A STRUCTURED DOUBLING ALGORITHM FOR DISCRETE-TIME ALGEBRAIC RICCATI EQUATIONS WITH SINGULAR CONTROL WEIGHTING MATRICES

CHUN-YUEH CHIANG*, HUNG-YUAN FAN†, AND WEN-WEI LIN‡

Abstract. In this paper we propose a structured doubling algorithm for solving discrete-time algebraic Riccati equations without the invertibility of control weighting matrices. In addition, we prove that the convergence of the SDA algorithm is linear with ratio less than $\frac{1}{2}$ when all unimodular eigenvalues of the closed-loop matrix are semisimple. Numerical examples are shown to illustrate the feasibility and efficiency of the proposed algorithm.

1. Introduction. The paper concerns with a structured doubling algorithm (SDA) for solving the symmetric almost stabilizing solution X_s of a discrete-time algebraic Riccati equation (DARE) of the form

$$\mathcal{R}(X) \equiv -X + A^\top XA + Q - (C + B^\top XA)^\top (R + B^\top XB)^{-1}(C + B^\top XA) = 0, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, $Q = Q^\top \in \mathbb{R}^{n\times n}$ and $R = R^\top \in \mathbb{R}^{m\times m}$, respectively. A symmetric solution $X \in \mathbb{R}^{n\times n}$ of (1.1) is called stabilizing (resp., almost stabilizing) if $R+B^\top XB$ is invertible and all the eigenvalues of the closed-loop matrix $A_F \equiv A+BF$ are in the open (resp., closed) unit disk, where

$$F = -(R + B^\top XB)^{-1}(B^\top XA + C). \tag{1.2}$$

Moreover, we say that DARE (1.1) has an almost stabilizing solution X with property **(P)** if X is an almost stabilizing solution to DARE (1.1) and all unimodular eigenvalues of A_F are semisimple. The DARE (1.1) arises, e.g., in (a) filtering or stochastic realization problems, and (b) linear quadratic control problems. In case (a), R is the measurement noise covariance and it is not uncommon for this kind of matrix to be singular. For (b), R is the control weighting matrix and in the discrete-time case, occasionally such a matrix can be singular as well. Therefore, we focus on the DARE (1.1) with a singular matrix R throughout the paper.

As is widely known, the DARE (1.1) and its stabilizing solution X_s originated in the discrete-time linear quadratic control problem (case (b) from above) formulated for the discrete-time system

$$x_{k+1} = Ax_k + Bu_k, \quad k = 0, 1, \ldots, \quad x_0 = \xi, \tag{1.3}$$

in which one wishes to minimize the cost functional

$$J(u) = \sum_{k=0}^{\infty} \begin{bmatrix} x_k^\top & u_k^\top \end{bmatrix} \begin{bmatrix} Q & C^\top \\ C & R \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}, \tag{1.4}$$

where usually $R > 0$ and $Q \geqslant 0$ are considered (see, e.g., [19] and references therein).

Necessary and sufficient conditions for the existence of the stabilizing solution X_s to the DARE (1.1) are derived in [13] without any assumptions on the invertibility of A or positivity

*Department of Mathematics, National Tsing Hua University, Hsinchu, 30013, Taiwan. d907201@oz.nthu.edu.tw

†Department of Applied Mathematics, Hsuan Chuang University, Hsinchu, 30092, Taiwan. hyfan@wmail.hcu.edu.tw

‡Department of Mathematics, National Tsing Hua University, Hsinchu, 30013, Taiwan. wwlin@math.nthu.edu.tw

of R or Q. Note that if the DARE (1.1) has a stabilizing solution X_s, then it is unique [13, Proposition 1].

