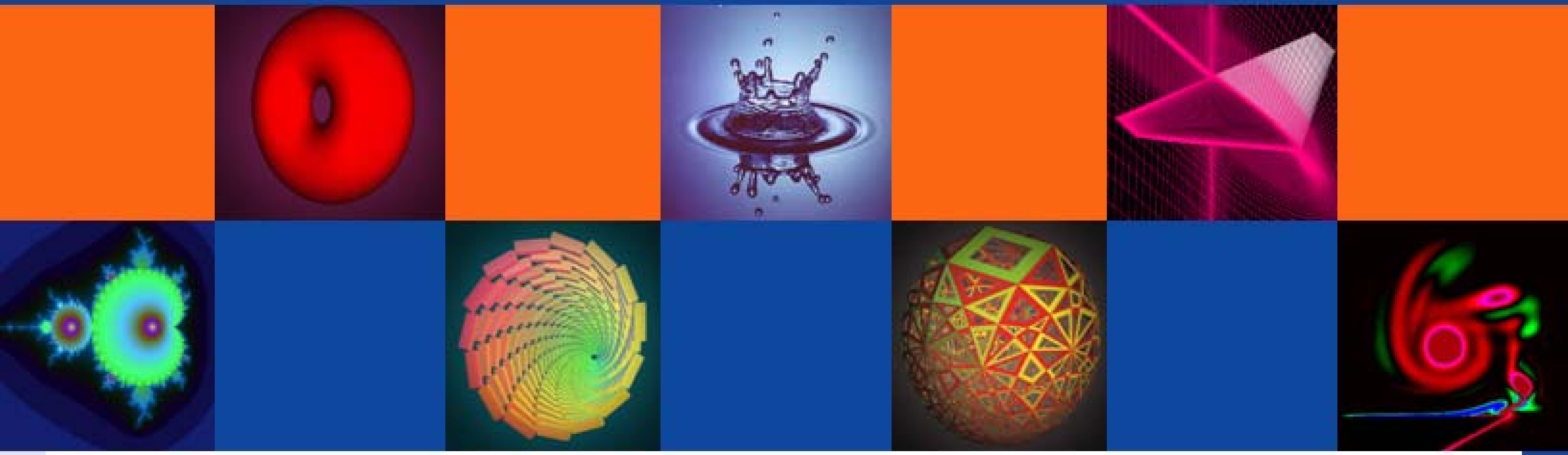
The University of Texas ARLINGTON

MATHEMATICS

Preprint Series



Low Rank ADI Solution of Sylvester Equation via Exact Shifts

Ninoslav Truhar

Zoran Tomljanovic

Ren-Cang Li

Technical Report 2009-03

http://www.uta.edu/math/preprint/

Low Rank ADI Solution of Sylvester Equation via Exact Shifts

Ninoslav Truhar *

Department of Mathematics, University J.J. Strossmayer, Trg Ljudevita Gaja 6, 31000 Osijek, Croatia

Zoran Tomljanović

Department of Mathematics, University J.J. Strossmayer, Trg Ljudevita Gaja 6, 31000 Osijek, Croatia

Ren-Cang Li [1]

Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019, USA

Abstract

The solution to a general Sylvester equation $AX - XB = GF^*$ with a low rank right-hand side is analyzed quantitatively through Low-rank Alternating-Directional-Implicit method (LR-ADI) with exact shifts. New bounds and perturbation bounds on X are obtained. A distinguished feature of these bounds is that they reflect the interplay between the eigenvalue decompositions of A and B and the right-hand side factors G and F. Numerical examples suggest that because of this inclusion of details, new perturbation bounds are much sharper than the existing ones.

Key words: Sylvester equation, Low-rank Alternating-Directional-Implicit (LR-ADI) method, upper bounds, perturbation bounds
1991 MSC: 15A24, 65F10, 65F30, 35B20

---

* Corresponding author.
Email addresses: ntruhar@mathos.hr (Ninoslav Truhar), ztomljan@mathos.hr (Zoran Tomljanović), rcli@uta.edu (Ren-Cang Li).

[1] Supported in part by NSF grant Nos. DMS-0702335 and DMS-0810506.

1 Introduction

In this paper we consider the properties of the solution of $m \times n$ Sylvester equation

$$AX - XB = C, \tag{1.1}$$

where A, B, and C are $m \times m$, $n \times n$, and $m \times n$, respectively, and unknown matrix X is $m \times n$. Equation (1.1) has a unique solution if and only if A and B have no common eigenvalues, which will be assumed throughout this paper.

Sylvester equations appear frequently in many areas of applied mathematics. We refer the reader to the elegant survey by Bhatia and Rosenthal [12] and references therein for a history of the equation and many interesting and important theoretical results. Sylvester equations are important in a number of applications such as matrix eigen-decompositions [21,41], control theory [16,32,41], model reduction [1,2,8,43], numerical solution of matrix differential Riccati equations [15,17,19], and many more.

There are several numerical algorithms for calculation of the solution of Sylvester equations. The standard ones are the Bartels-Stewart algorithm [6] and the Hessenberg-Schur method first described by Enright [19], but more often attributed to Golub, Nash, and Van Loan [20]. Another computationally efficient approach for the case that both A and $-B$ are (Hurwitz) stable, i.e., have all their eigenvalues in the open left half plane, is the sign function method [38]. All these methods are efficient for dense matrices A and B.

However, recent interest is directed more towards large and sparse matrices A and B, and $C = GF^*$ with very low rank, where G and F have only a few columns. For dense A, B, an approach based on the sign function method is suggested in [9] that exploits the low-rank structure of C. This approach is further used in [7] in order to solve large-scale Sylvester equations with data-sparse A, B, i.e., dense matrices A, B that can be represented by $\mathcal{O}(n \log(n))$ data. Common methods for sparse A, B are Krylov subspace based algorithms [4,5,18,25–30,39,40,42] and Alternating-Directional-Implicit (ADI) iterations [10,22,31,33,34,36,49].

On the other hand, the problem of the sensitivity of the solution of the Sylvester equation is also widely studied problem. There are several books which contains results about this, e.g., [37,23,45].

Our main concern is how would the solution behaves when C is a low rank matrix. Our investigation is through applying Low-rank Alternating-Directional-Implicit method (LR-ADI) with exact shifts, i.e., all or part of the eigenvalues of A and B.

Our results show that the right-hand side of Sylvester equation can sometimes greatly influence the norm of X and how it changes in the face of perturbations. The bound on the norm of X can be considered as a proper generalization of the results from [48].

Notation. Throughout this paper, $\mathbb{C}^{n\times m}$ is the set of all $n \times m$ complex matrices, $\mathbb{C}^n = \mathbb{C}^{n\times 1}$, and $\mathbb{C} = \mathbb{C}^1$. Similarly define $\mathbb{R}^{n\times m}$, $\mathbb{R}^n$, and $\mathbb{R}$ except replacing the word *complex* by *real*. I_n (or simply I if its dimension is clear from the context) is the $n \times n$ identity matrix, and e_j is its jth column. The superscript "$\cdot^*$" takes conjugate transpose while "$\cdot^{\mathrm{T}}$" takes transpose only. We shall also adopt MATLAB-like convention to access the entries of vectors and matrices. $i : j$ is the set of integers from i to j inclusive and $i : i = \{i\}$. For a vector u and a matrix X, $u_{(j)}$ is u's jth entry, $X_{(i,j)}$ is X's (i, j)th entry; X's submatrices $X_{(k:\ell,i:j)}$, $X_{(k:\ell,:)}$, and $X_{(:,i:j)}$ consist of intersections of row k to row ℓ and column i to column j, row k to row ℓ, and column i to column j, respectively. $\|\cdot\|$ and $\|\cdot\|_{\mathrm{F}}$ stands for the spectral norm and the Frobenius norm of a matrix, respectively. $\kappa(A) = \|A\|\,\|A^{-1}\|$ is the spectral condition number of A.

2 Low rank ADI for Sylvester Equation

Our first result is a generalization of the main results from [48] where some new estimates for the eigenvalue decay rate of the Lyapunov equation $AX+XA^{\mathrm{T}} = C$ with a low rank right-hand side C has been derived. Our main result will be a bound on the norm of the solution of the following $m\times n$ Sylvester equation

$$AX - XB = GF^*, \tag{2.1}$$

where A, B, G and F are $m \times m$, $n \times n$, $m \times r$ and $n \times r$, respectively, and unknown matrix X is $m \times n$. It is assumed $r \ll \min\{m, n\}$.

But before we continue, we will briefly describe the Low-rank Alternating-Directional-Implicit (LR-ADI) method for solving Sylvester equation (2.1) (more details can be found in [47] or [11]).

Given two sets of parameters $\{\alpha_i\}$ and $\{\beta_i\}$, ADI iteration for iteratively solving (2.1) goes as follows: For $i = 0, 1, \ldots,$

(1) solve $(A - \beta_i I)X_{i+1/2} = X_i(B - \beta_i I) + C$ for $X_{i+1/2}$;
(2) solve $X_{i+1}(B - \alpha_i I) = (A - \alpha_i I)X_{i+1/2} - C$ for X_{i+1},

for an initial guess X_0 which is assumed to be 0 in this paper. A rather straightforward implementation for ADI can be given based on this.

Note that parameters $\{\alpha_i\}$ and $\{\beta_i\}$ in (LR-ADI) method for solving Sylvester equation (2.1) should be chosen such that $\alpha_i \neq \beta_i$, for all i. In the case of Lyapunov equation $AX + XA^{\mathrm{T}} = C$ with the stable matrix A parameters $\{\alpha_i\}$ should be chosen such that $Re(\alpha_i) < 0$ and $\beta_i = -\alpha_i$ for all i, where $Re(z)$ denotes real part of z.

