cdots, m_g)$ is the symmetric positive definite mass matrix, $K \in \mathbb{R}^{g\times g}$ is the tridiagonal stiffness matrix, $D \in \mathbb{R}^{g\times g}$ is the tridiagonal damping matrix, and N is the matrix of constraints.

The spectral projectors P_l, P_r and the $\{2\}$-inverse P of E can be expressed by the blocks of E and A, see [40].

Table 1: Example 1. Performance comparison of KS, EKS and LR-ADI. Linear systems with A or $E - \mu A$ were solved by GMRES(20). Incomplete LU factorization with threshold 0.01 was used to construct preconditioners for GMRES(20) with tolerance 10^{-12}.

n	Method	DIM	CPU	RES
7700	KS	76	26.32	6.5582e-10
	EKS	50	11.59	6.6069e-11
	LR-ADI	20	19.07	1.6664e-10
14559	KS	86	79.46	5.8458e-10
	EKS	56	27.76	1.0602e-10
	LR-ADI	23	54.60	3.0183e-10
30400	KS	106	361.74	2.3407e-10
	EKS	64	125.31	3.1694e-10
	LR-ADI	32	243.23	4.1225e-10

In this experiment we take

$$\begin{aligned} m_1 &= m_2 = \cdots = m_g = 100, \\ k_1 &= k_2 = \cdots = k_{g-1} = 2, \\ \kappa_2 &= \kappa_3 = \cdots = \kappa_{g-1} = 2, \quad \kappa_1 = \kappa_g = 4, \\ d_1 &= d_2 = \cdots = d_{g-1} = 5, \\ \delta_2 &= \delta_3 = \cdots = \delta_{g-1} = 5, \quad \delta_1 = \delta_g = 10. \end{aligned}$$

For $g = 2000, 6000, 10000$, we obtain three descriptor systems of order $n = 4001, 12001, 200001$ with $B \in \mathbb{R}^n$.

The numerical results for Example 2 are reported in Table 2. Table 2 shows that for this example, in terms of the subspace dimension, the LR-ADI method is the best one while in terms of the CPU time, the EKS method has the best performance. Moreover, EKS+LU needs less CPU time than EKS+GMRES(20) for $n = 4001, 12001$. We also note that the EKS method requires the same subspace dimension as the KS method for this example.

4.3 Example 3

We now do the same experiment as in Example 2 except that

$$\begin{aligned} m_1 &= m_2 = \cdots = m_g = 100, \\ k_1 &= k_2 = \cdots = k_{g-1} = 2, \\ \kappa_1 &= \kappa_2 = \cdots = \kappa_g = 4, \\ d_1 &= d_2 = \cdots = d_{g-1} = 3, \\ \delta_1 &= \delta_2 = \cdots = \delta_g = 7. \end{aligned}$$

Table 2: Example 2. Performance of comparison of KS, EKS and LR-ADI. Linear systems with A or $E - \mu A$ were solved by GMRES(20). Incomplete LU factorization with threshold 0.01 was used to construct preconditioners for GMRES(20) with tolerance 10^{-12}.

n	Method	DIM	CPU	RES
4001	KS	40	2.03	1.0090e-09
	EKS	40	1.67	8.5209e-10
	LR-ADI	23	44.56	1.1226e-10
12001	KS	40	4.17	1.0090e-09
	EKS	40	3.82	8.5209e-10
	LR-ADI	23	252.95	1.1226e-10
20001	KS	40	7.68	1.0090e-09
	EKS	40	7.35	8.5209e-10
	LR-ADI	23	597.12	1.1226e-10

The numerical results for Example 3 are presented in Table 3. Table 3 shows that for this example, the EKS method with the inner iterative method GMRES(20) has the best performance in terms of the CPU time.

5 Conclusions

In this paper, we have proposed two iterative methods, the Krylov subspace method and the extended Krylov subspace method, to solve the projected continuous-time generalized Lyapunov equation. It has been shown that every one of the iterates generated by these methods satisfies the projection condition. Moreover, the residuals for each of these methods have simple expressions. Numerical experiments are presented for the performance comparison between the newly proposed iterative methods and the LR-ADI method. It is shown that in terms of the CPU time, the extended Krylov subspace outperforms the LR-ADI method.