For any $n\times n$ matrices A and B, the matrix pencil $A-\lambda B$ is called regular if $\det(A-\lambda B) \not\equiv 0$. We shall be concerned only with regular pencils throughout the paper. A k-dimensional subspace χ of $\mathbb{C}^n$ is called a deflating subspace for $A-\lambda B$ if there exists matrices $P_1, P_2 \in \mathbb{C}^{n\times k}$ and $Q_1, Q_2 \in \mathbb{C}^{k\times k}$ such that $AP_1 = P_2Q_1$, $BP_1 = P_2Q_2$ and the columns of P_1 span χ. A deflating subspace χ of the pencil $A-\lambda B$ is called stable if the spectrum of $A-\lambda B$ restricted to χ is contained in the open unit disk. On the other hand, a space $\mathcal{V} \in \mathbb{C}^{2n}$ is called isotropic if $x^H \mathcal{J} y = 0$ for any x, $y \in \mathcal{V}$, where the skew-symmetric matrix $\mathcal{J} \equiv \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$ and I_n is an $n\times n$ identity matrix. A deflating subspace $\mathcal{V} \subseteq \mathbb{C}^{2n}$ of $A-\lambda B \in \mathbb{C}^{2n\times 2n}$ is called a stable Lagrangian subspace if it is a maximal isotropic subspace and the spectrum of $A-\lambda B$ restricted to $\mathcal{V}$ is contained in the closed unit disk. For solving the symmetric stabilizing solution X_s to the DARE (1.1), one common approach is to compute the stable deflating subspace of the extended symplectic pencil (ESP) $\mathcal{M}-\lambda\mathcal{L}$ associated with the DARE (1.1), where

$$\mathcal{M} = \begin{bmatrix} A & 0 & B \\ -Q & I & -C^\top \\ C & 0 & R \end{bmatrix}, \quad \mathcal{L} = \begin{bmatrix} I & 0 & 0 \\ 0 & A^\top & 0 \\ 0 & -B^\top & 0 \end{bmatrix}. \tag{1.5}$$

This dilated pencil appears naturally when writing the canonical system associated to (1.3) and (1.4) in descriptor form. If DARE (1.1) has a symmetric solution, then after some steps of Gauss Elimination, we have

$$\mathcal{M}-\lambda\mathcal{L} \overset{eq.}{\sim} \begin{bmatrix} A+BF-\lambda I & 0 & B \\ 0 & \lambda I-(A+BF)^\top & 0 \\ 0 & \lambda B^\top & R+B^\top XB \end{bmatrix}. \tag{1.6}$$

According to (1.6), a unimodular number λ is an eigenvalue of $A+BF$ with algebraic multiplicity k if and only if it is an eigenvalue of $\mathcal{M}-\lambda\mathcal{L}$ with algebraic multiplicity $2k$. The following results give a complete characterization about the spectrum of ESP.

LEMMA 1.1. *[8] Let X be a solution of* (1.1) *with $R+B^\top XB > 0$. If*

$$\operatorname{rank}[\lambda I - A \;\; B] = n \tag{1.7}$$

for all $\lambda \in \mathbb{C}$ with $|\lambda| = 1$, then the elementary divisors of $A+BF$ corresponding to λ have degrees $k_1, k_2, \dots, k_s (1 \le k_1 \le \dots \le k_s \le n)$ if and only if the elementary divisors of $\mathcal{M}-\lambda\mathcal{L}$ corresponding to λ have degrees $2k_1, \dots, 2k_s$.

THEOREM 1.1. *[13] Suppose that the ESP (1.5) is regular, then we have:*

1. $\deg\det(\mathcal{M}-\lambda\mathcal{L}) \le 2m$.
2. *If $\lambda \neq 0$ is a generalized eigenvalue of $\mathcal{M}-\lambda\mathcal{L}$, then $1/\lambda$ is also a generalized eigenvalue of the same multiplicity.*
3. *If $\lambda = 0$ is a generalized eigenvalue of $\mathcal{M}-\lambda\mathcal{L}$ with multiplicity r, then $\lambda = \infty$ is a generalized eigenvalue of multiplicity $m+r$.*

If the stable deflating subspace χ is spanned by columns of an $(2n+m)\times n$ matrix

$$V = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} \begin{matrix} \}n \\ \}n \\ \}m \end{matrix}, \tag{1.8}$$

and V_1 is invertible, then $X_s = V_2V_1^{-1}$ is the stabilizing solution of DARE (1.1), see, e.g., [22, 7, 13]. Unfortunately, this algorithm does not take into account the symplectic structure of

$\mathcal{M} - \lambda\mathcal{L}$ in (1.5). Non-structure-preserving iterative processes loosen the symplectic structure, and this may cause the algorithm to fail or lose accuracy in adverse circumstances. When $(\mathcal{M}, \mathcal{L})$ has no unimodular eigenvalues and $R > 0$, the quadratically convergent SDA algorithms [6, 11], based on the viewpoint of the inverse-free iteration [1, 17], have been developed for solving the unique symmetric stabilizing solution X_s of DAREs and generalized DAREs, respectively. Extensive numerical experiments show that the SDA is more efficient and outperforms the other algorithms. Therefore, one of our main purposes is to develop a structured doubling algorithm for computing the symmetric stabilizing solution X_s to DARE (1.1) without the restriction $R > 0$.