Expressing X_{i+1} in terms of X_i, we have[2]

$$\begin{aligned} X_{i+1} = &(\beta_i - \alpha_i)(A - \beta_i I)^{-1} C (B - \alpha_i I)^{-1} \\ &+ (A - \alpha_i I)(A - \beta_i I)^{-1} X_i (B - \beta_i I)(B - \alpha_i I)^{-1}, \end{aligned}$$

and the error equation

$$\begin{aligned} X_{i+1} - X &= (A - \alpha_i I)(A - \beta_i I)^{-1}(X_i - X)(B - \beta_i I)(B - \alpha_i I)^{-1}, \\ &= \left[\prod_{j=0}^{i}(A - \alpha_j I)(A - \beta_j I)^{-1}\right](X_0 - X)\left[\prod_{j=0}^{i}(B - \beta_j I)(B - \alpha_j I)^{-1}\right], \end{aligned} \tag{2.2}$$

where X denotes the exact solution. If convergence occurs much earlier in the sense that it takes much fewer than $\min\{m, n\}/r$ steps, then ADI in the factored form as below would be more economical. Let $X_i = Z_i D_i Y_i^*$. We have

$$\begin{aligned} X_{i+1} = &\begin{pmatrix} (A - \beta_i I)^{-1} G & (A - \alpha_i I)(A - \beta_i I)^{-1} Z_i \end{pmatrix} \times \\ &\begin{pmatrix} (\beta_i - \alpha_i) I & \\ & D_i \end{pmatrix} \times \\ &\begin{pmatrix} F^*(B - \alpha_i I)^{-1} \\ Y_i^*(B - \beta_i I)(B - \alpha_i I)^{-1} \end{pmatrix} \\ \equiv\ & Z_{i+1} D_{i+1} Y_{i+1}^*, \end{aligned}$$

where

[2] It can be also gotten from the identity

$$(A - \beta I) X (B - \alpha I) - (A - \alpha I) X (B - \beta I) = (\beta - \alpha) C.$$

$$Z_{i+1} = \begin{pmatrix} (A-\beta_i I)^{-1}G & (A-\alpha_i I)(A-\beta_i I)^{-1}Z_i \end{pmatrix},$$
$$D_{i+1} = \begin{pmatrix} (\beta_i-\alpha_i)I & \\ & D_i \end{pmatrix},$$
$$Y_{i+1}^* = \begin{pmatrix} F^*(B-\alpha_i I)^{-1} \\ Y_i^*(B-\beta_i I)(B-\alpha_i I)^{-1} \end{pmatrix}.$$

After renaming the parameters $\{\alpha_i\}$ and $\{\beta_i\}$ as in [47] or [11], since $Z_0 = 0$ and $Y_0 = 0$, we can write,

$$Z_k = \begin{pmatrix} Z^{(1)} & Z^{(2)} & \cdots & Z^{(k)} \end{pmatrix},$$
$$\text{with} \quad \begin{cases} Z^{(1)} = (A-\beta_1 I)^{-1}G, \\ Z^{(i+1)} = (A-\alpha_i I)(A-\beta_{i+1}I)^{-1}Z^{(i)} \\ \qquad\quad = Z^{(i)} + (\beta_{i+1}-\alpha_i)(A-\beta_{i+1}I)^{-1}Z^{(i)}, \end{cases} \tag{2.3}$$

and $Y_k = \begin{pmatrix} Y^{(1)} & Y^{(2)} & \cdots & Y^{(k)} \end{pmatrix}$,

$$\text{with} \quad \begin{cases} Y^{(1)^*} = F^*(B-\alpha_1 I)^{-1}, \\ Y^{(i+1)^*} = Y^{(i)^*}(B-\alpha_{i+1}I)^{-1}(B-\beta_i I) \\ \qquad\quad = Y^{(i)^*} + (\alpha_{i+1}-\beta_i)Y^{(i)^*}(B-\alpha_{i+1}I)^{-1}. \end{cases} \tag{2.4}$$

All together give

$$X_k = Z_k D_k Y_k^*, \quad D_k = \operatorname{diag}\left((\beta_1-\alpha_1)I, \ldots, (\beta_k-\alpha_k)I\right),$$

or

$$X_k = \sum_{j=1}^{k} (\beta_j-\alpha_j) Z^{(j)} Y^{(j)^*}. \tag{2.5}$$

3 The diagonalizable case

In this section we will present an upper bound for the norm of the solution of Sylvester equation (2.1) for diagonalizable matrices A and B, an error bound for the kth LR-ADI solution with exact shifts, and a first order perturbation bound when A, B, G, and F are perturbed slightly.

Suppose A and B are diagonalizable, i.e.,

$$A = S\Lambda S^{-1}, \quad \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_m), \tag{3.1}$$
$$B = T\Omega T^{-1}, \quad \Omega = \mathrm{diag}(\mu_1, \dots, \mu_n). \tag{3.2}$$

Then the $(i+1)$st block of Z_k and Y_k^* can be written as

$$Z^{(i+1)} = S(\Lambda - \alpha_i I)(\Lambda - \beta_{i+1} I)^{-1} S^{-1} Z^{(i)}, \tag{3.3}$$
$$Y^{(i+1)*} = Y^{(i)*} T(\Omega - \alpha_{i+1} I)^{-1}(\Omega - \beta_i I)T^{-1}. \tag{3.4}$$

Define

$$\Delta_i = (\Lambda - \alpha_i\, I)(\Lambda - \beta_i\, I)^{-1},\ (\Delta_i)_{(j,j)} = \frac{\lambda_j - \alpha_i}{\lambda_j - \beta_i},$$
$$\Theta_i = (\Omega - \beta_i\, I)(\Omega - \alpha_i\, I)^{-1},\ (\Theta_i)_{(j,j)} = \frac{\mu_j - \beta_i}{\mu_j - \alpha_i}.$$

Equations (3.3) and (3.4) imply

$$\begin{aligned} Z^{(i+1)} &= S(\Lambda - \alpha_i I)(\Lambda - \beta_{i+1} I)^{-1}(\Lambda - \alpha_{i-1} I)(\Lambda - \beta_i I)^{-1} \cdots \\ &\qquad \cdots (\Lambda - \alpha_1 I)(\Lambda - \beta_2 I)^{-1}(\Lambda - \beta_1 I)^{-1} S^{-1} G \\ &= S(\Lambda - \beta_{i+1} I)^{-1} \Delta_i \Delta_{i-1} \cdots \Delta_1 S^{-1} G, \\ Y^{(i+1)*} &= F^* T(\Omega - \alpha_1 I)^{-1}(\Omega - \alpha_2 I)^{-1}(\Omega - \beta_1 I) \cdots \\ &\qquad \cdots (\Omega - \alpha_{i+1} I)^{-1}(\Omega - \beta_i I)T^{-1} \\ &= F^* T \Theta_1 \Theta_2 \cdots \Theta_i (\Omega - \alpha_{i+1} I)^{-1} T^{-1}. \end{aligned}$$

Notice $(\Lambda - \beta_{i+1} I)^{-1} \Delta_i \Delta_{i-1} \cdots \Delta_1$ is diagonal with its jth diagonal entry

$$\Big((\Lambda - \beta_{i+1} I)^{-1} \Delta_i \Delta_{i-1} \cdots \Delta_1\Big)_{(j,j)} = \prod_{\ell=1}^{i} \frac{\lambda_j - \alpha_\ell}{\lambda_j - \beta_\ell} \cdot \frac{1}{\lambda_j - \beta_{i+1}},$$

and similarly $\Theta_1 \Theta_2 \cdots \Theta_i (\Omega - \alpha_{i+1} I)^{-1}$ is also diagonal with its jth diagonal entry

$$\Big(\Theta_1 \Theta_2 \cdots \Theta_i (\Omega - \alpha_{i+1} I)^{-1}\Big)_{(j,j)} = \prod_{\ell=1}^{i} \frac{\mu_j - \beta_\ell}{\mu_j - \alpha_\ell} \cdot \frac{1}{\mu_j - \alpha_{i+1}},$$

Our first theorem gives an upper bound for the norm of the solution of Sylvester equation (2.1). But before we state the first theorem, we note that (2.2) implies that if $\{\alpha_j\}_{j=0}^{i}$ contains all of A's eigenvalues (multiple eigenvalues counted as many times as their algebraic multiplicities) or if $\{\beta_j\}_{j=0}^{i}$ contains all of B's eigenvalues, then $X_{i+1} - X \equiv 0$. This is because, by the Cayley-Hamilton theorem, $p(A) \equiv 0$ for A's characteristic polynomial

$p(\lambda) \stackrel{\text{def}}{=} \det(\lambda I - A)$ and $q(B) \equiv 0$ for $q(\lambda) \stackrel{\text{def}}{=} \det(\lambda I - B)$. Choose parameters such that

$$\text{either } \alpha_i = \lambda_{p_i} \text{ for } i = 1, \dots, m, \text{ or } \beta_j = \mu_{q_j} \text{ for } j = 1, \dots, n,$$

where $\{p_i\}$ and $\{q_j\}$ denote some permutations of indices $1, \dots, m$ and $1, \dots, n$, respectively. Then the solution X of the Sylvester equation (2.1) can be written as

$$X \equiv X_{n_0} = S \left(\sum_{j=1}^{n_0} (\mu_{q_j} - \lambda_{p_j}) \Phi^{(j)} \Psi^{(j)} \right) T^{-1}, \qquad n_0 = \min\{m, n\}, \tag{3.5}$$

where

$$\Phi^{(j)} = \operatorname{diag} \left(\frac{\sigma(1, j-1)}{\lambda_1 - \mu_{q_j}}, \dots, \frac{\sigma(m, j-1)}{\lambda_m - \mu_{q_j}} \right) S^{-1} G, \tag{3.6}$$

$$\sigma(i, j-1) = \prod_{s=1}^{j-1} \frac{\lambda_i - \lambda_{p_s}}{\lambda_i - \mu_{q_s}}, \qquad \sigma(i, 0) = 1, \quad i = 1, \dots, m, \tag{3.7}$$

and

$$\Psi^{(j)} = F^* T \operatorname{diag} \left(\frac{\tau(1, j-1)}{\mu_1 - \lambda_{p_j}}, \dots, \frac{\tau(n, j-1)}{\mu_n - \lambda_{p_j}} \right), \tag{3.8}$$

$$\tau(i, j-1) = \prod_{s=1}^{j-1} \frac{\mu_i - \mu_{q_s}}{\mu_i - \lambda_{p_s}}, \qquad \tau(i, 0) = 1, \quad i = 1, \dots, n.$$

Before we state our first result, we like to emphasize that equation (3.5) is a proper generalization of the similar equation for the solution of the Lyapunov equation from [48]. It is obtained as a by-product of the ADI Method for Sylvester Equations from [11] or [47]. A similar equation to (3.5) can also be obtained using the eigenvalue decompositions of A and B and the LU decomposition of the corresponding Cauchy matrix presented in [13].

Now, we can state our first result.

