SIAM J. MATRIX ANAL. APPL.
Vol. 24, No. 1, pp. 260–280
© 2002 Society for Industrial and Applied Mathematics

LOW RANK SOLUTION OF LYAPUNOV EQUATIONS*

JING-REBECCA LI† AND JACOB WHITE‡

Abstract. This paper presents the Cholesky factor–alternating direction implicit (CF–ADI) algorithm, which generates a low rank approximation to the solution X of the Lyapunov equation $AX + XA^T = -BB^T$. The coefficient matrix A is assumed to be large, and the rank of the right-hand side $-BB^T$ is assumed to be much smaller than the size of A. The CF–ADI algorithm requires only matrix-vector products and matrix-vector solves by shifts of A. Hence, it enables one to take advantage of any sparsity or structure in A.

This paper also discusses the approximation of the dominant invariant subspace of the solution X. We characterize a group of spanning sets for the range of X. A connection is made between the approximation of the dominant invariant subspace of X and the generation of various low order Krylov and rational Krylov subspaces. It is shown by numerical examples that the rational Krylov subspace generated by the CF–ADI algorithm, where the shifts are obtained as the solution of a rational minimax problem, often gives the best approximation to the dominant invariant subspace of X.

Key words. Lyapunov equation, alternating direction implicit iteration, low rank approximation, dominant invariant subspace, iterative methods

AMS subject classifications. 65F30, 65F10, 15A24, 93C05

PII. S0895479801384937

1. Introduction.

In this paper we present the Cholesky factor–alternating direction implicit (CF–ADI) algorithm, which is well suited to solving large-scale Lyapunov equations whose right-hand sides have low rank. A Lyapunov equation has the form

$$AX + XA^T = -BB^T, \qquad A \in \mathbb{R}^{n\times n},\ X \in \mathbb{R}^{n\times n}. \tag{1.1}$$

The unknown is the matrix X. We assume that the coefficient matrix A is large and stable, $\lambda_i(A) < 0\ \forall i$. Furthermore, we assume that the rank of the right-hand side $-BB^T$ is much smaller than n, or simply, $\operatorname{rank}(B) = r_b \ll n$. When A is stable, the matrix X is symmetric from the uniqueness of the solution to (1.1), and it is positive semidefinite [18]. Such Lyapunov equations occur in the analysis and model reduction of large, linear, time-invariant systems, where the number of inputs and the number of outputs are small compared to the system size.

The first contribution of this paper is the CF–ADI algorithm, which is a reformulation of the alternating direction implicit (ADI) algorithm for Lyapunov equations [5, 8, 23, 38, 39, 40] and gives exactly the same approximation. However, CF–ADI requires only matrix-vector products and matrix-vector solves by shifts of A. Hence, it enables one to take advantage of any sparsity or structure in the coefficient matrix

*Received by the editors February 12, 2001; accepted for publication (in revised form) by V. Mehrmann January 15, 2002; published electronically July 9, 2002.

http://www.siam.org/journals/simax/24-1/38493.html

†Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, Room 1118, New York, NY 10012-1185 (jingli@cims.nyu.edu). The research of this author was supported by the National Science Foundation and the Semiconductor Research Corporation/Hewlett-Packard Graduate Fellowships.

‡Research Laboratory of Electronics, Massachusetts Institute of Technology, Room 36-817, Cambridge, MA 02139-4307 (white@mit.edu).

A. The CF–ADI algorithm is intended to be used as a low rank algorithm to provide a low rank approximation to the exact solution matrix X. Frequently the exact solution X itself has low numerical rank [1, 28].

For some applications, it is sufficient to find the dominant invariant subspace of X. The complete knowledge of X is not necessary. For example, in the linear systems setting, the dominant invariant subspace of X may have physical meaning either as the span of the directions most sensitive to input or as the span of the directions to which the output is the most sensitive (see [6, 9, 33]). In fact, knowledge of the dominant invariant subspace of X is enough to produce the balanced truncation reduced model [26, 29] for symmetric systems [20, 21]. Hence, for some applications, approximating the dominant invariant subspace of X is as relevant as approximating X itself. In light of this, the second half of this paper is devoted to the approximation of the dominant invariant subspace of X.

The second contribution of this paper is making the connection between the approximation of the dominant invariant subspace of X and the generation of various low order Krylov and rational Krylov subspaces. It is shown that various methods of generating low rank approximations to X, including the CF–ADI algorithm, involve finding a low order Krylov or rational Krylov subspace to approximate the dominant invariant subspace of X. All these subspaces, when taken to order n, yield the full range of X. We compare the CF–ADI choice of a rational Krylov subspace, where the shifts are obtained by solving a rational minimax problem, with several other natural choices. We show by numerical examples that the subspace generated by CF–ADI often provides the best approximation to the dominant invariant subspace of X.

A preliminary form of the CF–ADI algorithm as applied to the model reduction problem can be found in [20, 21, 22]. In this paper we give details of the CF–ADI algorithm as relevant to the solution of (1.1). We also include complexity analysis, parameter selection procedure, stopping criteria, the use of real arithmetic, and numerical results on convergence, all of which appear for the first time in literature. Some early numerical results on using CF–ADI to approximate the dominant invariant subspace of X can be found in [22].

It has come to the authors' attention that another low rank reformulation of the ADI algorithm was independently proposed in [27]. However, in that version, the work required to produce a rank k approximation to X increases as $O(k^2)$, whereas for the CF–ADI algorithm presented in this paper, the work increases as $O(k)$. In fact, the algorithm in [27] appears as an intermediate step in deriving the final CF–ADI algorithm.

This paper is organized in the following way. Section 2 motivates the solution of the Lyapunov equation and the approximation of the dominant invariant subspace of the solution in the context of linear, time-invariant systems. Section 3 provides background on existing approaches to the solution of (1.1), including the ADI algorithm in some detail. Section 4 develops the CF–ADI algorithm. Section 5 contains a collection of definitions and useful results concerning Krylov and rational Krylov subspaces. Section 6 characterizes spanning sets for a subspace based on A and B. Section 7 shows that these spanning sets also span the range of X and uses that result to prove several properties of CF–ADI. Section 8 makes the connection between the approximation of the dominant invariant subspace of X and the generation of various low order Krylov and rational Krylov subspaces. We also make numerical comparisons of several different Krylov and rational Krylov subspace approximations. Section 9 contains the conclusions.

2. Motivation. Lyapunov equations with a low rank right-hand side occur in the analysis and model reduction of large, linear, time-invariant systems, where the system size is much larger than the number of inputs and the number of outputs. In this paper we focus on systems whose coefficient matrices are large and sparse. Such systems occur in interconnect modeling, solutions of PDEs, and other applications.

A linear, time-invariant system with realization (A, B, C) is characterized by the equations

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t), \tag{2.1}$$

$$y(t) = Cx(t). \tag{2.2}$$

The vector valued function $x(t) : \mathbb{R} \mapsto \mathbb{R}^n$ gives the state at time t and has n components. The input $u(t) : \mathbb{R} \mapsto \mathbb{R}^{r_b}$ and the output $y(t) : \mathbb{R} \mapsto \mathbb{R}^{r_c}$ have r_b and r_c components, respectively. The matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times r_b}$, $C \in \mathbb{R}^{r_c\times n}$ are the system matrix, the input coefficient matrix, and the output coefficient matrix, respectively. For single-input single-output (SISO) systems, $r_b = 1, r_c = 1$. Even for multiple-input, multiple-output (MIMO) systems, r_b and r_c are usually both very small compared to n.

If the system matrix A is stable, i.e., all the eigenvalues of A are in the open left half plane, then the controllability Gramian $P \in \mathbb{R}^{n\times n}$ and the observability Gramian $Q \in \mathbb{R}^{n\times n}$ associated with the system in (2.1)–(2.2) are the unique, symmetric, and positive semidefinite solutions to the following two Lyapunov equations (see, e.g., [6, 9, 18, 33]):

$$AP + PA^T = -BB^T, \tag{2.3}$$

$$A^TQ + QA = -C^TC. \tag{2.4}$$

If the number of inputs r_b is much smaller than the number of state components n, then $\operatorname{rank}(BB^T) = \operatorname{rank}(B) \le r_b \ll n$, and the right-hand side of (2.3) has low rank. Similarly, if the number of outputs r_c is much smaller than n, then the right-hand side of (2.4) has low rank.

The physical importance of the dominant eigenvectors of the Gramians P and Q is that they are the directions most sensitive to the input and the directions to which the output is the most sensitive, respectively (see [6, 9, 33]). In addition, for symmetric systems, where $A = A^T$ and $C = B^T$ in (2.1)–(2.2) and (2.3) and (2.4) are the same, knowledge of the dominant invariant subspace of $P = Q$ is sufficient to produce the balanced truncation reduced model [26, 29] for the system [20, 21]. Therefore, for some applications, approximating the dominant invariant subspace of the solution to (1.1) is as relevant as approximating the solution itself.

3. Previous methods. This section describes several existing methods for finding or approximating the solution X to the Lyapunov equation (1.1). Several of the algorithms described in this paper utilize the Cholesky factors of square matrices, and we give the definition of the Cholesky factor below.

DEFINITION 3.1. *A matrix Z is a* Cholesky factor *of $X \in \mathbb{R}^{n\times n}$ if it satisfies*

$$X = ZZ^T. \tag{3.1}$$

In this paper, the Cholesky factor Z is not required to be a square matrix nor have lower triangular structure.