Since the DARE (1.1) is a quadratic matrix equation, it is natural to apply Newton's method (NTM) to obtain an approximate solution. There are extensive literature on the application of Newton's method for the solution of algebraic Riccati equations, for both the continuous and the discrete case. Recently, an efficient Newton-type method has been proposed by [8] to solve the symmetric maximal solution $X_+ \in \mathbb{R}^{n\times n}$ of DARE (1.1). Sufficient conditions for the existence of the maximal solution X_+ to the DARE (1.1) are given in [8, Theorem 1.1]. For $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$, the pair (A, B) is said to be d-stabilizable if $\operatorname{rank}[\lambda I - A, B] = n$ for all $\lambda \in \mathbb{C}$ with $|\lambda| \geqslant 1$. For real symmetric matrices X and Y, we write $X \geqslant Y$ $(X > Y)$ if $X - Y$ is positive semidefinite (definite). A symmetric solution X_+ of (1.1) is called maximal if $X_+ \geqslant S$ for every symmetric solution S. Maximal and almost stabilizing solutions play important roles in applications, see, e.g., [8, 13, 14, 19] and references given therein. The following theorem tells us that the maximal solution is at least almost stabilizing [8].

THEOREM 1.2. *Let (A, B) be d-stabilizable pair and assume that there is a symmetric solution $\tilde{X}$ of the inequality $\mathcal{R}(X) \geqslant 0$ for which $R + B^\top \tilde{X} B > 0$. Then there exists a maximal symmetric solution X_+ of (1.1). Moreover, $R + B^\top X_+ B > 0$ and all the eigenvalues of the closed-loop matrix A_F lie in the closed unit disk.*

As in Theorem 1.4 of [8], Newton iteration converges quadratically to the symmetric maximal solution X_+ when the same conditions as in Theorem 1.2 are assumed and all eigenvalues of the associated closed-loop matrix A_F are in the open unit disk. In this case, the maximal solution is at least a stabilizing solution of DARE (1.1). Moreover, it has been proven in Theorem 4.3 of [8] that Newton's method converges linearly with ratio $\frac{1}{2}$ to the maximal solution X_+ of the DARE (1.1) under the same conditions as in Theorem 1.2 and the ESP in (1.5) satisfies the following assumption.

(A) All elementary divisors of unimodular eigenvalues of $\mathcal{M} - \lambda\mathcal{L}$ are of degree two.

The assumption **(A)** is equivalent to the condition that all eigenvalues of the closed-loop matrix $A_F = A - B(R + B^\top X_+ B)^{-1}(C + B^\top X_+ A)$ on the unit circle are semisimple if the conditions of Theorem 1.2 are satisfied and the first Fréchet derivative of $\mathcal{R}$ in (1.1) at the maximal solution X_+ is not invetible [8]. In Section 3, under the assumption **(A)**, we shall prove that the SDA algorithm converges linearly with ratio less than $\frac{1}{2}$ to an almost stabilizing solution X_s with property **(P)** of the DARE (1.1).

The paper is organized as follows. In Section 2, we propose a structured doubling algorithm for computing the symmetric stabilizing solution or the maximal solution of DARE (1.1) without the invertibility of R. Furthermore, a min-max optimization problem is proposed for selecting a suitable symmetric matrix Y. The convergence analysis of SDA for solving DAREs (1.1) is shown in Section 3. Some numerical examples with singular matrices R are given in Section 4 to illustrate the efficiency and feasibility of the SDA algorithm. Finally, concluding remarks are given in Section 5. Throughout this paper, we denote $A^H = \bar{A}^\top$ the conjugate transpose of $A \in \mathbb{C}^{n\times n}$, $\imath = \sqrt{-1}$, and $I \equiv I_n$ and $0 \equiv 0_n$ the identity and zero matrices, respectively, of order n. The jth column of I_n is denoted by e_j. $\|\cdot\|$ denotes any matrix norm, $\sigma(A)$ and $\rho(A)$ denote the spectrum and the spectral radius of A, respectively.