Theorem 3.1 *Assume A and B have eigen-decompositions* (3.1) *and* (3.2). *Let X be the solution of the Sylvester equation* (2.1). *Then the following inequality*

$$\|X\| \le \|S\| \|T^{-1}\| \sum_{j=1}^{n_0} |\mu_{q_j} - \lambda_{p_j}| \sum_{i=1}^{m} \frac{|\sigma(i, j-1)| \cdot \|\hat{g}_i\|}{|\lambda_i - \mu_{q_j}|} \sum_{\ell=1}^{n} \frac{|\tau(\ell, j-1)| \, \|\hat{f}_\ell\|}{|\mu_\ell - \lambda_{p_j}|}, \tag{3.9}$$

holds, where $\sigma(\cdot, \cdot)$ and $\tau(\cdot, \cdot)$ are defined as in (3.6) *and* (3.8), *respectively, and $\hat{g}_i$, $\hat{f}_\ell$ denote the ith row of $\hat{G} = S^{-1} G$ and the ℓth column of $\hat{F}^* = F^* T$, respectively.*

Proof. The proof simply follows after taking the norms of the both sides of (3.5) and using (3.6) and (3.8). □

REMARK **3.1** *Note that in the bound from* Theorem 3.1 *we combined each of* G *and* F^* *together with eigenvectors matrices of* A *and* B*. This approach has two advantages. First, if certain eigenvectors of* A *are nearly linearly dependent, then it is straightforward to see that for some* G *it is possible* $\|S^{-1}G\| \ll \|S^{-1}\|\|G\|$*. On the other hand if for some* i*, numbers* $\frac{\sigma(i,j-1)}{\lambda_i-\mu_{q_j}}$ *are large and corresponding* $\|\hat{g}_i\|$ *are small, then again it is possible*

$$\sum_{i=1}^{m} \frac{|\sigma(i,j-1)| \cdot \|\hat{g}_i\|}{|\lambda_i - \mu_{q_j}|} \ll \max_i \frac{\sigma(i,j-1)}{|\lambda_i - \mu_{q_j}|} \cdot \|S^{-1}G\| \,.$$

3.1 Error Bounds

In this section we are going to present an upper bound for the approximation of the solution of Sylvester equation (2.1).

Theorem 3.2 *Assume* A *and* B *have eigen-decompositions* (3.1) *and* (3.2)*. Let* X_k *be the* k*th approximation obtained by* (2.3) – (2.5) *with the set of ADI parameters corresponding to some subset of exact eigenvalues of* A *and* B*, i.e.* $\{\alpha_1, \alpha_2, \ldots, \alpha_k\} = \{\lambda_{p_1}, \lambda_{p_2}, \ldots, \lambda_{p_k}\}$ *and* $\{\beta_1, \beta_2, \ldots, \beta_k\} = \{\mu_{q_1}, \mu_{q_2}, \ldots, \mu_{q_k}\}$*. Then the following inequality*

$$\|X-X_k\| \leq \|S\|\|T^{-1}\| \sum_{j=k+1}^{n_0} |\mu_{q_j}-\lambda_{p_j}| \sum_{i=1}^{m} \frac{|\sigma(i,j-1)| \cdot \|\hat{g}_i\|}{|\lambda_i - \mu_{q_j}|} \sum_{\ell=1}^{n} \frac{|\tau(\ell,j-1)| \, \|\hat{f}_\ell\|}{|\mu_\ell - \lambda_{p_j}|}, \tag{3.10}$$

holds, where $\sigma(\cdot,\cdot)$ *and* $\tau(\cdot,\cdot)$ *are defined as in* (3.6) *and* (3.8)*, and* $\hat{g}_i$ *and* $\hat{f}_\ell$ *are as in* Theorem 3.1*, and* $\{\lambda_{p_j},\, j > k\}$ *and* $\{\lambda_{q_j},\, j > k\}$ *are the eigenvalue subsets of* A *and* B *complement to the ones already used as shifts for obtaining* X_k*.*

Proof. It follows from (2.5) that

$$X_k = \sum_{j=1}^{k} (\mu_{q_j} - \lambda_{p_j}) Z^{(j)} Y^{(j)^*}.$$

For $k = n_0 \overset{\text{def}}{=} \min\{m, n\}$, we have

$$X \equiv X_{n_0} = \sum_{j=1}^{n_0} (\mu_{q_j} - \lambda_{p_j}) Z^{(j)} Y^{(j)^*}.$$

Therefore

$$
\begin{aligned}
X - X_k &= \sum_{j=k+1}^{n_0} (\mu_{q_j} - \lambda_{p_j}) Z^{(j)} Y^{(j)^*} \\
&= S \left(\sum_{j=k+1}^{n_0} (\mu_{q_j} - \lambda_{p_j}) \Phi^{(j)} \Psi^{(j)} \right) T^{-1}. \qquad (3.11)
\end{aligned}
$$

Inequality (3.10) simply follows after taking the norm from both sides in the above equality. □

REMARK **3.2** The bound (3.10) can be used as the upper bound for the decay of singular values of the solution X. The similar bounds can be found in [35], [3] and [48]. Especially, in the case of Lyapunov equation it holds

$$\|X - X_k\| \leq \mathrm{trace}(X) - \mathrm{trace}(X_k),$$

thus one can see that bound (3.10) is a generalization of [48, Theorem 3.1].

Example 3.1 The following example will illustrate the quality of the bound (3.9). Let $A = S\Lambda S^{-1}$ with $\Lambda = \mathrm{diag}(30, 40, 60, 70, 90)$ and

Did you generate A and B using S, $\cdots$ exactly as shown here? If not, you should consider doing so. Then reader could repeat your examples. No need to print A and B. This applies to all examples.

$$
S = \begin{pmatrix}
10.2028 & 0.0153 & 0.4186 & 0.8381 & 0.5028 \\
0.1987 & 10.7468 & 0.8462 & 0.0196 & 0.7095 \\
0.6038 & 0.4451 & 10.5252 & 0.6813 & 0.4289 \\
0.2722 & 0.9318 & 0.2026 & 10.3795 & 0.3046 \\
0.1988 & 0.4660 & 0.6721 & 0.8318 & 10.1897
\end{pmatrix}.
$$

Let $B = T\Omega T^{-1}$ with $\Omega = \mathrm{diag}(31, 41, 61, 71, 91)$ and

$$
T = \begin{pmatrix}
20.1934 & 0.6979 & 0.4966 & 0.6602 & 0.7271 \\
0.6822 & 20.3784 & 0.8998 & 0.3420 & 0.3093 \\
0.3028 & 0.8600 & 20.8216 & 0.2897 & 0.8385 \\
0.5417 & 0.8537 & 0.6449 & 20.3412 & 0.5681 \\
0.1509 & 0.5936 & 0.8180 & 0.5341 & 20.3704
\end{pmatrix}.
$$

Then A and B are given as

$$A = \begin{pmatrix} 29.815 & -0.40683 & 0.99328 & 2.9505 & 2.8681 \\ -0.33502 & 39.816 & 1.4206 & -0.28644 & 3.4595 \\ -1.8052 & -0.94791 & 60.046 & 0.68913 & 1.3953 \\ -1.0330 & -2.6379 & 7.1608 \cdot 10^{-3} & 70.021 & 0.83157 \\ -0.99706 & -1.9850 & -1.7082 & -1.4306 & 90.302 \end{pmatrix},$$

$$B = \begin{pmatrix} 30.936 & 0.20752 & 0.58806 & 1.2338 & 2.0822 \\ -0.36662 & 40.936 & 0.83224 & 0.48647 & 0.72542 \\ -0.43344e & -0.87143 & 60.994 & 0.13814 & 1.2600 \\ -1.0334 & -1.2304 & -0.25766 & 71.042 & 0.62278 \\ -0.37317 & -1.3800 & -1.0989 & -0.47661 & 91.093 \end{pmatrix},$$

where all entries are properly rounded to 5 digits.

Further,

$$G = \begin{pmatrix} 408.112 & 7.952 & 24.152 & 10.892 & 7.952 \\ 0.308 & 214.938 & 8.904 & 18.638 & 9.322 \end{pmatrix}^{\mathrm{T}},$$

$$F = \begin{pmatrix} 0.2978 & -0.0093 & -0.0060 & -0.0092 & -0.0100 \\ -0.0322 & 0.9852 & -0.0409 & -0.0146 & -0.0117 \end{pmatrix}^{\mathrm{T}}.$$

It can be computed that

$$\|X\|_{\mathrm{F}} = 467.76,$$

while the bound (3.9) gives

$$\|X\|_{\mathrm{F}} \leq 213.2.$$

The bound will be tighter if eigenvector matrices S and T become more diagonaly dominant. On the other hand the bound will be attained for example if A and B are diagonal ($S = T = I$) and $F = G = \begin{pmatrix} 1, 0, \ldots, 0 \end{pmatrix}^T$.

3.2 Perturbation Bound

We'll present a perturbation bound for the solution of Sylvester equation (2.1) perturbed to

$$(A + \delta A)(X + \delta X) - (X + \delta X)(B + \delta B) = (G + \delta G)(F + \delta F)^*. \qquad (3.12)$$

Neglecting the second order terms and subtracting the unperturbed Sylvester equation from the perturbed one yield

$$A\,\delta X - \delta X\,B \approx G\,\delta F^* + \delta G\,F^* - \delta A\,X + X\,\delta B. \tag{3.13}$$

Note that we can approximate the solution δX of (3.13) by $\delta X \approx \delta X_1 + \delta X_2 + \delta X_3$, where

$$A\,\delta X_1 - \delta X_1\,B = -\delta A\,X, \tag{3.14}$$
$$A\,\delta X_2 - \delta X_2\,B = -X\delta\,B, \tag{3.15}$$
$$A\,\delta X_3 - \delta X_3\,B = (G\,\delta F^* + \delta G\,F^*). \tag{3.16}$$