The Bartels–Stewart method [2] first transforms A to real Schur form and then back solves for the solution of the transformed Lyapunov equation. The solution X is then obtained by a congruence transformation. Reducing a general, possibly sparse matrix to real Schur form requires $O(n^3)$ work, as does the congruence transformation to produce X.

The Hammarling method [12] also first transforms A to Schur form and has $O(n^3)$ complexity. It computes the lower triangular matrix Cholesky factor of the solution X rather than X itself.

The matrix sign function method [3, 30] exploits a simple connection between X and the matrix sign function of the $2n \times 2n$ matrix $\left[\begin{smallmatrix} A^T & 0 \\ BB^T & -A \end{smallmatrix}\right]$. The latter is found by Newton iteration. The complexity of this approach depends on the speed of the convergence of the Newton iteration but is at best $O(n^3)$. Low rank versions of the matrix sign function method can be found in [4, 19].

An approximate power iteration algorithm to determine the dominant invariant subspace of X is contained in [13], where approximations to the matrix-vector products Xv are computed. At each iteration, a Sylvester equation with a large left coefficient matrix and a small right coefficient matrix must be solved.

The low rank Smith(l) method in [27] gives the same approximation as the ADI method with cyclic parameters, and exploits the low rank of the right-hand side of the Lyapunov equation, but it is not as efficient as the CF–ADI algorithm to be derived in section 4. The main reason is that it is dependent on a low rank implementation of the ADI algorithm which is given in this paper in (4.6)–(4.7) and which is only an intermediate step in deriving the final CF–ADI algorithm.

3.1. Alternating direction implicit iteration. The ADI method [5, 39, 40, 41] is an iterative method and is given as Algorithm 1. The parameters $\{p_1, p_2, \ldots, p_J\}$, $\mathrm{Re}\{p_j\} < 0$, are called the ADI parameters. To keep the final ADI approximation

Algorithm 1. Alternating direction implicit algorithm.

INPUT: A, B.

1. If $v \mapsto Av, v \in \mathbb{R}^n$, is not $O(n)$ work, tridiagonalize A.
   a. Find $\tilde{A}$ tridiagonal, such that $\tilde{A} = SAS^{-1}$.
   b. Set $\tilde{B} := SB$.

   Otherwise, set $\tilde{A} := A, \tilde{B} := B$.
2. Choose ADI parameters, $\{p_1, \ldots, p_J\}$, $\mathrm{Re}\{p_i\} < 0$, (real or complex conjugate pairs), according to section 3.1.1 and references, using spectral bounds on $\tilde{A}$.
3. Initial guess,
   $$\tilde{X}_0 = 0_{n\times n}. \tag{3.2}$$
4. FOR $j = 1, 2, \ldots, J$, DO
   $$(\tilde{A} + p_j I)\tilde{X}_{j-\frac{1}{2}} = -BB^T - \tilde{X}_{j-1}(\tilde{A}^T - p_j I), \tag{3.3}$$
   $$(\tilde{A} + p_j I)\tilde{X}_j = -BB^T - \tilde{X}^T_{j-\frac{1}{2}}(\tilde{A}^T - p_j I). \tag{3.4}$$
   END
5. If A was tridiagonalized, recover solution,
   $$X^{adi}_J := S^{-1}\tilde{X}_J S^{-T}. \tag{3.5}$$
   Otherwise, $X^{adi}_J = \tilde{X}_J$.

OUTPUT: $X^{adi}_J \in \mathbb{R}^{n\times n}$, $X^{adi}_J \approx X$.

X_J^{adi} real, it is assumed that in the parameter list $\{p_1, p_2, \ldots, p_J\}$, each parameter is either real or comes as a part of a complex conjugate pair.

A general matrix A must be first reduced to tridiagonal form before proceeding with the ADI iteration in (3.3)–(3.4), to avoid the two full matrix-matrix products and two full matrix-matrix solves. However, it is well known that tridiagonalization of a general nonsymmetric matrix can be unstable (see, e.g., [10]).

The complexity of the ADI algorithm is $O(n^3) + O(Jn^2)$, where J is the total number of ADI iterations [23]. The $O(n^3)$ term comes from the tridiagonalization of a general matrix A, and the transformation in (3.5) to obtain the final ADI approximation. If A is already sparse or structured, there is no need to reduce A to tridiagonal form. In either case, the $O(Jn^2)$ term comes from J iterations of (3.3)–(3.4). In terms of complexity, the ADI method is competitive with the Bartels–Stewart and Hammarling methods, which are also $O(n^3)$ methods. However, the need in the ADI algorithm for the tridiagonalization of a general matrix A can pose a potentially serious problem.

If A is diagonalizable, then the ADI approximation X_J^{adi} has the following error bound [39]:

$$\|X_J^{adi} - X\|_F \leq \|T\|_2^2 \|T^{-1}\|_2^2 k(\mathbf{p})^2 \|X_0^{adi} - X\|_F,$$

$$k(\mathbf{p}) = \max_{x \in spec(A)} \left| \prod_{j=1}^{J} \frac{(p_j - x)}{(p_j + x)} \right|, \tag{3.6}$$

where T is a matrix whose columns are eigenvectors of A and $\mathbf{p} = \{p_1, p_2, \ldots, p_J\}$ are the ADI parameters.

3.1.1. ADI parameter selection. The selection of good parameters is vitally important to the successful application of the ADI algorithm. Optimal ADI parameters $\{p_1, p_2, \ldots, p_J\}$ are a function of J and solve the following rational minimax problem [40]:

$$\min_{p_1, p_2, \ldots, p_J} \max_{x \in \mathcal{R}} \left| \prod_{j=1}^{J} \frac{(p_j - x)}{(p_j + x)} \right|, \tag{3.7}$$

where $\mathcal{R}$ is a region in the open left half plane, and

$$\lambda_1(A), \ldots, \lambda_n(A) \in \mathcal{R} \subset \mathbb{C}^-.$$

If the eigenvalues of A are strictly real, then the solution to (3.7) is known (see [40]). The solution to (3.7) is not known when $\mathcal{R}$ is an arbitrary region in the open left half plane. The problem of finding optimal and near-optimal parameters was investigated in several papers [8, 15, 34, 35, 37, 40].

Here we summarize a parameter selection procedure given in [40]. Define the spectral bounds a, b, and α for the matrix A as

$$a = \min_i(\mathrm{Re}\{\lambda_i\}), \quad b = \max_i(\mathrm{Re}\{\lambda_i\}), \quad \alpha = \tan^{-1} \max_i \left| \frac{\mathrm{Im}\{\lambda_i\}}{\mathrm{Re}\{\lambda_i\}} \right|, \tag{3.8}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $-A$. It is assumed that the spectrum of $-A$ lies entirely inside the "elliptic function domain" determined by a, b, α, as defined

in [40]. If this assumption does not hold, one should try to apply a more general parameter selection algorithm. Let

$$\cos^2\beta = \frac{2}{1+\frac{1}{2}(\frac{a}{b}+\frac{b}{a})},$$

$$m = \frac{2\cos^2\alpha}{\cos^2\beta} - 1.$$

If $m < 1$, the parameters are complex and are given in [8, 40]. If $m \geq 1$, the parameters are real, and we define

$$k' = \frac{1}{m+\sqrt{m^2-1}}, \quad k = \sqrt{1-k'^2}.$$

Note that $k' = \frac{a}{b}$ if all the eigenvalues of A are real. Define the elliptic integrals K and v as

$$F[\psi,k] = \int_0^\psi \frac{dx}{\sqrt{1-k^2\sin^2 x}},$$

$$K = K(k) = F\left[\frac{\pi}{2},k\right], \qquad v = F\left[\sin^{-1}\sqrt{\frac{a}{bk'}},k'\right].$$

The number of ADI iterations required to achieve $k(\mathbf{p})^2 \leq \epsilon_1$ is $J = \lceil \frac{K}{2v\pi}\log\frac{4}{\epsilon_1}\rceil$, and the ADI parameters are given by

$$p_j = -\sqrt{\frac{ab}{k'}}\,dn\left[\frac{(2j-1)K}{2J},k\right], \quad j = 1,2,\ldots,J, \tag{3.9}$$

where $dn(u,k)$ is the elliptic function. It was noted in [23] that for many practical problems ADI converges in a few iterations with these parameters.

3.2. Low rank methods. In [14, 16], low rank approximations to X were proposed which have the form

$$X \approx X_J^{lr} := U_J X_{J\times J} U_J^T, \tag{3.10}$$

where the columns of $U_J \in \mathbb{R}^{n\times r_J}, r_J \leq Jr_b$, form an orthonormal basis for the block Krylov subspace

$$\mathcal{K}_J(A,B) := span\{B, AB, A^2B, \ldots, A^{J-1}B\}.$$

The columns of U_J, as well as the quantities $B_J := (U_J)^T B$ and $A_{J\times J} := U_J^T A U_J$, are obtained via the block Arnoldi process [7, 42].

If $\lambda_i(A_{J\times J}) + \bar{\lambda}_j(A_{J\times J}) \neq 0\ \forall\, i,j$, ensuring that a unique solution to (3.11) exists, then the residual of (1.1),

$$R_J(X_{J\times J}) := A(U_J X_{J\times J} U_J^T) + (U_J X_{J\times J} U_J^T)A^T + BB^T,$$

satisfies the Galerkin condition

$$U_J^T R_J(X_{J\times J}) U_J = 0$$

if and only if $X_{J\times J}$ satisfies

$$A_{J\times J}X_{J\times J} + X_{J\times J}A_{J\times J}^T + B_J B_J^T = 0 \tag{3.11}$$

[14, 16]. The more complicated linear matrix equation that $X_{J\times J}$ must satisfy in order to minimize the Frobenius norm of $R_J(X_{J\times J})$ was also given in [16].