Acknowledgements: The authors would like to thank Professor Gen-Qi Xu and the referees for their useful comments. Yujian Zhou is supported by the Science and Technology Planning Project of Hunan Province (grant No. 2010JT4042). Yiqin Lin is supported by the National Natural Science Foundation of China (grant No. 10801048), the Natural Science Foundation of Hunan Province (grant No. 11JJ4009), the Scientific Research Foundation of Education Bureau of Hunan Province for Outstanding Young Scholars in University (grant No. 10B038), and the

Table 3: Example 3. Performance of comparison of KS, EKS and LR-ADI. Linear systems with A or $E - \mu A$ were solved by GMRES(20). Incomplete LU factorization with threshold 0.01 was used to construct preconditioners for GMRES(20) with tolerance 10^{-12}.

n	Method	DIM	CPU	RES
4001	KS	34	1.73	7.9706e-10
	EKS	34	1.31	7.1816e-10
	LR-ADI	22	40.07	7.7008e-12
12001	KS	34	4.01	7.9706e-10
	EKS	34	3.46	7.1816e-10
	LR-ADI	22	215.28	7.7008e-12
20001	KS	34	7.64	7.9706e-10
	EKS	34	7.03	7.1816e-10
	LR-ADI	22	513.22	7.7008e-12

Chinese Postdoctoral Science Foundation (grant No. 2012M511386). Liang Bao is supported by the National Natural Science Foundation of China (grants No. 10926150 and No. 11101149).

References:

[1] B. Anderson and J. Moore, *Optimal Control–Linear Quadratic Methods*, Prentice–Hall, Englewood Cliffs, NJ 1990.

[2] L. Bao, Y. Lin and Y. Wei, Krylov subspace methods for the generalized Sylvester equation, *Appl. Math. Comput.* 175, 2006, pp. 557–573.

[3] L. Bao, Y. Lin and Y. Wei, A new projection method for solving large Sylvester equations, *Appl. Numer. Math.* 57, 2007, pp. 521–532.

[4] R. H. Bartels and G. W. Stewart, Solution of the equation $AX + XB = C$, *Comm. ACM* 15, 1972, pp. 820–826.

[5] U. Baur and P. Benner, Factorized solution of Lyapunov equations based on hierarchical matrix arithmetic, *Computing* 78, 2006, pp. 211–234.

[6] P. Benner, R. Li and N. Truhar, On the ADI method for Sylvester equations, *J. Comput. Appl. Math.* 233, 2009, pp. 1035–1045.

[7] P. Benner, H. Mena and J. Saak, On the parameter selection problem in the Newton-ADI iteration for large-scale Riccati equations, *Elect. Trans. Numer. Anal.* 29, 2008, pp. 136–149.

[8] P. Benner and E. S. Quintana-Ortí, Solving stable generalized Lyapunov equations with the matrix sign function, *Numer. Algorithms* 20, 1999, pp. 75–100.

[9] P. Benner, E. S. Quintana-Ortí and G. Quintana-Ortí, Solving stable Sylvester equations via rational iterative schemes, *J. Sci. Comput.* 28, 2006, pp. 51–83.

[10] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia 1997.

[11] V. Druskin and L. Knizhnerman, Extended Krylov subspaces: Approximation of the matrix square root and related functions, *SIAM J. Matrix Anal. Appl.* 19, 1998, pp. 755–771

[12] F. Gantmacher, *Theory of Matrices*, Chelsea, New York 1959.

[13] J. Gardiner, A. Laub, J. Amato and C. Moler, Solution of the Sylvester matrix equation $AXB^T + CXD^T = E$, *ACM Trans. Math. Software* 18, 1992, pp. 223–231.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, John Hopkins University Press, Baltimore, MD 1996.

[15] S. Gugercin, D. C. Sorensen and A. C. Antoulas, A modified low-rank Smith method for large-scale Lyapunov equations, *Numer. Algorithms* 32, 2003, pp. 27–55.