For (3.14), we again choose parameters

$$\text{either } \alpha_i = \lambda_i \text{ for } i = 1, \ldots, m, \text{ or } \beta_j = \mu_j \text{ for } j = 1, \ldots, n,$$

to get

$$\delta X_1 = S\left(\sum_{j=1}^{n_0}(\mu_j - \lambda_j)\hat{\Phi}^{(j)}\,\hat{\Psi}^{(j)}\right)T^{-1}, \quad n_0 = \min\{m, n\}, \tag{3.17}$$

where

$$\hat{\Phi}^{(j)} = \operatorname{diag}\left(\frac{\sigma(1, j-1)}{\lambda_1 - \mu_j}, \ldots, \frac{\sigma(m, j-1)}{\lambda_m - \mu_j}\right)S^{-1}\delta A, \quad \sigma(i, j-1) = \prod_{s=1}^{j-1}\frac{\lambda_i - \lambda_s}{\lambda_i - \mu_s}$$

and

$$\hat{\Psi}^{(j)} = XT\operatorname{diag}\left(\frac{\tau(1, j-1)}{\mu_1 - \lambda_j}, \ldots, \frac{\tau(n, j-1)}{\mu_n - \lambda_j}\right), \quad \tau(i, j-1) = \prod_{s=1}^{j-1}\frac{\mu_i - \mu_s}{\mu_i - \lambda_s}.$$

Define

$$\Delta_{\Phi^{(j)}} = \operatorname{diag}\left(\frac{\sigma(1, j-1)}{\lambda_1 - \mu_j}, \ldots, \frac{\sigma(m, j-1)}{\lambda_m - \mu_j}\right), \tag{3.18}$$
$$\Delta_{\Psi^{(j)}} = \operatorname{diag}\left(\frac{\tau(1, j-1)}{\mu_1 - \lambda_j}, \ldots, \frac{\tau(n, j-1)}{\mu_n - \lambda_j}\right), \tag{3.19}$$

and

$$\sigma_{\max}^{(j)} = \|\Delta_{\Phi^{(j)}}\| = \max_i \frac{|\sigma(i, j-1)|}{|\lambda_i - \mu_j|}, \qquad \tau_{\max}^{(j)} = \|\Delta_{\Psi^{(j)}}\| = \max_i \frac{|\tau(i, j-1)|}{|\mu_i - \lambda_j|}. \tag{3.20}$$

Equation (3.17) implies

$$\|\delta X_1\| \leq \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \|S\Delta_{\Phi^{(j)}}\| \|S^{-1}\delta A\| \|X\| \|T\Delta_{\Psi^{(j)}}T^{-1}\|,$$

and thus

$$\frac{\|\delta X_1\|}{\|X\|} \leq \|S^{-1}\delta A\| \|S\| \kappa(T) \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \tau_{\max}^{(j)} \, \sigma_{\max}^{(j)}. \tag{3.21}$$

Similarly for δX_2, we have

$$\|\delta X_2\| \leq \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \|S\Delta_{\Phi^{(j)}}S^{-1}\| \|\delta BT\| \|X\| \|\Delta_{\Psi^{(j)}}T^{-1}\|$$

and consequently

$$\frac{\|\delta X_2\|}{\|X\|} \leq \|\delta BT\| \|T^{-1}\| \kappa(S) \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \tau_{\max}^{(j)} \, \sigma_{\max}^{(j)}. \tag{3.22}$$

Finally for δX_3

$$\|\delta X_3\| \leq \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \|S\Delta_{\Phi^{(j)}}\| \|S^{-1}(G\delta F^* + \delta G F^*)T\| \|\Delta_{\Psi^{(j)}}T^{-1}\|$$

which implies

$$\|\delta X_3\| \leq \|S^{-1}(G\delta F^* + \delta G F^*)T\| \|S\| \|T^{-1}\| \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \tau_{\max}^{(j)} \, \sigma_{\max}^{(j)}. \tag{3.23}$$

Now we can state the new theorem for the perturbation of the Sylvester equation (2.1) perturbed as in (3.12).

Theorem 3.3 *Assume A and B have eigen-decompositions* (3.1) *and* (3.2). *Let X be the solution of the Sylvester equation* (2.1) *and let $X + \delta X$ be the solution of the perturbed Sylvester equation* (3.12). *For sufficiently small*

$$\epsilon = \max\{\|\delta A\|, \|\delta B\|, \|\delta G\|, \|\delta F\|\},$$

we have

$$\frac{\|\delta X\|}{\|X\|} \leq \Bigg(\kappa(T)\|S\| \|S^{-1}\delta A\| + \kappa(S)\|T^{-1}\| \|\delta BT\| + \|S\| \|T^{-1}\| \frac{\|S^{-1}(G\delta F^* + \delta G F^*)T\|}{\|X\|} \Bigg) \gamma + \mathcal{O}(\epsilon^2), \tag{3.24}$$

where $\gamma = \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \tau_{\max}^{(j)} \sigma_{\max}^{(j)}$ *and* $\sigma_{\max}^{(j)}$, $\tau_{\max}^{(j)}$ *are defined as in* (3.20).

Proof. Take the norm of $\delta X \approx \delta X_1 + \delta X_2 + \delta X_3$ and use bounds (3.21), (3.22), (3.23) to get (3.24). □

REMARK **3.3** From (3.17), and similar equalities for δX_2 and δX_3, we can obtain a sharper bound

$$\frac{\|\delta X\|}{\|X\|} \leq \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \, \|S \Delta_{\Phi^{(j)}} S^{-1}\| \, \|T \Delta_{\Psi^{(j)}} T^{-1}\| \times \left(\|\delta A\| + \|\delta B\| + \frac{\|S^{-1}(G\delta F^* + \delta G F^*)T\|}{\|X\|} \right) + \mathcal{O}(\epsilon^2),$$

where diagonal matrices $\Delta_{\Phi^{(j)}}$ and $\Delta_{\Psi^{(j)}}$ are defined as in (3.18) and (3.19).

REMARK **3.4** Using $\|AB\|_{\mathrm{F}} \leq \|A\| \|B\|_{\mathrm{F}}$ and $\|AB\|_{\mathrm{F}} \leq \|A\|_F \|B\|$ (see, e.g., [45, Theorem II. 3.9.]) similarly like in the previous calculations one can obtain following bound for Frobenius norm

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq \left(\kappa(T) \|S\| \|S^{-1} \delta A\| + \kappa(S) \|T^{-1}\| \|\delta B \, T\| + \|S\| \|T^{-1}\| \frac{\|S^{-1}(G\delta F^* + \delta G F^*)T\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \right) \gamma + \mathcal{O}(\epsilon^2), \tag{3.25}$$

where $\gamma = \sum_{j=1}^{n_0} |\mu_j - \lambda_j| \tau_{\max}^{(j)} \sigma_{\max}^{(j)}$ and $\sigma_{\max}^{(j)}$, $\tau_{\max}^{(j)}$ are defined as in (3.20).

As an illustration of the quality of the perturbation bound (3.25) we will present a comparison between this bound and two bounds from [23]. The first one is

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq \sqrt{3} \Psi \epsilon + \mathcal{O}(\epsilon^2), \tag{3.26}$$

where

$$\Psi = \|P^{-1}[\alpha(X^{\mathrm{T}} \otimes I_m) \ - \beta(I_n \otimes X) \ - \delta I_{mn}]\| / \|X\|_{\mathrm{F}}, \qquad P = I_n \otimes A - B^{\mathrm{T}} \otimes I_m$$

and $\epsilon = \max\left\{ \frac{\|\delta A\|_{\mathrm{F}}}{\alpha}, \frac{\|\delta B\|_{\mathrm{F}}}{\beta}, \frac{\|\delta C\|_{\mathrm{F}}}{\delta} \right\}$, while α, β and δ are scaling factors as in [23]. The second bound is a weaker version of (3.26):

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq \sqrt{3} \Phi \epsilon + \mathcal{O}(\epsilon^2), \tag{3.27}$$

where $\Phi = \|P^{-1}\| \dfrac{(\alpha+\beta)\|X\|_{\mathrm{F}} + \delta}{\|X\|_{\mathrm{F}}}$.

As pointed out in [23], the perturbation bound (3.27) with $\alpha = \|A\|_{\mathrm{F}}, \beta = \|B\|_{\mathrm{F}}$ and $\delta = \|C\|_{\mathrm{F}}$ is the one that is usually quoted in the literature for the Sylvester equation. For example, it is also in [16].

The following example compares new bound (3.25) with bounds (3.26) and (3.27).

Example 3.2 Let $A = S\Lambda S^{-1}$ with $\Lambda = \mathrm{diag}(30, 40, 60, 70, 90)$ and

$$S = \begin{pmatrix} 0.052 & 0.494 & 0.935 & 0.592 & 0.827 \\ 0.264 & 0.094 & 0.220 & 0.388 & 0.387 \\ 0.757 & 0.971 & 0.189 & 0.935 & 0.445 \\ 0.435 & 0.424 & 0.280 & 0.826 & 0.009 \\ 0.506 & 0.270 & 0.674 & 0.051 & 0.280 \end{pmatrix}.$$

Let $B = T\Omega T^{-1}$ with $\Omega = \mathrm{diag}(100, 200, 300, 400, 450)$ and

$$T = \begin{pmatrix} 85610 & 89420 & 42300 & 31670 & 97380 \\ 9709 & 64180 & 89230 & 36420 & 28030 \\ 0.3910 & 0.2030 & 0.9910 & 0.6150 & 0.1740 \\ 0.3760 & 0.0680 & 0.4000 & 0.5970 & 0.9210 \\ 0.5150 & 0.8000 & 0.3400 & 0.6850 & 0.2600 \end{pmatrix}.$$

Then A and B are given as

$$A = \begin{pmatrix} 62.226 & 75.155 & -3.0035 & -25.328 & -16.256 \\ 11.270 & 78.342 & 0.14503 & -11.108 & -17.048 \\ 19.526 & 51.318 & 33.357 & 5.7715 & -38.765 \\ 11.559 & 11.144 & -17.062 & 77.153 & -22.014 \\ 16.626 & 8.2033 & -5.7441 & -7.3482 & 38.922 \end{pmatrix},$$

$$B = \begin{pmatrix} 79.750 & 303.69 & -3.0939 \cdot 10^{7} & 3.7518 \cdot 10^{7} & -6.2612 \cdot 10^{6} \\ -65.058 & 291.66 & -1.0842 \cdot 10^{6} & 1.2139 \cdot 10^{7} & -8.3818 \cdot 10^{5} \\ -1.6190 \cdot 10^{-3} & 1.2461 \cdot 10^{-3} & 163.87 & 165.84 & 76.064 \\ -1.3364 \cdot 10^{-3} & 1.9579 \cdot 10^{-3} & -235.44 & 570.39 & 20.569 \\ -2.0331 \cdot 10^{-3} & 1.1994 \cdot 10^{-3} & -130.45 & 232.94 & 344.33 \end{pmatrix},$$

where all entries are properly rounded to 5 digits.