4. CF–ADI. A major contribution of this paper is the development of the CF–ADI algorithm, which is presented in this section. For the low rank right-hand side Lyapunov equation (1.1), CF–ADI produces the same approximation as the ADI method described in section 3 but is much more efficient because it iterates on the Cholesky factor of the ADI approximation rather than on the ADI approximation itself.

For simplicity, all quantities in Algorithm 1 with tildes will be written in this section without the tildes. It is assumed that B has full column rank. Otherwise, we replace B with $\tilde{B}$, where $\tilde{B}$ has full column rank, and $\tilde{B}\tilde{B}^T = BB^T$.

There are two matrix-matrix products and two matrix-matrix solves in (3.3)–(3.4) of Algorithm 1. The need for matrix-matrix operations rather than simply matrix-vector operations at each ADI step makes Algorithm 1 extremely expensive. The first step in developing CF–ADI is to combine (3.3) and (3.4) to obtain

$$\begin{aligned} X_j = &-2p_j(A+p_jI)^{-1}BB^T(A+p_jI)^{-T} \\ &+ (A+p_jI)^{-1}(A-p_jI)X_{j-1}(A-p_jI)^T(A+p_jI)^{-T}. \end{aligned} \tag{4.1}$$

From (4.1) and the fact that $X_0 = 0_{n\times n}$, it can be seen that X_j is symmetric $\forall\ j \in \mathbb{Z}$, and that $\operatorname{rank}(X_j) \le \operatorname{rank}(X_{j-1}) + \operatorname{rank}(B)$. Since iteration begins with the zero matrix initial guess, $\operatorname{rank}(X_j) \le jr_b$, where r_b is the number of columns in B. Therefore, X_j can be represented as an outer product,

$$X_j = Z_j Z_j^T, \tag{4.2}$$

where Z_j has jr_b columns. The matrix Z_j is a Cholesky factor of $X_j \in \mathbb{R}^{n\times n}$.

Replacing X_j with $Z_j {Z_j}^T$ in (4.1) results in

$$Z_0 = 0_{n\times p}, \tag{4.3}$$

$$\begin{aligned} Z_jZ_j^T = &-2p_j\big\{(A+p_jI)^{-1}B\big\}\big\{(A+p_jI)^{-1}B\big\}^T \\ &+ \big\{(A+p_jI)^{-1}(A-p_jI)Z_{j-1}\big\}\big\{(A+p_jI)^{-1}(A-p_jI)Z_{j-1}\big\}^T. \end{aligned} \tag{4.4}$$

The left-hand side of (4.4) is an outer product, and the right-hand side is the sum of two outer products. Thus, Z_j on the left-hand side of (4.4) can be obtained simply by combining the two factors in the two outer products on the right:

$$Z_j = \left[\sqrt{-2p_j}\big\{(A+p_jI)^{-1}B\big\},\ \big\{(A+p_jI)^{-1}(A-p_jI)Z_{j-1}\big\}\right]. \tag{4.5}$$

Thus, the ADI algorithm can be reformulated in terms of the Cholesky factor Z_j of X_j. There is no need to calculate or store X_j at each iteration—only Z_j is needed.

The preliminary form of CF–ADI which iterates on the Cholesky factor Z_j of X_j is

$$Z_1 = \sqrt{-2p_1}(A+p_1I)^{-1}B, \quad Z_1 \in R^{n\times r_b}, \tag{4.6}$$

$$Z_j = \left[\sqrt{-2p_j}(A+p_jI)^{-1}B, (A+p_jI)^{-1}(A-p_jI)Z_{j-1}\right], \quad Z_j \in R^{n\times jr_b}. \tag{4.7}$$

In this formulation, at each iteration, the previous Cholesky factor $Z_{j-1} \in \mathbb{R}^{n\times(j-1)r_b}$ needs to be modified by multiplication on the left by $(A+p_jI)^{-1}(A-p_jI)$. Thus, the number of columns which need to be modified at each iteration increases by r_b. The implementation in (4.6)–(4.7) was independently developed in [27].

In this paper, a further step is taken to keep constant the number of columns modified at each iteration.

The Jr_b columns of Z_J, the Cholesky factor of the Jth ADI approximation, can be written out explicitly:

$$Z_J = \left[S_J\sqrt{-2p_J}B, \quad S_J\left(T_JS_{J-1}\right)\sqrt{-2p_{J-1}}B, \ldots, S_JT_J\cdots S_2\left(T_2S_1\right)\sqrt{-2p_1}B \right],$$

where

$$S_i = (A + p_iI)^{-1}, \quad T_i = (A - p_iI). \tag{4.8}$$

Note that the S_i's and the T_i's commute:

$$S_iS_j = S_jS_i, \quad T_iT_j = T_jT_i, \quad S_iT_j = T_jS_i \quad \forall i, j.$$

The Cholesky factor Z_J then becomes

$$Z_J = [z_J, \quad P_{J-1}(z_J), \quad P_{J-2}(P_{J-1}z_J), \quad \ldots, P_1(P_2\cdots P_{J-1}z_J)], \tag{4.9}$$

where

$$z_J := \left(\sqrt{-2p_J}\right) S_JB = \sqrt{-2p_J}(A + p_JI)^{-1}B, \tag{4.10}$$

$$\begin{aligned} P_l := \left(\frac{\sqrt{-2p_l}}{\sqrt{-2p_{l+1}}}\right) S_lT_{l+1} &= \frac{\sqrt{-2p_l}}{\sqrt{-2p_{l+1}}}(A + p_lI)^{-1}(A - p_{l+1}I) \\ &= \left(\frac{\sqrt{-2p_l}}{\sqrt{-2p_{l+1}}}\right)[I - (p_{l+1} + p_l)\,(A + p_lI)^{-1}]. \end{aligned} \tag{4.11}$$

Since there is no significance to the order in which the ADI parameters appear, the index $1, \ldots, J$ in (4.9) can be reversed. The CF–ADI algorithm which comprises (4.9)–(4.11) with the index reversed is given as Algorithm 2.

Algorithm 2. The CF–ADI algorithm.

INPUT: A, B.

1. Choose CF–ADI parameters, $\{p_1, \ldots, p_{J_{max}}\}$, $\mathrm{Re}\{p_i\} < 0$, (real or complex conjugate pairs).
2. Define: $P_i = \left(\frac{\sqrt{-2p_{i+1}}}{\sqrt{-2p_i}}\right)[I - (p_{i+1} + p_i)(A + p_{i+1}I)^{-1}]$.

   a. $z_1 = \left(\sqrt{-2p_1}\right)(A + p_1I)^{-1}B,$ (4.12)

   b. $Z_1^{cfadi} = \left[\ z_1\ \right].$ (4.13)

3. FOR $j = 2, 3, \ldots, J_{max}$

   a. $z_j = P_{j-1}z_{j-1},$ (4.14)

   b. If ($\|z_j\|_2 > tol_1$ or $\frac{\|z_j\|_2}{\|Z_{j-1}\|_2} > tol_2$) and ($j \le J_{max}$)

   $$Z_j^{cfadi} = \left[\ Z_{j-1}^{cfadi} \quad z_j\ \right].$$ (4.15)

   Otherwise, $J = j - 1$, stop.

   END

OUTPUT: $Z_J^{cfadi} \in \mathbb{C}^{n\times Jr_b}$, $X \approx X_J^{cfadi} := Z_J^{cfadi}(Z_J^{cfadi})^T \in \mathbb{R}^{n\times n}$.

We now show that CF–ADI produces the same approximation as the ADI method.

THEOREM 4.1. *If X_J^{adi} is obtained by running J steps of Algorithm 1 with the ADI parameters $\{p_1, p_2, \ldots, p_J\}$ and Z_J^{cfadi} is obtained by running J steps of Algorithm 2 with the same parameters in any order, then*

$$X_J^{adi} = Z_J^{cfadi}(Z_J^{cfadi})^T. \tag{4.16}$$

Proof. From the derivation of CF–ADI, it is clear that (4.16) is true when the order of the parameters is reversed. The fact that parameter order does not matter in either algorithm is shown by

$$X_j = (A + p_jI)^{-1}(A + p_{j-1}I)^{-1}\Big((A - p_jI)(A - p_{j-1}I)X_{j-2}(A - p_jI)^T(A - p_{j-1}I)^T$$
$$- 2(p_j + p_{j-1})(ABB^TA^T + p_jp_{j-1}BB^T)\Big)(A + p_jI)^{-T}(A + p_{j-1}I)^{-T}.$$

Clearly, this expression does not depend on the order of p_j and p_{j-1}. Any ordering of $\{p_1, \ldots, p_J\}$ can be obtained by exchanging neighboring parameters. ∎

As a matter of notation, define

$$X_J^{cfadi} := Z_J^{cfadi}(Z_J^{cfadi})^T. \tag{4.17}$$

Both X_J^{cfadi} and Z_J^{cfadi} will be referred to as the Jth CF–ADI approximation—which one is meant will be clear from context. The full matrix X_J^{cfadi} is usually not explicitly calculated. It will be used in subsequent sections for analysis purposes only.