2. SDA and Newton's method for DAREs. In this section we first introduce a structured doubling algorithm for solving the almost stabilizing solution X_s of DARE (1.1) with the control-weighting matrix R is singular. In general, if a symmetric matrix $X \in \mathbb{R}^{n\times n}$ satisfies the DARE (1.1) and all eigenvalues of the closed-loop matrix A_F are in the closed unit disk, then we have

$$\mathcal{M}\begin{bmatrix} I \\ X \\ F \end{bmatrix} = \mathcal{L}\begin{bmatrix} I \\ X \\ F \end{bmatrix}\Phi, \tag{2.1}$$

where the matrix F is as in (1.2) and $\Phi = A_F = A + BF$ with $\rho(\Phi) \leqslant 1$. Since the control matrix $B \in \mathbb{R}^{n\times m}$ is usually of full column rank in many applications of control system theory, we can select an appropriate matrix $Y = Y^\top \in \mathbb{R}^{n\times n}$ such that $R + B^\top Y B$ is invertible. After some elementary block row operators are applied on both sides of (2.1), we obtain

$$\begin{bmatrix} (I-G_0Y)A - B\widehat{R}^{-1}C & 0 & 0 \\ -Q + C^\top\widehat{R}^{-1}(C+B^\top YA) & I & 0 \\ C + B^\top YA & 0 & \widehat{R} \end{bmatrix}\begin{bmatrix} I \\ X \\ F \end{bmatrix} = \begin{bmatrix} I - G_0Y & G_0 & 0 \\ C^\top\widehat{R}^{-1}B^\top Y & A^\top - C^\top\widehat{R}^{-1}B^\top & 0 \\ B^\top Y & -B^\top & 0 \end{bmatrix}\begin{bmatrix} I \\ X \\ F \end{bmatrix}\Phi, \tag{2.2}$$

where $\widehat{R} = R + B^\top Y B$ and $G_0 = B\widehat{R}^{-1}B^\top$, respectively.

Next, post-multiplying the second columns of the matrix pair in (2.2) by Y, and then adding them to the first columns, it follows that

$$\begin{bmatrix} (I-G_0Y)A - B\widehat{R}^{-1}C & 0 & 0 \\ \widetilde{Q} & I & 0 \\ C + B^\top YA & 0 & \widehat{R} \end{bmatrix}\begin{bmatrix} I \\ X-Y \\ F \end{bmatrix} = \begin{bmatrix} I & G_0 & 0 \\ A^\top Y & A^\top - C^\top\widehat{R}^{-1}B^\top & 0 \\ 0 & -B^\top & 0 \end{bmatrix}\begin{bmatrix} I \\ X-Y \\ F \end{bmatrix}\Phi \tag{2.3}$$

with $\widetilde{Q} = -Q + C^\top\widehat{R}^{-1}(C + B^\top YA) + Y$. Then the above matrix pair in (2.3) is pre-multiplied by the following block elementary matrix

$$\mathcal{E} = \begin{bmatrix} I & 0 & 0 \\ -A^\top Y & I & 0 \\ 0 & 0 & I \end{bmatrix},$$

we thus have

$$\begin{bmatrix} A_0 & 0 & 0 \\ -H_0 & I & 0 \\ C + B^\top YA & 0 & \widehat{R} \end{bmatrix}\begin{bmatrix} I \\ X-Y \\ F \end{bmatrix} = \begin{bmatrix} I & G_0 & 0 \\ 0 & A_0^\top & 0 \\ 0 & -B^\top & 0 \end{bmatrix}\begin{bmatrix} I \\ X-Y \\ F \end{bmatrix}\Phi \tag{2.4}$$

with

$$A_0 = (I - G_0Y)A - B\widehat{R}^{-1}C, \tag{2.5a}$$

$$G_0 = B\widehat{R}^{-1}B^\top, \tag{2.5b}$$

$$H_0 = Q - Y - C^\top\widehat{R}^{-1}(C + B^\top YA) + A^\top Y(I - G_0Y)A - A^\top YB\widehat{R}^{-1}C. \tag{2.5c}$$