Further,

$$G = \begin{pmatrix} -10 & -60 & 0 & 0 & 0 & 0 \\ -60 & 400 & 0 & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \qquad F = G.$$

Perturbations are

$$\delta B = 10^{-9} \cdot \operatorname{diag}\left(10^{-5}, 10^{-5}, 1, 1, 1\right), \quad \delta A = 0,$$
$$\delta F = \delta G = 0.$$

It can be computed that

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} = 5.05 \cdot 10^{-11},$$

while the perturbation bound (3.26) and (3.27) give

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq 5.42 \cdot 10^{-5}, \quad \frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq 1.01 \cdot 10^{-4},$$

respectively. On the other hand the perturbation bound (3.25) gives

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq 9.19 \cdot 10^{-9}.$$

The above example illustrates that the structure of the matrices F and G from the right-hand side in the Sylvester equation (2.1) sometimes can greatly influence the perturbation of the solution.

Before we explain this influence and a reason why the perturbation bounds (3.24) or (3.25) prevail the bounds (3.26) and (3.27) respectively, recall that all these bounds can be considered as the perturbation bounds for the solution of the linear system

$$Px = c, \quad \text{perturbed to} \quad (P + \delta P)(x + \delta x) = c + \delta c, \tag{3.28}$$

where

$$P = I_n \otimes A - B^* \otimes I_m, \quad c = \mathrm{vec}(C),$$

$$P + \delta P = I_n \otimes (A + \delta A) - (B + \delta B)^* \otimes I_m \quad \text{and} \quad c + \delta c = \mathrm{vec}(C + \delta C).$$

The reason why (3.24) or (3.25) prevail over the bounds (3.26) and (3.27) lie in the fact that they include the range of P and $P + \delta P$ and the orientation of c with the respect to the range of P.

The last property (which illustrates the influence of the right-hand side) of the perturbation bounds also can be found in [14], where Chan and Foulser have shown that if we consider the linear system

$$Px = c, \quad \text{perturbed to} \quad P(x + \delta x) = c + \delta c,$$

then the following bound holds:

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\sigma_{N+1-k}}{\sigma_N} \left(\frac{\|\mathcal{P}_k c\|}{\|c\|} \right)^{-1} \frac{\|\delta c\|}{\|c\|}, \quad k = 1, 2, \ldots, N, \quad N = n \cdot m, \tag{3.29}$$

where $\sigma_1 \geq \ldots \geq \sigma_N$ are singular values of the matrix P, that is

$$P = U \Sigma V^T; \quad \text{and} \quad \mathcal{P}_k = U_k U_k^T, \quad U_k = [\, u_{N+1-k}, \ldots, u_N \,].$$

The bound (3.25) and (3.29) are based on the similar ideas, although (3.25) is more general. As an illustration, in Example 3.2 reset $\delta B = 0$ and

$$\delta G = 10^{-6} \cdot \begin{pmatrix} 1 \; 1 \; 0 \; 0 \; 0 \\ 1 \; 1 \; 0 \; 0 \; 0 \end{pmatrix}^T, \quad \delta F = \delta G.$$

Recall that $\delta C = (G\,\delta F^* + \delta G\, F^*)$, and in (3.29) $\delta c = \mathrm{vec}(\delta C)$. In this case (3.25) gives

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq 2.59 \cdot 10^{-6},$$

while (3.29) gives

$$\frac{\|\delta x\|}{\|x\|} \leq 1.92 \cdot 10^{-7}$$

which illustrates that the both bounds may be of the same quality.

On the other hand Chan and Foulser also obtained a perturbation bound for (3.28) when P is perturbed to $P + \delta P$ and $\delta c = 0$. This bound, given in [14, Theorem 2], for Example 3.2 gives

$$\frac{\|\delta x\|}{\|x\|} \leq 1.84 \cdot 10^{-2},$$

which is obvious less sharp then (3.25) and (3.26).

Since in the application often eigenvalues of A and B are close, interesting question arises: what will happen with the bound (3.9) in that case. As was expected the bound (3.9) are not sensitive on clustered eigenvalues, that is if some eigenvalues of A are close to some of eigenvalues of B the bound will increase proportionally with the norm of solution. On the other hand, the perturbation bound (3.24) is more sensitive to the clustered eigenvalues then the bound (3.9), but for example if the range of the matrices on the right-hand side is close to the range of the eigenvector matrices the bound (3.9) would remain very tight.

REMARK **3.5** It is important to note that bounds (3.24) and (3.25) have more theoretical meaning than practical one. In practice, when applying ADI, one does not necessarily choose the eigenvalues of A and B as the ADI shifts. One would use Ritz values (see for example [11] or [47]) as opposed to a some subset of the exact eigenvalues to extract information. Thus, the natural question arise, *is it possible to obtained a certain version of the upper bound in* (3.24) *that uses arbitrary shifts instead of the exact eigenvalues*? Unfortunately at this moment we do not have an answer to this question. But we would like to emphasize that certain attempt in the same direction for the Lyapunov equation has been done in [46]. The bound obtained there is very complicated and hard to apply thus from this point of view if one would like to derive a version of the upper bound in (3.24) some different approach should be implemented.

4 The non-diagonalizable case

This section accomplishes the same tasks as the previous section but for non-diagonalizable matrices A and B.

Recall that equation (2.1) has a unique solution if and only if A and B have no common eigenvalues, which has been assumed throughout the paper. For the sake of simplicity, we will consider matrices A and B whose Jordan blocks are at most 2×2. The idea is easily extensible to more general cases, except more complicated results. On the other hand, the case with only up to 2×2 Jordan blocks does happen in practice. For example a simple planar spacecraft model [44] has two 2×2 Jordan blocks. This makes our case more interesting for investigation.

Let the Jordan canonical form of A be

$$A = SJ_AS^{-1}, \quad S \in \mathbb{C}^{m\times m}, \quad J_A = J_{A,1} \oplus \ldots \oplus J_{A,k_A}, \tag{4.1}$$

where $J_{A,i} \oplus J_{A,j}$ stands for a direct sum of $J_{A,i}$ and $J_{A,j}$, and

$$
\begin{aligned}
J_{A,i} &= \lambda_i \qquad\qquad \text{for} \quad i = 1, \ldots, \ell_A\,, \\
J_{A,i} &= \begin{pmatrix} \lambda_i & 1 \\ 0 & \lambda_i \end{pmatrix} \equiv \lambda_i I_2 + N_2\,, \ \text{for} \quad i = \ell_A + 1, \ldots, k_A\,,
\end{aligned}
$$

$2(k_A - \ell_A) + \ell_A = m$, I_2 is the 2×2 identity matrix, and $N_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Similarly, let the Jordan canonical form of B be

$$
\begin{aligned}
&B = TJ_BT^{-1},\ T \in \mathbb{C}^{n \times n},\ J_B = J_{B,1} \oplus \ldots \oplus J_{B,k_B}, \\
&J_{B,i} = \mu_i \qquad\qquad \text{for} \quad i = 1, \ldots, \ell_B\,, \\
&J_{B,i} = \begin{pmatrix} \mu_i & 1 \\ 0 & \mu_i \end{pmatrix} \equiv \mu_i I_2 + N_2\,, \ \text{for} \quad i = \ell_B + 1, \ldots, k_B\,,
\end{aligned}
\tag{4.2}
$$

where $2(k_B - \ell_B) + \ell_B = n$.

Theorem 4.1 *Assume A and B have Jordan canonical decompositions* (4.1) *and* (4.2). *Let X be the solution of the Sylvester equation* (2.1), *obtained by* (2.3) – (2.5) (LR-ADI) *with the set of ADI parameters*

$$\{\alpha_1, \alpha_2, \ldots, \alpha_m\} = \{\lambda_{p_1}, \lambda_{p_2}, \ldots, \lambda_{p_m}\}$$

and

$$\{\beta_1, \beta_2, \ldots, \beta_n\} = \{\mu_{q_1}, \mu_{q_2}, \ldots, \mu_{q_n}\},$$

where each eigenvalue appears as many times as its algebraic multiplicity. Then

$$\|X\| \le \|S\| \|T^{-1}\| \sum_{j=1}^{n_0} |\mu_{q_j} - \lambda_{p_j}| \sum_{i=1}^{k_A} \frac{||\eta(i, j-1)|| \cdot \|\hat{g}_i\|}{|\lambda_i - \mu_{q_j}|} \sum_{s=1}^{k_B} \frac{||\vartheta(s, j-1)||\, \|\hat{f}_s\|}{|\mu_s - \lambda_{p_j}|}\,, \tag{4.3}$$

where $n_0 = \min\{m, n\}$,

$$\eta(i,j) = \sigma(i,j) \quad \text{for } i = 1, \dots \ell_A, \tag{4.4}$$

$$\eta(i,j) = \left(I_2 - \frac{1}{\lambda_i - \mu_{q_{j+1}}} N_2\right) [\sigma(i,j) I_2 - \mu(i,j)\, N_2] \ \text{ for } i = \ell_A + 1, \dots k_A, \tag{4.5}$$

$$\sigma(i,j-1) = \prod_{s=1}^{j-1} \frac{\lambda_i - \lambda_{p_s}}{\lambda_i - \mu_{q_s}} \ \text{ and } \sigma(i,0) = 1, \quad \text{for all } i, \tag{4.6}$$

$$\mu(i,j) = \sum_{\ell=1}^{j} \prod_{\substack{t=1 \\ t \neq \ell}}^{j} \frac{\lambda_i - \lambda_{p_\ell}}{\lambda_i - \mu_{q_\ell}} \frac{\lambda_{p_t} - \mu_{q_t}}{(\lambda_i - \mu_{q_t})^2} \ \text{ for } i = \ell_A + 1, \dots k_A, \tag{4.7}$$

$$\vartheta(i,j) = \tau(i,j) \quad \text{for} \quad i = 1, \dots \ell_B, \tag{4.8}$$

$$\vartheta(i,j) = \left(I_2 - \frac{1}{\mu_i - \lambda_{p_{j+1}}} N_2\right) [\tau(i,j) I_2 - \nu(i,j)\, N_2] \ \text{ for } i = \ell_B + 1, \dots k_B, \tag{4.9}$$

$$\tau(i,j-1) = \prod_{s=1}^{j-1} \frac{\mu_i - \mu_{q_s}}{\mu_i - \lambda_{p_s}} \ \text{ and } \tau(i,0) = 1, \quad \text{for all } i, \tag{4.10}$$

$$\nu(i,j) = \sum_{\ell=1}^{j} \prod_{\substack{t=1 \\ t \neq \ell}}^{j} \frac{\mu_i - \mu_{q_\ell}}{\mu_i - \lambda_{p_\ell}} \frac{\mu_{q_t} - \lambda_{p_t}}{(\mu_i - \lambda_{p_t})^2} \ \text{ for } i = \ell_B + 1, \dots k_B, \tag{4.11}$$

and $\hat{g}_i$ *(for* $i = 1, \dots, \ell_A$*) denotes the ith* $1 \times r$ *submatrix and* $2 \times r$ *(for* $i = \ell_A + 1, \dots, k_A$*) submatrix of the matrix* $\hat{G} = S^{-1}G$*, respectively. Similarly,* $\hat{f}_j$ *denotes the jth,* $r \times 1$ *(for* $j = 1, \dots, \ell_B$*) submatrix and* $r \times 2$ *(for* $j = \ell_B + 1, \dots, k_B$*) submatrix of the matrix* $\hat{F}^* = F^*T$*, respectively.*