4.1. Stopping criteria and parameter selection. The stopping criterion $\|X_j^{cfadi} - X_{j-1}^{cfadi}\|_2 \leq tol^2$ can be implemented as $\|z_j\|_2 \leq tol$, since

$$\|Z_jZ_j^T - Z_{j-1}Z_{j-1}^T\|_2 = \|z_jz_j^T\|_2 = \|z_j\|_2^2.$$

It is not necessarily true that a small z_j implies that all further z_{j+k} will be small, but this has been observed in practice. Relative error can also be used, in which case the stopping criterion is $\frac{\|z_j\|_2}{\|Z_{j-1}\|_2} \leq tol$. The 2-norm of Z_{j-1}, which is also its largest singular value, can be estimated by performing power iterations to estimate the largest eigenvalue of $Z_{j-1}Z_{j-1}^T$, taking advantage of the fact that $j \ll n$. This cost is still high, and this estimate should be used only after each segment of several iterations.

The criterion for picking CF–ADI parameters, $\{p_1, \ldots, p_{J_{max}}\}$, is exactly the same as for ADI parameters, namely, they should solve the rational minimax problem (3.7). Section 3.1.1 gives a parameter selection procedure based on three spectral bounds of A in (3.8). These three bounds for A may be estimated using the power and inverse power iterations, or Gershgorin's circles (see [10]). Power and inverse power iterations can be done at the cost of a few matrix-vector products and solves. A numerical comparison of different choices of parameters is given in section 8.1.

4.2. Complexity. The following definition is helpful when B has more than one column.

DEFINITION 4.2. *An r_b-vector $v \in \mathbb{R}^{n \times r_b}$ is a matrix that has r_b columns.*

The final CF–ADI approximation Z_J^{cfadi} can be obtained from the starting r_b-vector z_1 after $J - 1$ products of the form P_iz_i. The cost of applying P_i to a vector

TABLE 4.1
ADI and CF–ADI complexity comparison when A is sparse.

	CF–ADI	ADI
Sparse A	$O(Jr_bn)$	$O(Jn^2)$
Full A	$O(Jr_bn^2)$	$O(n^3)+O(Jn^2)$

is that of a matrix-vector solve. The starting r_b-vector z_1 is obtained after r_b matrix-vector solves with the columns of $B \in \mathbb{R}^{n\times r_b}$ as the right-hand sides. Each succeeding r_b-vector in Z_J^{cfadi} is obtained from the previous r_b-vector at the cost of r_b matrix-vector solves. Thus, the work per iteration has been reduced from two matrix-matrix products and two matrix-matrix solves in (3.2)–(3.3) in the original ADI method to r_b matrix-vector solves in (4.14) in the CF–ADI algorithm.

The Cholesky factor of the Lyapunov solution is precisely what is needed in the model reduction of linear, time-invariant systems [26, 32, 36]. In general, if Z_J^{cfadi} is available, it is not necessary to calculate $X_J^{cfadi} = Z_J^{cfadi}(Z_J^{cfadi})^T$, whereas if the ADI approximation X_J^{adi} is available, it is often necessary to calculate its Cholesky factor in the subsequent model reduction procedure.

If the matrix A is sparse enough so that $v \mapsto Av$ as well as $v \mapsto (A+p_iI)^{-1}v$ have $O(n)$ complexity, where v is a vector, then Table 4.1 gives the complexity comparison between ADI and CF–ADI. Since r_b, the number of inputs, is by assumption much smaller than n, CF–ADI always results in substantial savings when A is sparse, reducing the work from $O(n^2)$ to $O(n)$.

4.3. Real CF–ADI for complex parameters. Algorithm 2 will result in a complex Cholesky factor $Z_J \in \mathbb{C}^{n\times Jr_b}$ if there are complex ADI parameters, although $Z_JZ_J^T \in \mathbb{R}^{n\times n}$ is guaranteed to be real if the parameters come in complex conjugate pairs.

A version of CF–ADI which uses only operations with real numbers can be implemented by noting that analogous to the matrices associated with a real parameter p_i, given in (4.8), the matrices associated with a complex conjugate pair $\{p_i, \bar{p_i}\}$ are

$$Q_i := (A^2 - \sigma_iA + \tau_iI)^{-1}, \quad R_i := (A^2 + \sigma_iA + \tau_iI), \tag{4.18}$$

$$\sigma_i = 2\mathrm{Re}\{-p_i\}, \quad \tau_i = |p_i|^2, \tag{4.19}$$

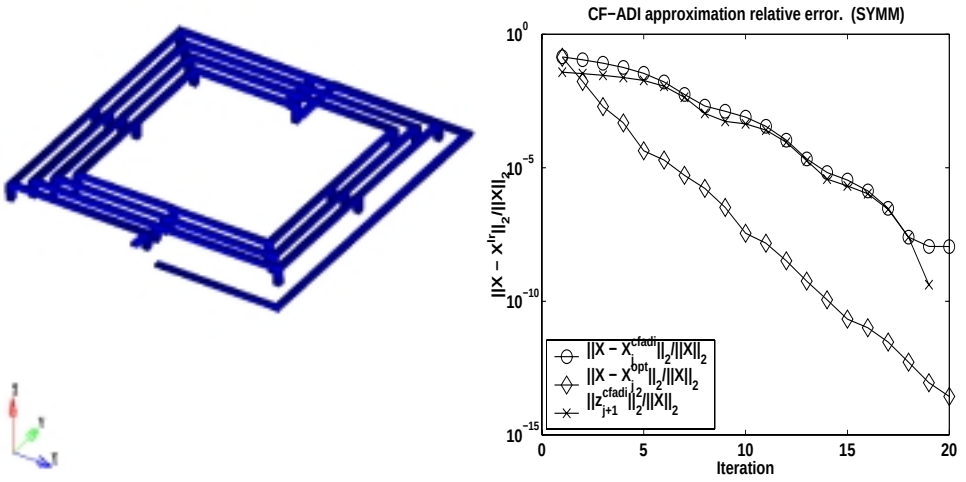
which involve only real quantities.

4.4. Numerical results. This section gives numerical results on the CF–ADI approximation to the solution to (1.1).

The example in Figure 4.1(b) comes from the inductance extraction of an on-chip planar square spiral inductor suspended over a copper plane [17], shown in Figure 4.1(a). The original order 500 system has been symmetrized according to [25]. The matrix A is a symmetric 500×500 matrix, and the input coefficient matrix $B \in \mathbb{R}^n$ has one column.

Because A is symmetric, the eigenvalues of A are real, and good CF–ADI parameters are easy to find. The procedure given in section 3.1.1 was followed. CF–ADI was run to convergence in this example, which took 20 iterations.

Figure 4.1(b) shows the relative 2-norm error of the CF–ADI approximation, $\frac{\|X-X_j^{cfadi}\|_2}{\|X\|_2}$, for $j = 1,\ldots,20$. At $j = 20$, relative error has reached 10^{-8}, which is about the same size as the error of the optimal [10] rank 11 approximation. The error estimate $\|z_{j+1}^{cfadi}\|_2^2$ approximates the actual error $\|X - X_j^{cfadi}\|$ closely $\forall\ j$.



(a) Spiral inductor (b) CF–ADI approximation

FIG. 4.1. *Spiral inductor, a symmetric system.*

5. Krylov and rational Krylov subspace results. This section contains a collection of definitions and results concerning Krylov and rational Krylov subspaces which will be used in subsequent sections.

We begin by giving definitions of Krylov and rational Krylov subspaces.

DEFINITION 5.1. *An* order m *Krylov subspace* $\mathcal{K}_{\mathbf{m}}(\mathbf{A}, \mathbf{z_1})$, $A \in \mathbb{R}^{n\times n}, z_1 \in \mathbb{R}^n$, *is the subspace*

$$\mathcal{K}_m(A, z_1) := span\left\{z_1,\ Az_1,\ A^2z_1, \ldots,\ A^{m-1}z_1\right\}. \tag{5.1}$$

DEFINITION 5.2. *An* order m *rational Krylov subspace* $\mathcal{K}^{\mathbf{rat}}_{\mathbf{m}}(\mathbf{A}, \mathbf{z_1}, \mathbf{p_{m-1}})$, $A \in \mathbb{R}^{n\times n}, z_1 \in \mathbb{R}^n$, $\mathbf{p_{m-1}} = \{p_1, \ldots, p_{m-1}\}$, $p_i \in \mathbb{R}$, *is the subspace*

(5.2)

$$\mathcal{K}_m^{rat}(A, z_1, \mathbf{p_{m-1}})$$
$$:= span\left\{z_1,\ (A+p_1I)^{-1}z_1,\ (A+p_2I)^{-1}(A+p_1I)^{-1}z_1, \ldots,\ \prod_{i=1}^{m-1}(A+p_iI)^{-1}z_1\right\}.$$

Note that for both Krylov and rational Krylov subspaces, the dimension of the subspace may be strictly smaller than the order m. The sets $\{z_1, \ldots, A^{m-1}z_1\}$ and $\{z_1, (A+p_1I)^{-1}z_1, \ldots, \prod_{i=1}^{m-1}(A+p_i)^{-1}z_1\}$ are *spanning sets* for $\mathcal{K}_m(A, z_1)$ and $\mathcal{K}_m^{rat}(A, z_1, \mathbf{p_{m-1}})$, respectively.

The following well-known result can be found in many standard textbooks, including [10].

PROPOSITION 5.3. *If* $m > n$, *then* $\mathcal{K}_m(A, B) = \mathcal{K}_n(A, B)$.