Consider the matrix pair $(\mathcal{M}_0, \mathcal{L}_0)$ in standard symplectic form (SSF), where

$$\mathcal{M}_0 = \begin{bmatrix} A_0 & 0 \\ -H_0 & I \end{bmatrix}, \quad \mathcal{L}_0 = \begin{bmatrix} I & G_0 \\ 0 & A_0^\top \end{bmatrix} \tag{2.6}$$

which satisfies $\mathcal{M}_0\mathcal{J}\mathcal{M}_0^\top = \mathcal{L}_0\mathcal{J}\mathcal{L}_0^\top$, where G_0 and H_0 are symmetric matrices. By Theorem 1.1, Lemma 1.1, (**A**) and (2.4). It is obvious that the spectrum of $(\mathcal{M}_0, \mathcal{L}_0)$ is the same of $(\mathcal{M}, \mathcal{L})$ except m infinite eigenvalues, i.e., the generalized eigenvalues of $(\mathcal{M}_0, \mathcal{L}_0)$ can be arranged as

$$\underbrace{0,\ldots,0}_{r},\ \lambda_{r+1},\ldots,\lambda_\ell,\ \underbrace{\omega_1,\omega_1,\ldots\omega_{n-\ell},\omega_{n-\ell}}_{\text{unimodular ew.}},\overline{\lambda_\ell^{-1}},\ldots,\overline{\lambda_{r+1}^{-1}},\underbrace{\infty,\ldots,\infty}_{r},$$

where the eigenvalues λ_i be inside the unit circle except the origin, $i = r+1,\ldots,n$. From (2.4)–(2.5c), we immediately obtain

$$\mathcal{M}_0 \begin{bmatrix} I \\ X-Y \end{bmatrix} = \mathcal{L}_0 \begin{bmatrix} I \\ X-Y \end{bmatrix} \Phi. \tag{2.7}$$

The DARE associated with the symplectic matrix pair $(\mathcal{M}_0, \mathcal{L}_0)$ in SSF is

$$\widehat{X} = A_0^\top \widehat{X}(I + G_0\widehat{X})^{-1}A_0 + H_0. \tag{2.8}$$

on which the efficient SDA algorithm [6, 16] can be applied. Note that if $\widehat{X}$ is the symmetric solution to the above DARE (2.8), then $X = \widehat{X} + Y$ is the symmetric solution to the DARE (1.1). Using some elementary block row operations and one column operation to transform the DARE (1.1) to an equivalent DARE (2.8) with the associated symplectic matrix pair $(\mathcal{M}_0, \mathcal{L}_0)$ in SSF. As derived in [6], for given any SSF form (2.6), we construct

$$\mathcal{M}_* = \begin{bmatrix} A_0(I+G_0H_0)^{-1} & 0 \\ -A_0^\top(I+H_0G_0)^{-1}H_0 & I \end{bmatrix}, \quad \mathcal{L}_* = \begin{bmatrix} I & A_0G_0(I+H_0G_0)^{-1} \\ 0 & A_0^\top(I+H_0G_0)^{-1} \end{bmatrix} \tag{2.9}$$

and consequently deduce that

$$\mathcal{M}_*\mathcal{L}_0 = \mathcal{L}_*\mathcal{M}_0. \tag{2.10}$$

We now compute $\mathcal{L}_*\mathcal{L}_0$ and $\mathcal{M}_*\mathcal{M}_0$, and apply the Sherman-Morrison-Woodbury formula to produce

$$\widehat{\mathcal{M}} = \begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix} = \mathcal{M}_*\mathcal{M}_0, \quad \widehat{\mathcal{L}} = \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^\top \end{bmatrix} = \mathcal{L}_*\mathcal{L}_0, \tag{2.11}$$

where

$$\widehat{A} = A_0(I+G_0H_0)^{-1}A_0, \tag{2.12a}$$
$$\widehat{G} = G_0 + A_0G_0(I+H_0G_0)^{-1}A_0^\top, \tag{2.12b}$$
$$\widehat{H} = H_0 + A_0^\top(I+G_0H_0)^{-1}H_0A_0. \tag{2.12c}$$

Equations in (2.11) show that the matrix pair $(\widehat{\mathcal{M}}, \widehat{\mathcal{L}})$ is again in SSF form. From (2.10)–(2.11), the pair $(\widehat{\mathcal{M}}, \widehat{\mathcal{L}})$ satisfies the doubling property: If $\mathcal{M}_0x = \lambda\mathcal{L}_0x$, then $\widehat{\mathcal{M}}x = \lambda\mathcal{M}_*\mathcal{L}_0x = \lambda\mathcal{L}_*\mathcal{M}_0x = \lambda^2\mathcal{L}_*\mathcal{L}_0x = \lambda^2\widehat{\mathcal{L}}x$. (2.12a)–(2.12c) to make up the iterations of SDA [6, 16], the SDA then can be modified to the following algorithm for DAREs (1.1):

ALGORITHM 2.1 (SDA for DAREs).