Proof. Similarly to (3.3), one gets

$$\begin{aligned} Z^{(j)} &= S(J_A - \lambda_{p_{j-1}} I)(J_A - \mu_{q_j} I)^{-1} S^{-1} Z^{(j-1)} \\ &= S(J_A - \lambda_{p_{j-1}} I)(J_A - \mu_{q_j} I)^{-1} (J_A - \lambda_{p_{j-2}} I)(J_A - \mu_{q_{j-1}} I)^{-1} \dots \\ &\quad \dots (J_A - \lambda_{p_1} I)(J_A - \mu_{q_2} I)^{-1} (J_A - \mu_{q_1} I)^{-1} S^{-1} G \\ &= S(J_A - \mu_{q_j} I)^{-1} \Gamma_{j-1} \Gamma_{j-2} \dots \Gamma_1 S^{-1} G, \end{aligned} \tag{4.12}$$

where

$$\Gamma_j = (J_A - \lambda_{p_j}\, I)(J_A - \mu_{q_j}\, I)^{-1}.$$

Similarly to (3.4), one gets

$$\begin{aligned} Y^{(j)^*} &= Y^{(j-1)^*} T (J_B - \lambda_{p_j} I)^{-1} (J_B - \mu_{q_{j-1}} I) T^{-1} \\ &= F^* T (J_B - \lambda_{p_1} I)^{-1} (J_B - \lambda_{p_2} I)^{-1} (J_B - \mu_{q_1} I) \cdots \\ &\quad \cdots (J_B - \lambda_{p_j} I)^{-1} (J_B - \mu_{q_{j-1}} I) T^{-1} \\ &= F^* T \Xi_1 \Xi_2 \dots \Xi_{j-1} (J_B - \lambda_{p_j} I)^{-1} T^{-1}, \end{aligned}$$

where

$$\Xi_j = (J_B - \mu_{q_j} I)(J_B - \lambda_{p_j} I)^{-1}\,.$$

Further, note that matrices Γ_j and Ξ_j are block diagonal matrices, and the ith diagonal block of Γ_j is

$$(\Gamma_j)_{(i,i)} = \frac{\lambda_i - \lambda_{p_j}}{\lambda_i - \mu_{q_j}} \qquad \text{for} \quad i = 1, \dots, \ell_A, \tag{4.13}$$

$$(\Gamma_j)_{(i,i)} = \frac{\lambda_i - \lambda_{p_j}}{\lambda_i - \mu_{q_j}} I_2 + \frac{\lambda_{p_j} - \mu_{q_j}}{(\lambda_i - \mu_{q_j})^2} N_2 \quad \text{for} \quad i = \ell_A + 1, \dots, k_A, \tag{4.14}$$

while the ith diagonal block of Ξ_j is

$$(\Xi_j)_{(i,i)} = \frac{\mu_i - \mu_{q_j}}{\mu_i - \lambda_{p_j}} \quad \text{for} \quad i = 1, \dots, \ell_B, \tag{4.15}$$

$$(\Xi_j)_{(i,i)} = \frac{\mu_i - \mu_{q_j}}{\mu_i - \lambda_{p_j}} I_2 + \frac{\mu_{q_j} - \lambda_{p_j}}{(\mu_i - \lambda_{p_j})^2} N_2 \quad \text{for} \quad i = \ell_B + 1, \dots, k_B. \tag{4.16}$$

Now from (4.13) and (4.14), it follows that the ith diagonal block of $(J_A - \mu_{q_j} I)^{-1}\Gamma_{j-1}\Gamma_{j-2}\dots\Gamma_1$ is

$$\begin{aligned}\left((J_A - \mu_{q_j} I)^{-1}\Gamma_{j-1}\Gamma_{j-2}\dots\Gamma_1\right)_{(i,i)} &= \frac{1}{\lambda_i - \mu_{q_j}} \prod_{s=1}^{j-1} \frac{\lambda_i - \lambda_{p_s}}{\lambda_i - \mu_{q_s}} \\ &= \frac{\sigma(i, j-1)}{\lambda_i - \mu_{q_j}} \quad \text{for} \quad i = 1, \dots, \ell_A,\end{aligned} \tag{4.17}$$

and for $i = \ell_A + 1, \dots, k_A$ it is

$$\begin{aligned}&\left(I_2 - \frac{1}{\lambda_i - \mu_{q_j}} N_2\right)\left(\frac{1}{\lambda_i - \mu_{q_j}} \prod_{s=1}^{j-1} \frac{\lambda_i - \lambda_{p_s}}{\lambda_i - \mu_{q_s}} I_2 + \frac{1}{\lambda_i - \mu_{q_j}} \sum_{\ell=1}^{j-1} \prod_{\substack{t=1 \\ t\neq\ell}}^{j-1} \frac{\lambda_i - \lambda_{p_\ell}}{\lambda_i - \mu_{q_\ell}} \, \frac{\lambda_{p_t} - \mu_{q_t}}{(\lambda_i - \mu_{q_t})^2} N_2\right) \\ &= \left(I_2 - \frac{1}{\lambda_i - \mu_{q_j}} N_2\right)\left(\frac{\sigma(i, j-1)}{\lambda_i - \mu_{q_j}} I_2 \; + \; \frac{\mu(i, j-1)}{\lambda_i - \mu_{q_j}} N_2\right),\end{aligned} \tag{4.18}$$

where $\sigma(i, j-1)$ and $\mu(i, j-1)$ are defined as in (4.6)–(4.7).

Similarly from (4.15) and (4.16), it follows that the ith diagonal block of $\Xi_1\Xi_2\dots\Xi_{j-1}(J_B - \lambda_{p_j} I)^{-1}$ is

$$\left(\Xi_1\Xi_2\ldots\Xi_{j-1}(J_B-\lambda_{p_j}I)^{-1}\right)_{(i,i)} = \frac{1}{\mu_i-\lambda_{p_j}}\prod_{s=1}^{j-1}\frac{\mu_i-\mu_{q_s}}{\mu_i-\lambda_{p_s}} = \frac{\tau(i,j-1)}{\mu_i-\lambda_{p_j}} \quad \text{for} \quad i=1,\ldots,\ell_B, \tag{4.19}$$

and for $i=\ell_B+1,\ldots,k_B$ it is

$$\left(I_2-\frac{1}{\mu_i-\lambda_{p_j}}N_2\right)\left(\frac{1}{\mu_i-\lambda_{p_j}}\prod_{s=1}^{j-1}\frac{\mu_i-\mu_{q_s}}{\mu_i-\lambda_{p_s}}I_2+\frac{1}{\mu_i-\lambda_{p_j}}\sum_{\ell=1}^{j-1}\prod_{\substack{t=1\\t\neq\ell}}^{j-1}\frac{\mu_i-\mu_{q_\ell}}{\mu_i-\lambda_{p_\ell}}\frac{\mu_{q_t}-\lambda_{p_t}}{(\mu_i-\lambda_{p_t})^2}N_2\right)$$
$$=\left(I_2-\frac{1}{\mu_i-\lambda_{p_j}}N_2\right)\left(\frac{\tau(i,j-1)}{\mu_i-\lambda_{p_j}}I_2+\frac{\nu(i,j-1)}{\mu_i-\lambda_{p_j}}N_2\right), \tag{4.20}$$

where $\tau(i,j-1)$ and $\nu(i,j-1)$ are defined as in (4.10)–(4.11).

Now, from (2.5) it follows that

$$X=\sum_{j=1}^{n_0}(\mu_{q_j}-\lambda_{p_j})Z^{(j)}Y^{(j)^*}$$
$$=S\left(\sum_{j=1}^{n_0}(\mu_{q_j}-\lambda_{p_j})(J_A-\mu_{q_j}I)^{-1}\Gamma_{j-1}\Gamma_{j-2}\ldots\Gamma_1\hat{G}\hat{F}^*\Xi_1\Xi_2\ldots\Xi_{j-1}(J_B-\lambda_{p_j}I)^{-1}\right)T^{-1}. \tag{4.21}$$

Finally, (4.3) simply follows by taking the norms at the both sides of the above equation and taking into consideration the definitions (4.4)–(4.11). □

4.1 Error Bound

Theorem 4.2 *Assume A and B having Jordan canonical decompositions* (4.1) *and* (4.2)*. Let X_k be the kth approximation obtained by* (2.3) – (2.5) *with the set of ADI parameters corresponding to subsets of exact eigenvalues of A and B, i.e. $\{\alpha_1,\alpha_2,\ldots,\alpha_k\}=\{\lambda_{p_1},\lambda_{p_2},\ldots,\lambda_{p_k}\}$ and $\{\beta_1,\beta_2,\ldots,\beta_k\}=\{\mu_{q_1},\mu_{q_2},\ldots,\mu_{q_k}\}$. Then the following inequality holds*

$$\|X-X_k\|\leq\|S\|\|T^{-1}\|\sum_{j=k+1}^{n_0}|\mu_{q_j}-\lambda_{p_j}|\sum_{i=1}^{k_A}\frac{||\eta(i,j-1)||\cdot\|\hat{g}_i\|}{|\lambda_i-\mu_{q_j}|}\sum_{\ell=1}^{k_B}\frac{||\vartheta(\ell,j-1)||\,\|\hat{f}_\ell\|}{|\mu_\ell-\lambda_{p_j}|},$$

where $\eta(i,j)$ and $\vartheta(\ell,j)$ are defined as in (4.4)–(4.11)*, respectively, and $\hat{g}_i$, $\hat{f}_\ell$ are defined as in* Theorem 4.1*, and $\{\lambda_{p_j},\ j>k\}$ and $\{\lambda_{q_j},\ j>k\}$ are the eigenvalue subsets of A and B complement to the ones already used as shifts for obtaining X_k.*

Proof. An equation similar to (3.11) holds. Taking norms and using formulas from (4.12) to (4.20), we obtain the assertion of the theorem. □

4.2 Perturbation Bound

We'll consider the same problem as we did in Subsection 3.2, except A and B are no longer diagonalizable but have the Jordan canonical forms as in (4.1) and (4.2), respectively. It can be seen that the development there up to (3.16) remains valid.