Theorem 5.4 characterizes the rational Krylov subspace $\mathcal{K}_m^{rat}(A, (A+p_1I)^{-1}B, \{p_2, \ldots, p_m\})$ as the direct sum of l rational Krylov subspaces, where l is the number of distinct parameters in the list $\{p_1, \ldots, p_m\}$.

THEOREM 5.4. *Let* $\mathcal{K}_m^{rat}(A,(A+p_1I)^{-1}B,\{p_2,\dots,p_m\})$ *be such that no* $(A+p_iI)$ *is singular. Then*

$$\begin{aligned}
&\mathcal{K}_m^{rat}\left(A,(A+p_1I)^{-1}B,\{p_2,\dots,p_m\}\right)\\
&= span\left\{(A+p_1I)^{-1}B,\dots,\prod_{i=1}^{j}(A+p_iI)^{-1}B,\dots,\prod_{i=1}^{m}(A+p_iI)^{-1}B\right\}\\
&= \sum_{i=1}^{l} span\{(A+p_iI)^{-1}B,\dots,(A+p_iI)^{-s_i}B\}\\
&= \sum_{i=1}^{l}\mathcal{K}_{s_i}^{rat}\left((A+p_iI),\ (A+p_iI)^{-1}B,\ \mathbf{0_{s_i-1}}\right),
\end{aligned}$$

where $s_1+\cdots+s_l=m$, *each* p_i *appears in* $\{p_1,\dots,p_m\}$ *a total of* s_i *times, and the summation sign denotes direct sum of subspaces.*

Proof. If the parameters are distinct, the proof follows from the partial fractions expansion

$$\prod_{i=1}^{j}(A+p_iI)^{-1}=\sum_{i=1}^{j}\left(\prod_{k\neq i}\left(\frac{1}{p_k-p_i}\right)\right)(A+p_iI)^{-1},\quad p_1\neq p_2\neq\cdots\neq p_J.$$

A slightly different expansion taking into account repeated parameters can be calculated to give the general statement of the theorem. ∎

6. Spanning sets of $\mathcal{L}(A,B)$. In this section we prove Theorem 6.1, which shows the equivalence of an infinite number of order n Krylov and rational Krylov subspaces based on A and B. For simplicity we assume B has only one column. Most of the results in this section can be easily generalized to the case when B has more than one column.

THEOREM 6.1. *Let* $A\in\mathbb{R}^{n\times n}$ *be invertible,* $B\in\mathbb{R}^n$, $B\neq 0$, $\mathbf{p}=\{\dots,p_{-2},p_{-1},p_0,p_1,p_2,\dots\}$, $p_i\in\mathbb{R}$, *and define the subspace* $\mathcal{L}(A,B,\mathbf{p})$ *as*

$$\begin{aligned}
&\mathcal{L}(A,B,\mathbf{p})\\
&:= span\Big\{\dots,\ \prod_{i=-j}^{-1}(A+p_iI)^{-1}B,\dots,\ (A+p_{-2}I)^{-1}(A+p_{-1}I)^{-1}B,\\
&\qquad (A+p_{-1}I)^{-1}B,\ B,\ (A+p_0I)B,\\
&\qquad (A+p_1I)(A+p_0I)B,\dots,\ \prod_{i=0}^{j-1}(A+p_iI)B,\dots\Big\}\\
&= span\big\{\dots,v_{-j}(A,B,\mathbf{p}),\dots,v_{-2}(A,B,\mathbf{p}),\ v_{-1}(A,B,\mathbf{p}),\ v_0(A,B,\mathbf{p}),\\
&\qquad v_1(A,B,\mathbf{p}),\ v_2(A,B,\mathbf{p}),\dots,v_j(A,B,\mathbf{p}),\dots\big\},
\end{aligned}\tag{6.1}$$

where

$$v_j(A,B,\mathbf{p})=\begin{cases}B, & j=0,\\ \prod_{i=0}^{j-1}(A+p_iI)B, & j>0,\\ \prod_{i=j}^{-1}(A+p_iI)^{-1}B, & j<0,\end{cases}\tag{6.2}$$

and where all matrix inverses in (6.1) *are well defined. Then* $\forall s \in \mathbb{Z}$, $\forall \mathbf{p}$, $\forall \mathbf{r} = \{\dots, r_{-1}, r_0, r_1, \dots\}$, $\forall \mathbf{q} = \{\dots, q_{-1}, q_0, q_1, \dots\}, r_i, q_i \in \mathbb{R}$,

$$\mathcal{L}(A, B, \mathbf{p}) = span\{v_s(A, B, \mathbf{p}),\ v_{s+1}(A, B, \mathbf{p}), \dots, v_{s+(n-1)}(A, B, \mathbf{p})\} \tag{6.3}$$

$$= span\{B, AB, \dots,\ A^{n-1}B\} \tag{6.4}$$

$$= \mathcal{L}(A, v_s(A, B, \mathbf{r}), \mathbf{q}) \tag{6.5}$$

if all matrix inverses in (6.5) *are well defined.*

REMARK 1. *We refer to* B *in* $\mathcal{L}(A, B, \mathbf{p})$ *as the base vector. Because of* (6.5), $\mathcal{L}(A, B) := \mathcal{L}(A, v_s(A, B, \mathbf{r}), \mathbf{q})$ *may be written without referring to the base vector* $v_s(A, B, \mathbf{r})$ *or the shifts* $\mathbf{q}$.

The proof of Theorem 6.1 needs the following lemmas. The dependence of the v_i's on $A, B, \mathbf{p}$ will be suppressed in the proofs unless needed.

LEMMA 6.2. *Let the* v_j*'s be defined as in* (6.2). *Then*

$$v_l \in span\{v_s,\ v_{s+1},\ v_{s+2}, \dots, v_{s+(n-1)}\} \tag{6.6}$$

whenever $l > s + (n-1)$.

Proof. From (6.2), it can be seen that $v_j = (A + p_{j-1}I)v_{j-1}\ \forall j$; hence,

$$span\{v_{j-1}, v_j\} = span\{v_{j-1}, Av_{j-1}\}$$

and

$$span\{v_s,\ v_{s+1},\ v_{s+2}, \dots, v_l\} = span\{v_s, Av_s, \dots, A^{l-s}v_s\} = \mathcal{K}_{l-s+1}(A, v_s). \tag{6.7}$$

From Proposition 5.3,

$$\begin{aligned} span\{v_s,\ v_{s+1},\ v_{s+2}, \dots, v_l\} &= \mathcal{K}_{l-s+1}(A, v_s) \\ &= \mathcal{K}_n(A, v_s) = span\{v_s,\ v_{s+1},\ v_{s+2}, \dots, v_{s+(n-1)}\}. \end{aligned}$$

The result follows. ∎

LEMMA 6.3. *Let the* v_j*'s be defined as in* (6.2)*; then*

$$v_l \in span\{v_s, v_{s+1}, v_{s+2}, \dots, v_{s+(n-1)}\} \tag{6.8}$$

whenever $l < s$.

Proof. First we show that the lemma is true for $l = s - 1$. Equivalently, because of (6.7), show that

$$(A + p_{s-1}I)^{-1}v_s \in span\{v_s, Av_s, \dots, A^{n-1}v_s\}. \tag{6.9}$$

Shifts can be added in the right-hand side of (6.9),

$$span\{v_s, Av_s, \dots, A^{n-1}v_s\} = span\{v_s, (A + p_{s-1}I)v_s, \dots, (A + p_{s-1}I)^{n-1}v_s\},$$

without affecting its column span. Because $\{v_{s-1}, v_s, \dots, v_{s+(n-1)}\}$ are $n+1$ vectors in $\mathbb{R}^n$, there exist coefficients, $c_0, \dots, c_n$, not all zero, such that

$$c_0 v_s + c_1(A + p_{s-1}I)v_s + \cdots + c_{n-1}(A + p_{s-1}I)^{n-1}v_s + c_n(A + p_{s-1}I)^{-1}v_s = 0. \tag{6.10}$$

If $c_n \neq 0$, (6.9) is proven. Otherwise, since $B \neq 0$, we can choose $0 \leq j < n-1$ such that $c_j \neq 0$ and $c_i = 0\ \forall i < j$. Then multiply (6.10) by $(A + p_{s-1}I)^{-(j+1)}$ to obtain

$$\begin{aligned} &c_j(A + p_{s-1}I)^{-1}v_s + c_{j+1}v_s + \cdots + c_{n-1}(A + p_{s-1}I)^{n-2-j}v_s = 0 \\ &\implies c_j(A + p_{s-1}I)^{-1}v_s = -c_{j+1}v_s - \cdots - c_{n-1}(A + p_{s-1}I)^{n-2-j}v_s. \end{aligned}$$

Thus, (6.9) is proven, and (6.8) holds for $l = s-1$. If $l < s-1$,

$$v_l \in \mathit{span}\{v_{l+1}, v_{l+2}, \ldots, v_{l+n}\} \tag{6.11}$$

$$\subseteq \mathit{span}\{v_{l+2}, \ldots, v_{l+n+1}\} \tag{6.12}$$

$$\vdots \tag{6.13}$$

$$\subseteq \mathit{span}\{v_s, \ldots, v_{s+n-1}\}. \tag{6.14}$$

Relation (6.12) follows because each vector $v_{l+1}, \ldots, v_{l+n}$ is in $\mathit{span}\{v_{l+2}, \ldots, v_{l+n+1}\}$. ∎