[16] S. J. Hammarling, Numerical solution of the stable non-negative definite Lyapunov equation, *IMA J. Numer. Anal.* 2, 1982, pp. 303–323.

[17] U. Helmke and J. Moore, *Optimization and Dynamical Systems*, Springer, London 1994.

[18] D. Y. Hu and L. Reichel, Krylov-subspace methods for the Sylvester equation, *Linear Algebra Appl.* 172, 1992, pp. 283–313.

[19] I. M. Jaimoukha and E. M. Kasenally, Krylov subspace methods for solving large Lyapunov equations, *SIAM J. Numer. Anal.* 31, 1994, pp. 227–251.

[20] K. Jbilou and A. J. Riquet, Projection methods for large Lyapunov matrix equations, *Linear Algebra Appl.* 415, 2006, pp. 344–358.

[21] D. Kressner, Block variants of Hammarling's method for solving Lyapunov equations, *ACM Trans. Math. Software* 34, 2008, pp. 1–15.

[22] J. Li and J. White, Low rank solution of Lyapunov equations, *SIAM J. Matrix Anal. Appl.* 24, 2002, pp. 260–280.

[23] A. Lu and E. Wachspress, Solution of Lyapunov equations by alternating direction implicit iteration, *Comput. Math. Appl.* 21, 1991, pp. 43–58.

[24] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*, Springer, Heidelberg 1991.

[25] T. Penzl, Numerical solution of generalized Lyapunov equations, *Adv. Comput. Math.* 8, 1998, pp. 33–48.

[26] T. Penzl, A cyclic low-rank Smith method for large sparse Lyapunov equations, *SIAM J. Sci. Comput.* 21, 2000, pp. 1401–1418.

[27] P. Petkov, N. Christov and M. Konstantinov, *Computational Methods for Linear Control Systems*, Prentice-Hall, Hertfordshire, UK 1991.

[28] A. Ruhe, Numerical aspects of Gram-Schmidt orthogonalization of vectors, *Linear Algebra Appl.* 52, 1983, pp. 591–601.

[29] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia 2003.

[30] Y. Saad and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.* 7, 1986, pp. 856–869.

[31] V. Simoncini, A new iterative method for solving large-scale Lyapunov matrix equations, *SIAM J. Sci. Comput.* 29, 2007, pp. 1268–1288.

[32] R. Smith, Matrix equation $XA + BX = C$, *SIAM J. Appl. Math.* 16, 1968, pp. 198–201.

[33] V. I. Sokolov, *Contributions to the Minimal Realization Problem for Descriptor Systems*, Ph.D. thesis, Fakultät für Mathematik, Technische Universität Chemnitz, Chemnitz 2006.

[34] D. C. Sorensen and Y. Zhou, Direct methods for matrix Sylvester and Lyapunov equations, *J. Appl. Math.* 6, 2003, pp. 277–303.

[35] T. Stykel, Numerical solution and perturbation theory for generalized Lyapunov equations, *Linear Algebra Appl.* 349, 2002, pp. 155–185.

[36] T. Stykel, Stability and inertia theorems for generalized Lyapunov equations, *Linear Algebra Appl.* 355, 2002, pp. 297–314.

[37] T. Stykel, Gramian-based model reduction for descriptor systems, *Math. Control Signals Systems* 16, 2004, pp. 297–319.

[38] T. Stykel, Balanced truncation model reduction for semidiscretized Stokes equation, *Linear Algebra Appl.* 415, 2006, pp. 262–289.

[39] T. Stykel, *A modified matrix sign function method for projected Lyapunov equations*, *Systems Control Lett.* 56, 2007, pp. 695–701.

[40] T. Stykel, Low-rank iterative methods for projected generalized Lyapunov equations, *Elect. Trans. Numer. Anal.* 30, 2008, pp. 187–202.

[41] E. Wachspress, Iterative solution of the Lyapunov matrix equation, *Appl. Math. Lett.* 107, 1988, pp. 87–90.

[42] G. Wang, Y. Wei and S. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing/New York 2004.