Choose parameters such that

$$\text{either } \alpha_i = \lambda_{p_i} \text{ for } i = 1, \ldots, m, \text{ or } \beta_j = \mu_{q_j} \text{ for } j = 1, \ldots, n,$$

where $p_i = i$ for $1 \leq i \leq \ell_A$ and $p_{\ell_A+2i-1} = p_{\ell_A+2i} = \ell_A + i$ for $1 \leq i \leq 2(k_A - \ell_A)$, and $q_i = i$ for $1 \leq i \leq \ell_B$ and $q_{\ell_B+2i-1} = q_{\ell_B+2i} = \ell_B + i$ for $1 \leq i \leq 2(k_B - \ell_B)$, recall that $\alpha_i \neq \beta_j$, for all i, j.

Then if we apply formula (4.21) to the Sylvester's equation

$$A\delta X_1 - \delta X_1 B = -\delta A X\,,$$

we have

$$\delta X_1 = -S\left(\sum_{j=1}^{n_0}(\mu_{q_j} - \lambda_{p_j})(J_A - \mu_{q_j} I)^{-1}\Gamma_{j-1}\ldots\Gamma_1 S^{-1}\delta A\, X\, T\Xi_1\Xi_2\cdots\Xi_{j-1}(J_B - \lambda_{p_j} I)^{-1}\right)T^{-1}\,,$$

from which we have

$$\|\delta X_1\| \leq \|S\|\kappa(T)\|S^{-1}\delta A\|\|X\|\left(\sum_{j=1}^{n_0}|\mu_{q_j} - \lambda_{p_j}|\,\eta_{\max}^{(j)}\,\vartheta_{\max}^{(j)}\right), \tag{4.22}$$

where $\eta_{\max}^{(j)} = \|(J_A - \mu_{q_j} I)^{-1}\Gamma_{j-1}\Gamma_{j-2}\ldots\Gamma_1\|$ and $\vartheta_{\max}^{(j)} = \|\Xi_1\Xi_2\cdots\Xi_{j-1}(J_B - \lambda_{p_j} I)^{-1}\|$. Further using (4.17), (4.18), (4.4) and (4.5) we obtain

$$\begin{aligned}\eta_{\max}^{(j)} &= \left\|\operatorname{diag}\left(\frac{\eta(1, j-1)}{\lambda_1 - \mu_{q_j}}, \frac{\eta(2, j-1)}{\lambda_2 - \mu_{q_j}}, \ldots, \frac{\eta(k_A, j-1)}{\lambda_{k_A} - \mu_{q_j}}\right)\right\| \\ &= \max_k \frac{\|\eta(k, j-1)\|}{|\lambda_k - \mu_{q_j}|}.\end{aligned} \tag{4.23}$$

Similarly from (4.19), (4.20), (4.8) and (4.9) we have

$$\begin{aligned}\vartheta_{\max}^{(j)} &= \left\| \mathrm{diag}\left(\frac{\vartheta(1,j-1)}{\mu_1 - \lambda_{p_j}}, \frac{\vartheta(2,j-1)}{\mu_2 - \lambda_{p_j}}, \ldots, \frac{\vartheta(k_B,j-1)}{\mu_{k_B} - \lambda_{p_j}} \right) \right\| \\ &= \max_k \frac{\|\vartheta(k,j-1)\|}{|\mu_k - \lambda_{p_j}|}. \end{aligned} \tag{4.24}$$

For the solutions of the Sylvester equations (3.15) and (3.16), one can obtain the following bounds

$$\|\delta X_2\| \le \kappa(S)\|T^{-1}\|\|\delta BT\|\|X\| \left(\sum_{j=1}^{n_0} |\mu_{q_j} - \lambda_{p_j}|\, \eta_{\max}^{(j)}\, \vartheta_{\max}^{(j)} \right), \tag{4.25}$$

$$\|\delta X_3\| \le \|S\|\|T^{-1}\|\|S^{-1}(G\delta F^* + \delta G F^*)T\| \left(\sum_{j=1}^{n_0} |\mu_{q_j} - \lambda_{p_j}|\, \eta_{\max}^{(j)}\, \vartheta_{\max}^{(j)} \right). \tag{4.26}$$

Theorem 4.3 *Assume A and B having Jordan canonical decompositions* (4.1) *and* (4.2). *Let X be the solution of the Sylvester equation* (2.1) *and let $X+\delta X$ be the solution of the perturbed Sylvester equation* (3.12). *If*

$$\epsilon = \max\{\|\delta A\|, \|\delta B\|, \|\delta G\|, \|\delta F\|\}$$

is sufficiently small, then

$$\begin{aligned}\frac{\|\delta X\|}{\|X\|} \le \Big(&\kappa(T)\|S\|\|S^{-1}\delta A\| + \kappa(S)\|T^{-1}\|\|\delta BT\| \\ &+ \|S\|\|T^{-1}\| \frac{\|S^{-1}(G\delta F^* + \delta G F^*)T\|}{\|X\|} \Big) \gamma + \mathcal{O}(\epsilon^2), \end{aligned} \tag{4.27}$$

where $\gamma = \sum_{j=1}^{n_0} |\mu_{q_j} - \lambda_{p_j}| \vartheta_{\max}^{(j)}\, \eta_{\max}^{(j)}$, and $\eta_{\max}^{(j)}$ and $\vartheta_{\max}^{(j)}$ are defined as in (4.23) *and* (4.24).

Proof. From $\|\delta X\| \le \|\delta X_1\| + \|\delta X_2\| + \|\delta X_3\| + \mathcal{O}(\epsilon^2)$ using previous bounds (4.22), (4.25) and (4.26) for $\|\delta X_1\|$, $\|\delta X_2\|$ and $\|\delta X_3\|$, respectively we obtain assertion of the theorem. □

REMARK **4.1** Again, using $\|AB\|_{\mathrm{F}} \le \|A\|\|B\|_{\mathrm{F}}$ and $\|AB\|_{\mathrm{F}} \le \|A\|_F\|B\|$ ([45, Theorem II. 3.9.], similarly like in the previous calculations, one can obtain following bound for Frobenius norm

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq \left(\kappa(T)\|S\|\|S^{-1}\delta A\| + \kappa(S)\|T^{-1}\|\|\delta BT\| \right.$$
$$\left. + \|S\|\|T^{-1}\| \frac{\|S^{-1}(G\delta F^* + \delta GF^*)T\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \right) \gamma + \mathcal{O}(\epsilon^2)\,. \qquad (4.28)$$

where $\gamma = \sum_{j=1}^{n_0} |\mu_{q_j} - \lambda_{p_j}| \vartheta_{\max}^{(j)}\, \eta_{\max}^{(j)}$, and $\eta_{\max}^{(j)}$ and $\vartheta_{\max}^{(j)}$ are defined as in (4.23) and (4.24).

The bounds (4.27) and (4.28) share the same properties as the bounds for the diagonalizable case. In order to compare new bound (4.28) with bounds (3.26) and (3.27), we consider following example.

Example 4.1 Let $A = SJ_AS^{-1}$ with $J_A = 10 \oplus 20 \oplus (40I_2 + N_2) \oplus (60I_2 + N_2)$ and

$$S = \begin{pmatrix} 10.1610 & -0.2270 & -0.0810 & -0.0520 & 0.0120 & -1.3400 \\ -0.0990 & 2.5300 & -0.0450 & -0.0110 & 0.0310 & -1.5770 \\ 0.0040 & 0.0002 & 0.0110 & 0 & 0.0010 & -0.0490 \\ 0.0001 & 0.0001 & 0.0200 & 0.0060 & 0 & -0.0060 \\ 0.0100 & 0.0010 & 0 & 0.0003 & 0.0060 & -0.1070 \\ -0.5000 & -0.0720 & -0.0260 & -0.0030 & -0.1060 & 5.3420 \end{pmatrix}.$$

Let $B = TJ_BT^{-1}$ with $J_B = 1.1 \oplus 2.2 \oplus 3.3 \oplus 4.4 \oplus (5.5I_2 + N_2)$ and

$$T = \begin{pmatrix} 0.1000 & 0.0100 & 0.0073 & 0.0092 & 0.0018 & 0.0010 \\ 0.0100 & 0.4000 & 0.0200 & 0.0024 & 0.0078 & 0.0087 \\ 0.5230 & 2.0000 & 90.0000 & 3.0000 & 0.1266 & 0.9222 \\ 0.8643 & 0.5264 & 3.0000 & 160.0000 & 4.0000 & 0.5596 \\ 0.9746 & 0.0016 & 0.3493 & 4.0000 & 250.0000 & 5.0000 \\ 0.7726 & 0.7227 & 0.0952 & 0.9762 & 5.0000 & 0.7091 \end{pmatrix}.$$

Then A and B are given as

$$A = \begin{pmatrix} 9.3438 & -1.3180 & 240.46 & -279.17 & -200.96 & -15.227 \\ -0.50249 & 19.609 & -16.181 & -55.113 & -6.3517 & -12.382 \\ -2.1456 \cdot 10^{-2} & -9.5733 \cdot 10^{-3} & 36.534 & 1.5135 & 0.13783 & -0.21881 \\ -1.7008 \cdot 10^{-3} & -2.8580 \cdot 10^{-3} & -6.2176 & 43.294 & -0.47523 & -8.6585 \cdot 10^{-2} \\ -5.0208 \cdot 10^{-2} & -2.0441 \cdot 10^{-2} & 2.1992 & -1.4633 & 59.780 & -3.3920 \cdot 10^{-2} \\ 2.5271 & 1.3994 & 19.810 & 30.361 & 9.8637 & 61.441 \end{pmatrix},$$

$$B = \begin{pmatrix} 1.0248 & 1.1159 \cdot 10^{-2} & 1.6806 \cdot 10^{-4} & 1.3669 \cdot 10^{-4} & -1.6080 \cdot 10^{-4} & 9.5199 \cdot 10^{-3} \\ -0.56104 & 2.1026 & 2.6065 \cdot 10^{-4} & -2.8371 \cdot 10^{-4} & -1.1092 \cdot 10^{-3} & 6.1180 \cdot 10^{-2} \\ -38.622 & -11.312 & 3.3019 & 1.9604 \cdot 10^{-3} & -7.3187 \cdot 10^{-2} & 3.7450 \\ -85.894 & -14.725 & -3.2912 \cdot 10^{-2} & 4.3613 & -0.13780 & 7.8559 \\ -3170 & -689.19 & 5.7388 \cdot 10^{-2} & -2.2285 & -2.9559 & 426.79 \\ -96.947 & -19.144 & 2.9388 \cdot 10^{-3} & -4.9494 \cdot 10^{-2} & -0.17122 & 14.165 \end{pmatrix},$$

where all entries are properly rounded to 5 digits.