Proof of Theorem 6.1. Lemmas 6.2 and 6.3 show that for any $\mathbf{p}$,

$$\mathcal{L}(A, B, \mathbf{p}) = \mathit{span}\{v_s(A, B, \mathbf{p}),\ v_{s+1}(A, B, \mathbf{p}), \ldots, v_{s+(n-1)}(A, B, \mathbf{p})\}$$

holds for any s. Equation (6.4) follows from the fact that for any $\mathbf{p}$, with the choice of $s = 0$,

$$\mathit{span}\{v_0(A, B, \mathbf{p}), v_1(A, B, \mathbf{p}), \ldots, v_{n-1}(A, B, \mathbf{p})\} = \mathit{span}\{B, AB, \ldots, A^{n-1}B\}.$$

Equation (6.5) follows from

$$\begin{aligned} \mathcal{L}(A, B, \mathbf{p}) &= \mathit{span}\{B, AB, \ldots, A^{n-1}B\} = \mathcal{L}(A, B, \mathbf{r}) \\ &= \mathit{span}\{v_s(A, B, \mathbf{r}), v_{s+1}(A, B, \mathbf{r}), \ldots, v_{s+(n-1)}(A, B, \mathbf{r})\} \\ &= \mathit{span}\{v_s(A, B, \mathbf{r}), Av_s(A, B, \mathbf{r}), \ldots, A^{n-1}v_s(A, B, \mathbf{r})\} \\ &= \mathcal{L}(A, v_s(A, B, \mathbf{r}), \mathbf{q}) \quad \forall \mathbf{p}, \quad \forall \mathbf{r}, \quad \forall \mathbf{q}. \end{aligned}$$ ∎

Remark 2. *Special cases of Theorem* 6.1 *can be found in many references, including* [11, 31].

7. Lyapunov solution and rational Krylov subspaces. In this section we characterize the range of the Lyapunov solution as order n Krylov and rational Krylov subspaces with different starting vectors and different sets of shifts. We also state several properties of the CF–ADI approximation.

Proposition 7.1 is a well-known result which makes the connection between the range of the Lyapunov solution X and the Krylov subspace $\mathcal{K}_n(A, B)$ (see [6, 33]).

Proposition 7.1. *Let X be the solution to* (1.1). *Then*

$$\text{range}(X) = \mathit{span}\{B, AB, \ldots, A^{n-1}B\} = \mathcal{K}_n(A, B). \tag{7.1}$$

The following corollary of Theorem 6.1 gives a more complete characterization of the range of X as Krylov and rational Krylov subspaces.

Corollary 7.2. *With the same notation as in Theorem* 6.1,

$$\text{range}(X) = \mathcal{L}(A, v_t(A, B, \mathbf{r}), \mathbf{q}) \quad \forall t \in \mathbb{Z}, \quad \forall \mathbf{r}, \quad \forall \mathbf{q}. \tag{7.2}$$

Theorem 6.1 and Corollary 7.2 together imply that any n consecutive vectors $\{w_s, \ldots, w_{s+n-1}\}$, $s \in \mathbb{Z}$, in the infinite spanning set for $\mathcal{L}(A, v_t(A, B, \mathbf{r}), \mathbf{q})$ are a spanning set for range(X).

We now state some properties of the CF–ADI approximation and omit the proofs.

PROPOSITION 7.3. *Let Z_j^{cfadi} be the jth CF–ADI approximation. Then its column span has the following characterization:*

$$colsp(Z_j^{cfadi}) = \mathcal{K}_j^{rat}(A,\ (A + p_1 I)^{-1} B,\ \{p_2, \ldots, p_j\}). \tag{7.3}$$

PROPOSITION 7.4. *Let $Z_j^{cfadi} = [z_1, \ldots, z_j]$ be the jth CF–ADI approximation, and let $B \in \mathbb{R}^n$. If z_{j+1} is a linear combination of $\{z_1, \ldots, z_j\}$, then z_l is a linear combination of $\{z_1, \ldots, z_j\}$ whenever $l \geq j + 1$.*

PROPOSITION 7.5. *Let $Z_n^{cfadi} = [z_1, \ldots, z_n]$ be the nth CF–ADI approximation. Then*

$$\begin{aligned} colsp(Z_n^{cfadi}) &= \mathcal{K}_n^{rat}(A, (A - p_1 I)^{-1} B, \{p_2, \ldots, p_n\}) \\ &= \text{range}(X). \end{aligned}$$

REMARK 3. *Proposition* 7.5 *states that if CF–ADI is run n steps, the range of X emerges.*

PROPOSITION 7.6. *If z_{j+1} at the $(j+1)$st step of the CF–ADI iteration is a linear combination of the previous iterates, $z_1, \ldots, z_j$, and $B \in \mathbb{R}^n$, then*

$$span\{z_1, \ldots, z_j\} = \text{range}(X).$$

REMARK 4. *If the goal is to find the range of the exact solution X, then iteration can stop when z_{j+1} is a linear combination of the previous columns. If, however, the goal is to approximate X by $Z_j^{cfadi}(Z_j^{cfadi})^T$, then iteration may have to continue, since even if $Z_j^{cfadi}(Z_j^{cfadi})^T$ has the same range as X, they may not be close as matrices.*

8. Rational Krylov subspace approximation to dominant invariant subspace. In this section we are concerned with the approximation of the dominant invariant subspace of the Lyapunov solution. In particular, we make the connection between approximating the dominant invariant subspace of the solution X and the generation of various low order Krylov and rational Krylov subspaces. As described in section 2, for some important applications it is sufficient to find the dominant invariant subspace of X. The complete knowledge of X is not necessary.

Corollary 7.2 in section 7 shows that range$(X) = \mathcal{L}(A, v_t(A, B, \mathbf{r}), \mathbf{q})\ \forall t,\ \forall \mathbf{r},\ \forall \mathbf{q}$. The range of X can also be characterized in terms of its eigenvectors. Let

$$X = [u_1, \ldots, u_n] \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{bmatrix} [u_1, \ldots, u_n]^T$$

be the eigenvalue (singular value) decomposition of X, with the eigenvalues ordered so that

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Then the eigenvectors of X associated with the nonzero eigenvalues, $u_1, \ldots, u_r$, span the range of X,

$$\text{range}(X) = span\,\{u_1, \ldots, u_r\}\,. \tag{8.1}$$

Combining Corollary 7.2 and (8.1) gives spanning sets for the invariant subspace, $span\,\{u_1, \ldots, u_r\}$, of X,

$$span\{u_1, \ldots, u_r\} = span\{w_s, \ldots, w_{s+n-1}\}, \tag{8.2}$$

where w_i, $i = s, \ldots, s+n-1$, are n consecutive vectors in the infinite spanning set for $\mathcal{L}(A, v_t(A, B, \mathbf{r}), \mathbf{q})$.

It is then natural to approximate the J dimensional dominant invariant subspace of X, $span\{u_1, \ldots, u_J\}$, $J \le r \le n$, by $span\{v_1, \ldots, v_J\}$,

$$span\{u_1, \ldots, u_J\} \approx span\{v_1, \ldots, v_J\}, \tag{8.3}$$

where $\{v_1, \ldots, v_J\}$ is a subset of the order n spanning set $\{w_s, \ldots, w_{s+n-1}\}$ for some choice of s, t, $\mathbf{r}$, $\mathbf{q}$. Since only the matrix A and the vector B are given, from practical concerns the subset $\{v_1, \ldots, v_J\}$ should contain consecutive components of $\{w_s, \ldots, w_{s+n-1}\}$. Without loss of generality, we choose

$$\{v_1, \ldots, v_J\} = \{w_s, \ldots, w_{s+J-1}\}. \tag{8.4}$$

The set $\{v_1, \ldots, v_J\}$ may be generated as for an order J Krylov subspace based on the matrix A and the vector v_1 or may be generated in reverse order as for a rational Krylov subspace based on A and v_J.

A basis for any Krylov or rational Krylov subspace choice in (8.4) may be generated stably via the Arnoldi algorithm [7, 42]. The subspace $span\{u_1, \ldots, u_r\}$ will emerge in the same number of Arnoldi steps, which is at most n, for any subspace choice in (8.4). Because it is not practical to run any of these Krylov subspace-based approaches to n Arnoldi steps, we focus on the case when $J \ll n$.

A few examples of the approximation we consider in section 8.1 are $span\{v_1, \ldots, v_J\} =$

$$\text{a. } \mathcal{K}_J(A, B) = span\{B, AB, \ldots, A^{J-1}B\}, \tag{8.5}$$

$$\text{b. } \mathcal{K}_J^{rat}(A,\ A^{-1}B,\ \mathbf{0_{J-1}}) = span\{A^{-1}B, A^{-2}B, \ldots, A^{-J}B\}, \tag{8.6}$$

$$\text{c. } \mathcal{K}_J^{rat}\left(A, (A + p_1 I)^{-1}B, \{p_2, \ldots, p_J\}\right) \quad \text{for any } \{p_1, \ldots, p_J\}. \tag{8.7}$$

The choice in (8.5) was utilized in [14, 16]. If we choose the shifts $\{p_1, \ldots, p_J\}$ to be CF–ADI parameters in (8.7), we obtain the CF–ADI approximation to the dominant invariant subspace of X. Clearly, the shifts in (8.7) may be chosen in other ways. It is also possible to realize the choice in (8.7) as the direct sum of shifted rational Krylov subspaces due to Theorem 5.4.

The answer to the question of which choice in (8.5)–(8.7) best satisfies (8.3) depends on A, B, J, and the shift parameters $\{p_1, \ldots, p_J\}$. However, since there is more freedom in the choice in (8.7) than in (8.5) or (8.6), in general, one expects (8.7) to be a better choice if the shift parameters are chosen well. One answer to how to choose the shifts in (8.7) is to use the CF–ADI parameters, which are the solution of the rational minimax problem (3.7). The justification is that these parameters minimize the norm of the error $\|X - X_J^{cfadi}\|$.

8.1. Numerical results. In this section we give numerical examples of approximating the dominant invariant subspace of X by the Krylov and rational Krylov subspace choices in (8.5)–(8.7), including several natural choices of shifts in (8.7). Some preliminary numerical results on using CF–ADI to approximate the dominant invariant subspace of X can be found in [22], but the subspaces comparisons have not appeared before in literature.

The measure of the closeness of two subspaces is provided by the concept of principal angles between subspaces (see [10]).

DEFINITION 8.1. *Let S^1 and S^2 be two subspaces, of dimension d_1 and d_2, respectively, and assume $d_1 \geq d_2$. Then the d_2* principal angles *between S^1 and S^2 are $\theta_1, \ldots, \theta_{d_2}$ such that*

$$\cos(\theta_j) = \max_{u^1 \in S^1, \|u^1\|=1} \max_{u^2 \in S^2, \|u^2\|=1} (u^1)^T u^2 = (u_j^1)^T u_j^2$$

under the constraints that

$$(u^1)^T u_i^1 = 0, \quad (u^2)^T u_i^2 = 0, \quad i = 1 : j-1.$$

REMARK 5. *If the columns of U^1 are an orthonormal basis for S^1, the columns of U^2 are an orthonormal basis for S^2, and $(U^1)^T U^2$ has singular value decomposition $(U^1)^T U^2 = U \Sigma V^T$, then*

$$\cos(\theta_j) = \Sigma(j,j), \quad u_j^1 = U^1 U(:,j), \quad u_j^2 = U^2 V(:,j).$$

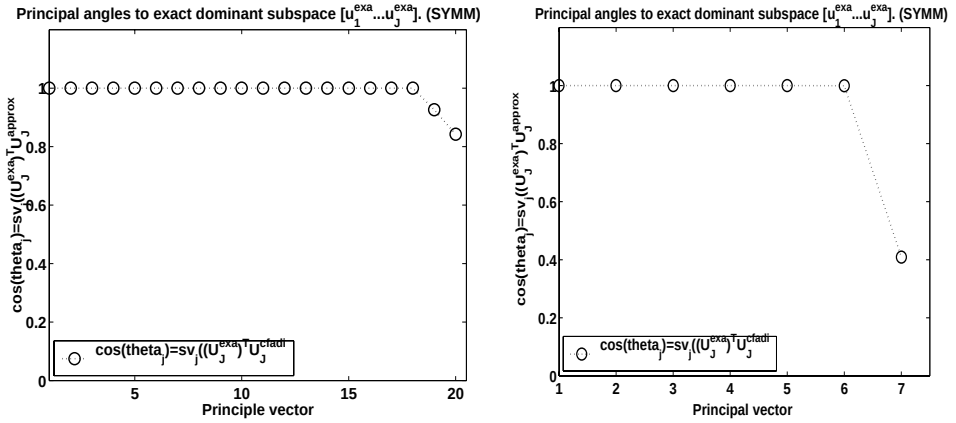
It can be seen that if $S^1 = S^2$, then $\cos(\theta_j) = 1$, $j = 1, \ldots, d_1 = d_2$, and if $S^1 \perp S^2$, then $\cos(\theta_j) = 0$, $j = 1, \ldots, d_2$.

The two bases $\{u_1^1, \ldots, u_{d_2}^1\}$ and $\{u_1^2, \ldots, u_{d_2}^2\}$ are mutually orthogonal, $(u_i^1)^T u_j^2 = 0$, if $i \neq j$. And $(u_i^1)^T u_i^2 = \cos(\theta_i)$ indicates the closeness of u_i^1 and u_i^2. A basis for the intersection of S^1 and S^2 is given by those basis vectors whose principal angle is 0. Thus, the closeness of two subspaces can be measured by how many of their principal angles are close to 0.

The example in Figure 8.1 comes from the spiral inductor problem considered in section 4.4. The matrix A is symmetric, 500×500, and B has one column. CF–ADI was run for 20 iterations and the results are shown in Figure 8.1(a). The relative error after 20 iterations is $\frac{\|X - X_j^{cfadi}\|_2}{\|X\|_2} = 10^{-8}$. The cosines of 18 of the principal angles between the exact invariant subspace and the approximate subspace are 1, and the cosines of the last 2 are above 0.8, indicating close match of all dominant eigenvectors. In contrast, Figure 8.1(b) shows the results after CF–ADI was run for only 7 iterations. The relative error $\frac{\|X - X_7^{cfadi}\|_2}{\|X\|_2}$ is 4.0×10^{-3}. However, it can be seen that the cosines of 6 principal angles are 1. Thus, dominant eigenspace information about X can emerge, even when CF–ADI has not converged.

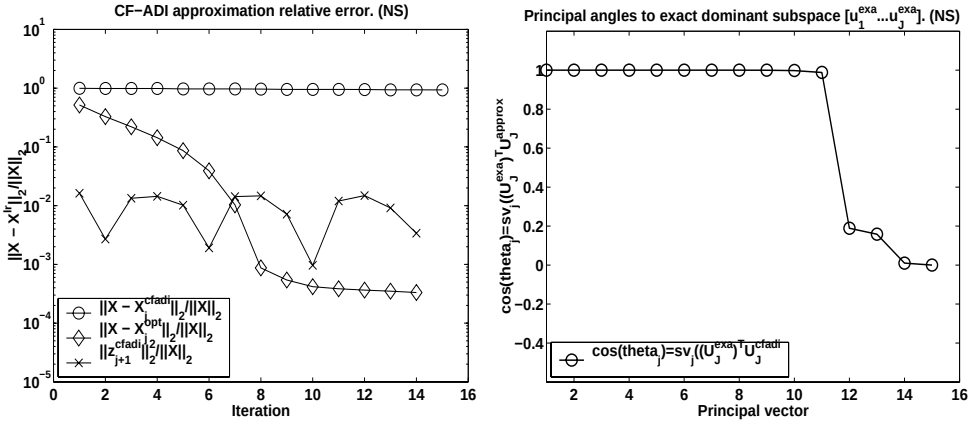
Figure 8.2 shows another example of running CF–ADI only a small number of steps, before convergence occurs. It comes from a discretized transmission line example [24]. The system matrix A is nonsymmetric, 256×256, and the input matrix B has one column. The parameter selection procedure in [40] was followed and the resulting CF–ADI parameters were complex.

Figure 8.2(a) shows that the CF–ADI error was not decreasing at all during the 15 iterations. The relative error stagnates at 1. However, Figure 8.2(b) shows that the intersection of the 15 dimensional exact dominant invariant subspace and the 15 dimensional CF–ADI approximation has dimension 10 (almost 11).



(a) 20 CF–ADI iterations (b) 7 CF–ADI iterations

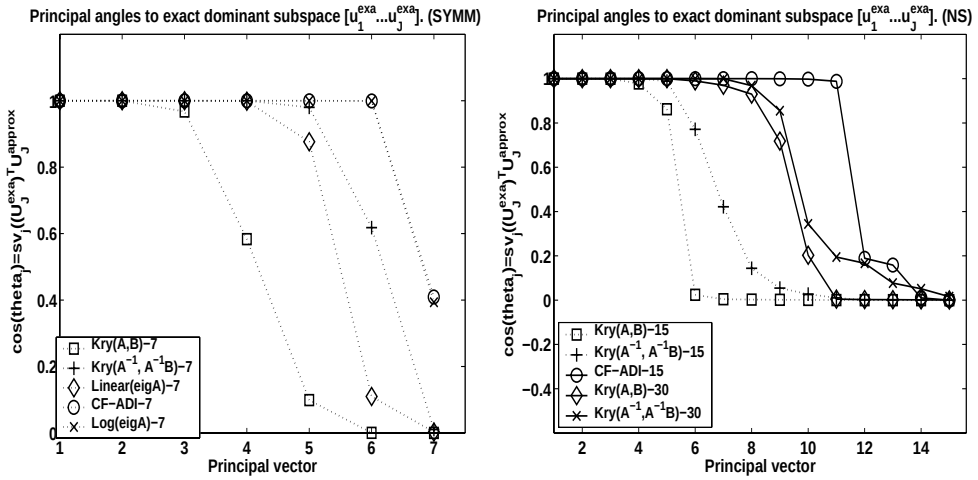
FIG. 8.1. *Symmetric matrix,* $n = 500$. *Principal angles.*



(a) Convergence (b) Principal angle

FIG. 8.2. *Nonsymmetric matrix,* $n = 256$, 15 *CF–ADI iterations, not converged.*

In Figure 8.3 we make comparison of the Krylov and rational Krylov subspace choices in (8.5)–(8.7). In Figure 8.3(a) we compare different rank 7 approximations to the exact dominant invariant subspace for the symmetric spiral inductor example. The shifted rational Krylov subspace is compared with the unshifted Krylov subspace, $\mathcal{K}_J(A, B)$, and the unshifted rational Krylov subspace, $\mathcal{K}_J(A^{-1}, A^{-1}B)$, for $J = 7$. Three choices of shift parameters for the rational Krylov subspace, $\mathcal{K}_J^{rat}(A, (A + p_1 I)^{-1}B, \{p_2, \ldots, p_J\})$, are compared. They are linearly and logarithmicaly spaced points on the eigenvalue interval of A and CF–ADI parameters from the solution of rational minimax problem (3.7). Figure 8.3(a) shows that $\mathcal{K}_7(A, B)$ provides the worst approximation. A better approximation is $\mathcal{K}_7^{rat}(A, (A + p_1 I)^{-1}B, \{p_2, \ldots, p_7\})$, with $\{p_1, \ldots, p_7\}$ linearly spaced points on the eigenvalue interval of A. A better approxi-



(a) Symmetric system

(b) Nonsymmetric system

FIG. 8.3. *Comparison of various low rank approximations to the exact dominant invariant subspace.*

mation than that is the unshifted rational Krylov subspace, $\mathcal{K}_7(A^{-1}, A^{-1}B)$. Finally, for this example, using the CF–ADI parameters and using logarithmically spaced points in $\mathcal{K}_7^{rat}(A, (A+p_1I)^{-1}B, \{p_2, \ldots, p_7\})$ both provide the best approximation.