Further,

$$G = \begin{pmatrix} 20 & 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \qquad F = G.$$

Perturbations are

$$\delta A = 10^{-10} \cdot \mathrm{diag}\,(0.2, 0.2, 0.0002, 0.0002, 0.0002, 0.0002)\,, \quad \delta B = 5\,\delta A,$$
$$\delta G = 10^{-6} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \delta F = \delta G.$$

It can be computed that

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} = 1.05 \cdot 10^{-7},$$

while the perturbation bounds (3.26) and (3.27) give

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \le 0.11, \quad \frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \le 1.45,$$

respectively. On the other hand, the perturbation bound (4.28) gives

$$\frac{\|\delta X\|_{\mathrm{F}}}{\|X\|_{\mathrm{F}}} \leq 2.84 \cdot 10^{-6}.$$

As in the diagonalizable case, the above example shows that the structure of the matrices F and G from the Sylvester equation (2.1) sometimes can greatly influence the perturbation of the solution.

5 Concluding remarks

We have analyzed the solution to a general Sylvester equation $AX - XB = GF^*$ with a low rank-right hand side. LR-ADI with the exact shifts provides us the tool to do so. Our new results contain considerably more detailed information on the eigen-properties of A and B and the right-hand side GF^* as opposed to the existing ones. Because of this, our new bounds are sharper and provides better understanding of the solution structure, but are messier as a tradeoff.

Although we tackled the general case by considering when A and B have Jordan blocks of orders only up to 2, the technique is readily applicable to Jordan blocks of orders higher than 2 with little changes.

Acknowledgement. The authors would like to thank the referees for their remarks which have helped to clarify some of our statements and arguments and improved the quality of the paper.

References

[1] R. Aldhaheri, Model order reduction via real Schur-form decomposition, Internat. J. Control 53 (3) (1991) 709–716.

[2] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, Advances in Design and Control, SIAM, Philadelphia, PA, 2005.

[3] A. C. Antoulas, D. C. Sorensen, and Y. Zhou, *On the decay rate of hankel singular values and related issues*, Systems Control Lett. 46 (2002) 323–342.

[4] L. Bao, Y. Lin, Y. Wei, Krylov subspace methods for the generalized Sylvester equation, Appl. Math. Comput. 175 (2006) 557–573.

[5] L. Bao, Y. Lin, Y. Wei, A new projection method for solving large Sylvester equations, Appl. Numer. Math. 57 (5–7) (2007) 521–532.

[6] R. H. Bartels and G. W. Stewart, *Algorithm 432: The solution of the matrix equation $AX - BX = C$*, Commun. ACM 8 (1972) 820–826.

[7] U. Baur, Low rank solution of data-sparse Sylvester equations, Numer. Lin. Alg. Appl. (2008) to appear.

[8] U. Baur, P. Benner, Cross-gramian based model reduction for data-sparse systems, Tech. rep., Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, FRG, submitted for publication (Dec. 2007).

[9] P. Benner, Factorized solution of Sylvester equations with applications in control, in: Proc. Intl. Symp. Math. Theory Networks and Syst. MTNS 2004 (CD-ROM), Leuven, Belgium, 2004, (10 pages).

[10] P. Benner, J.-R. Li, T. Penzl, Numerical solution of large Lyapunov equations, Riccati equations, and linear-quadratic control problems, Numer. Lin. Alg. Appl. (2008) to appear, Reprint of unpublished manuscript, University of Calgary, 1999.

[11] P. Benner, R.-C. Li, N. Truhar, On ADI Method for Sylvester Equations, submitted for publication

[12] R. Bhatia and P. Rosenthal, *How and why to solve the operator equation $AX - XB = Y$*, Bull. London Math. Soc. 29 (1997) 1–21.

[13] D. Calvetti and L. Reichel, *Factorizations of Cauchy matrices*, Journal of Computational and Applied Mathematics 86 (1997) pp. 103–123.

[14] T. F. Chan and D. E. Foulser, *Effectively well-conditioned linear systems*, SIAM J. Sci. Stat. Comput. 6 (9) (1988) 963–969.

[15] C. H. Choi and A. J. Laub, *Efficient matrix-valued algorithm for solving stiff Riccati differential equations*, IEEE Trans. Automat. Control 35 (1990) 770–776.

[16] B. Datta, Numerical Methods for Linear Control Systems, Elsevier Academic Press, 2004.

[17] L. Dieci, *Numerical integration of the differential Riccati equation and some related issues*, SIAM J. Numer. Anal. 29 (1992) 781–815.

[18] A. El Guennouni, K. Jbilou, J. Riquet, Block Krylov subspace methods for solving large Sylvester equations, Numer. Algorithms 29 (2002) 75–96.

[19] W. Enright, Improving the efficiency of matrix operations in the numerical solution of stiff ordinary differential equations, ACM Trans. Math. Softw. 4 (1978) 127–136.

[20] G. H. Golub, S. Nash, and C. F. Van Loan, *Hessenberg-Schur method for the problem $AX + XB = C$*, IEEE Trans. Automat. Control, AC-24 (1979) 909–913.

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

[22] S. Gugercin, D. Sorensen, and A. Antoulas, *A modified low-rank Smith method for large-scale Lyapunov equations*, Numerical Algorithms, 32 (2003) 27–55.

[23] Nicholas J. Higham, Accuracy and Stability of Numerical Algorithms, SIAM, Philadelphia, PA, 1996.

[24] Nicholas J. Higham, Perturbation theory and backward error for $AX = XB = C$, BIT, 33 (1) (1993) 124–136.

[25] M. Hochbruck, G. Starke, Preconditioned Krylov subspace methods for Lyapunov matrix equations, SIAM J. Matrix Anal. Appl. 16 (1) (1995) 156–171.

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

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

[28] I. M. Jaimoukha and E. M. Kasenally, *Oblique projectiob methods for large scale model reduction*, SIAM J. Matrix Anal. Appl. 16 (1995) 602–627.

[29] K. Jbilou, *Low rank approximate solutions to large Sylvester matrix equations*, Appl. Math. Comput. 177 (2006) 365–376.

[30] K. Jbilou, A. Messaoudi, H. Sadok, Global FOM and GMRES algorithms for matrix equations, Appl. Numer. Math. 31 (1999) 49–63.

[31] J.-R. Li and J. White, *Low-rank solution of Lyapunov equations*, SIAM J. Matrix Anal. Appl. 24 (2002) 260–280.

[32] A. Locatelli, Optimal Control: an Introduction, Birkhäuser, Basel, Boston, Berlin, 2001.

[33] A. Lu, E. Wachspress, Solution of Lyapunov equations by ADI iteration, Comp. Math. Appl. 21 (1991) 43–58.

[34] T. Penzl, *A cyclic low-rank smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comput. 21 (2000) 1401–1418.

[35] ——, *Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case*, Systems Control Lett., 40 (2000) 139–144.

[36] ——, *LYAPACK: A MATLAB toolbox for large Lyapunov and Riccati equations, model reduction problems, and linear-quadratic optimal control problems, users' guide (ver. 1.0).* Available at `www.tu-chemnitz.de/sfb393/lyapack/`, 2000.

[37] M. Konstantinov, D. Wei Gu, V. Mehrmann, P. Petkov, Perturbation Theory for Matrix Equations, Elsevier Science B.V.; 2003

[38] J. Roberts, Linear model reduction and solution of the algebraic Riccati equation by use of the sign function, Internat. J. Control 32 (1980) 677–687, reprint of Technical Report No. TR-13, CUED/B-Control, Cambridge University, Engineering Department, 1971.

[39] Y. Saad, *Numerical solution of large Lyapunov equations*, in Signal Processing, Scattering, Operator Theory and Numerical Methods, M. A. Kaashoek, J. H. V. Shuppen, and A. Ran, eds., Birkhaüser, Boston, MA, 1990 503–511.

[40] D. Salkuyeh, F. Toutounian, New approaches for solving large Sylvester equations, Appl. Math. Comput. 173 (1) (2006) 9–18.

[41] V. Sima, Algorithms for Linear-Quadratic Optimization, vol. 200 of Pure and Applied Mathematics, Marcel Dekker, Inc., New York, NY, 1996.

[42] ——, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput. 29 (2007) 1268–1288.

[43] D. Sorensen, A. Antoulas, The Sylvester equation and approximate balanced reduction, Linear Algebra Appl. 351/352 (2002) 671–700.

[44] G. Strang, Linear Algebra and Its Applications, third ed., Harcourt Brace Jovanovich, Philadelphia, PA, 1988.

[45] G. W. Stewart, Ji-guang Sun, Matrix Perturbation Theory, Academic Press, Harcourt Brace Jovanovich, 1990.

[46] N. Truhar, The perturbation bound for the solution of the Lyapunov equation, Mathematical Communication 12 (1) (2007) 83–94.

[47] N. Truhar, R.-C. Li, On ADI Method for Sylvester Equations, Technical Report 2008-02, department of Mathematics, University of Texas at Arlington, Arlington, TX. USA.

[48] N. Truhar and K. Veselić, *Bounds on the trace of solution to the Lyapunov equation with a general stable matrix*, Systems Control Lett. 56 (2007) 493–503.

[49] E. L. Wachspress, *Iterative solution of the Lyapunov matrix equation*, Appl. Math. Lett. 1 (1988) 87–90.