In Figure 8.3(b) comparison is made for the nonsymmetric transmission line example. Order 15 and 30 unshifted Krylov and rational Krylov subspaces, $\mathcal{K}_J(A, B)$, $\mathcal{K}_J(A^{-1}, A^{-1}B)$, $J = 15, 30$, are compared with the order 15 shifted rational Krylov subspace, $\mathcal{K}_{J_{cfadi}}^{rat}(A, (A+p_1I)^{-1}B, \{p_2, \ldots, p_{J_{cfadi}}\})$, $J_{cfadi} = 15$, where $\{p_1, \ldots, p_{J_{cfadi}}\}$ are an approximate solution to the complex region rational minimax problem (3.7), obtained by the procedure described in [40].

Figure 8.3(b) shows that $\mathcal{K}_{15}(A, B)$ gives the worst approximation, followed by $\mathcal{K}_{15}(A^{-1}, A^{-1}B)$. Finding order 30 unshifted subspaces, $\mathcal{K}_{30}(A, B)$ and $\mathcal{K}_{30}(A^{-1}, A^{-1}B)$, to match the 15 dimensional exact dominant invariant subspace offers improvement. But clearly the order 15 subspace, $\mathcal{K}_{15}^{rat}(A, (A+p_1I)^{-1}B, \{p_2, \ldots, p_{15}\})$, using the CF–ADI parameters, gives the best approximation.

9. Conclusions. In this paper we developed the CF–ADI algorithm to generate a low rank approximation to the solution to the Lyapunov equation. CF–ADI requires only matrix-vector products and linear solves. Hence, it enables one to take advantage of any sparsity or structure in the coefficient matrix. The range of the CF–ADI approximation is a low order shifted rational Krylov subspace, where the shifts are the solution of a rational minimax problem.

We characterized the range of the solution to the Lyapunov equation as order n Krylov and rational Krylov subspaces with various starting vectors and various sets of shifts. A connection is made between the approximation of the dominant invariant subspace of the Lyapunov solution and the generation of low order Krylov and rational Krylov subspaces.

It is shown that the rational Krylov subspace generated by the CF–ADI algorithm

frequently gives the most accurate approximation to the dominant invariant subspace of the exact solution to the Lyapunov equation, which is needed in many engineering applications.

REFERENCES

[1] A. Antoulas, D. Sorensen, and Y. Zhou, *On the Decay Rate of Hankel Singular Values and Related Issues*, Technical report, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 2001.

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

[3] A. N. Beavers, Jr. and E. D. Denman, *A new solution method for the Lyapunov matrix equation*, SIAM J. Appl. Math., 29 (1975), pp. 416–421.

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

[5] G. Birkhoff, R. S. Varga, and D. Young, *Alternating direction implicit methods*, in Advances in Computers, Vol. 3, Academic Press, New York, 1962, pp. 189–273.

[6] P. C. Chandrasekharan, *Robust Control of Linear Dynamical Systems*, Harcourt Brace, London, San Diego, CA, 1996.

[7] I. Elfadel and D. Ling, *A block rational Arnoldi algorithm for multipoint passive model-order reduction of multiport RLC networks*, in Proceedings of the International Conference on Computer-Aided Design, San Jose, CA, 1997, pp. 66–71.

[8] N. S. Ellner and E. L. Wachspress, *Alternating direction implicit iteration for systems with complex spectra*, SIAM J. Numer. Anal., 28 (1991), pp. 859–870.

[9] K. Glover, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds*, Internat. J. Control, 39 (1984), pp. 1115–1193.

[10] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.

[11] E. Grimme, *Krylov Projection Methods for Model Reduction*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1997.

[12] S. J. Hammarling, *Numerical solution of the stable, nonnegative definite Lyapunov equation*, IMA J. Numer. Anal., 2 (1982), pp. 303–323.

[13] A. S. Hodel, B. Tenison, and K. R. Poolla, *Numerical solution of the Lyapunov equation by approximate power iteration*, Linear Algebra Appl., 236 (1996), pp. 205–230.

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

[15] M.-P. Istace and J.-P. Thiran, *On the third and fourth Zolotarev problems in the complex plane*, SIAM J. Numer. Anal., 32 (1995), pp. 249–259.

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

[17] M. Kamon, F. Wang, and J. White, *Recent improvements for fast inductance extraction and simulation [packaging]*, in Proceedings of the IEEE 7th Topical Meeting on Electrical Performance of Electronic Packaging, West Point, NY, 1998, pp. 281–284.

[18] P. Lancaster, *Explicit solutions of linear matrix equations*, SIAM Rev., 12 (1970), pp. 544–566.

[19] V. B. Larin and F. A. Aliev, *Construction of square root factor for solution of the Lyapunov matrix equation*, Systems Control Lett., 20 (1993), pp. 109–112.

[20] J.-R. Li, F. Wang, and J. White, *An efficient Lyapunov equation-based approach for generating reduced-order models of interconnect*, in Proceedings of the 36th Design Automation Conference, New Orleans, LA, 1999, pp. 1–6.

[21] J.-R. Li and J. White, *Efficient model reduction of interconnect via approximate system gramians*, in Proceedings of the IEEE/ACM International Conference on Computer-Aided Design, San Jose, CA, 1999, pp. 380–383.

[22] J.-R. Li and J. White, *Reduction of large circuit models via low rank approximate gramians*, Int. J. Appl. Math. Comput. Sci., 11 (2001), pp. 1151–1171.

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

[24] N. Marques, M. Kamon, J. White, and L. Silveira, *A mixed nodal-mesh formulation for efficient extraction and passive reduced-order modeling of 3D interconnects*, in Proceedings of the 35th ACM/IEEE Design Automation Conference, San Francisco, CA, June 1998, pp. 297–302.

[25] L. Miguel Silveira, M. Kamon, I. Elfadel, and J. White, *A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of RLC circuits*, in Proceedings of the IEEE/ACM International Conference on Computer-Aided Design, San Jose, CA, 1996, pp. 288–294.
[26] B. C. Moore, *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Automat. Control, 26 (1981), pp. 17–32.
[27] T. Penzl, *A cyclic low-rank Smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comput., 21 (2000), pp. 1401–1418.
[28] T. Penzl, *Eigenvalue decay bounds for solutions of Lyapunov equations: The symmetric case*, Systems Control Lett., 40 (2000), pp. 139–144.
[29] L. Pernebo and L. M. Silverman, *Model reduction via balanced state space representations*, IEEE Trans. Automat. Control, 27 (1982), pp. 382–387.
[30] J. D. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control, 32 (1980), pp. 677–687.
[31] A. Ruhe, *The rational Krylov algorithm for nonsymmetric eigenvalue problems.* III. *Complex shifts for real matrices*, BIT, 34 (1994), pp. 165–176.
[32] M. G. Safonov and R. Y. Chiang, *A Schur method for balanced-truncation model reduction*, IEEE Trans. Automat. Control, 34 (1989), pp. 729–733.
[33] E. D. Sontag, *Mathematical Control Theory*, 2nd ed., Springer-Verlag, New York, 1998.
[34] G. Starke, *Optimal alternating direction implicit parameters for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 28 (1991), pp. 1431–1445.
[35] G. Starke, *Fejér-Walsh points for rational functions and their use in the ADI iterative method*, J. Comput. Appl. Math., 46 (1993), pp. 129–141.
[36] M. S. Tombs and I. Postlethwaite, *Truncated balanced realization of a stable nonminimal state-space system*, Internat. J. Control, 46 (1987), pp. 1319–1330.
[37] E. L. Wachspress, *Optimum alternating-direction-implicit iteration parameters for a model problem*, J. Soc. Indust. Appl. Math., 10 (1962), pp. 339–350.
[38] E. L. Wachspress, *Iterative solution of the Lyapunov matrix equation*, Appl. Math. Lett., 1 (1988), pp. 87–90.
[39] E. L. Wachspress, *ADI iterative solution of Lyapunov equations*, in Iterative Methods in Linear Algebra (Brussels, 1991), North-Holland, Amsterdam, 1992, pp. 229–231.
[40] E. L. Wachspress, *The ADI Model Problem*, Self published, Windsor, CA, 1995.
[41] O. B. Widlund, *On the rate of convergence of an alternating direction implicit method in a noncommutative case*, Math. Comp., 20 (1966), pp. 500–515.
[42] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford Sci. Publ., Oxford University Press, New York, 1988.