Input: A, B, C, Q, R; τ *(a small tolerance); k=0, err=1;*
Output: *a symmetric stabilizing solution* X *to DARE (1.1).*
Select a symmetric matrix Y *such that* $\widehat{R} \equiv R + B^\top Y B$ *is invertible;*
Put $A_0 := (I - GY)A - B\widehat{R}^{-1}C,$
$G_0 := B\widehat{R}^{-1}B^\top,$
$H_0 := Q - Y - C^\top \widehat{R}^{-1}B^\top YA - A^\top YB\widehat{R}^{-1}C - C^\top \widehat{R}^{-1}C + A^\top Y(I - GY)A;$
While $err > \tau$,
Put $A_{k+1} := A_k(I + G_k H_k)^{-1}A_k^\top,$
$G_{k+1} := G_k + A_k G_k (I + H_k G_k)^{-1}A_k^\top,$
$H_{k+1} := H_k + A_k^\top (I + H_k G_k)^{-1}H_k A_k,$
$err := \|H_{k+1} - H_k\|;$
If $I + G_k H_k$ *is ill-conditioned, then break down,*
Else set $k := k + 1;$
End if
End
$X := H_k + Y.$

The Newton Method in [8, 14] is developed to solve the DARE (1.1) by solving a discrete-time Lyapunov equation (or Stein equation) at each iteration. The convergence of Newton's method is shown to either quadratic or linear with the common ratio $\frac{1}{2}$. Specifically, the Newton method can be stated as follows. Here we use the Matlab command `dlyap` to solve the Stein equation [18].

ALGORITHM 2.2 (NTM for DAREs).

Input: A, B, C, Q, R; τ *(a small tolerance); k=0, err=1;*
Output: *a symmetric stabilizing solution* X *to DARE.*
Choose L_0 *such that* $A_0 \equiv A - BL_0$ *is d-stable;*
Solve $X_0 := dlyap(A_0^\top, Q + L_0^\top RL_0 - C^\top L_0 - L_0^\top C);$
While $err > \tau$,
Put $L_{k+1} := (R + B^\top X_k B)^{-1}(C + B^\top X_k A)$ *and* $A_{k+1} := A - BL_k;$
Solve $X_{k+1} := dlyap(A_k^\top, Q + L_{k+1}^\top RL_{k+1} - C^\top L_{k+1} - L_{k+1}^\top C);$
Put $err := \|X_{k+1} - X_k\|;$
Set $k := k + 1;$
End
$X := X_k.$

2.1. Selection of Y**.** For simplicity, we choose $Y = \gamma I \in \mathbb{R}^{n\times n}$ such that $R_\gamma \equiv R + \gamma B^\top B$ is invertible for $\gamma > 0$. We first derive the forward error bounds of matrices A_0, G_0 and H_0 given in (2.5a)–(2.5c), respectively. According to these forward errors, we can design a numerical scheme to determine an appropriate value $\hat{\gamma} > 0$. In the following roundoff analysis, we use $fl(\cdot)$ to denote computed floating point values. The quantity u is the *unit roundoff* (or machine precision), which is typically of order 10^{-8} or 10^{-16} in single and double precision computer arithmetic, respectively. When A and B are $m \times n$ real matrices, the matrix $B := |A|$ if $b_{ij} = |a_{ij}|$ for all i, j, and $A \leq B$ if $a_{ij} \leq b_{ij}$ for all i, j. The 1- and ∞- matrix norms are denoted by $\|\cdot\|_1$ and $\|\cdot\|_\infty$, respectively.

Since $Y = \gamma I$, it follows from (2.5a)–(2.5c) that

$$A_0 = A - \gamma B R_\gamma^{-1} B^\top A - B R_\gamma^{-1} C, \tag{2.13}$$

$$G_0 = B R_\gamma^{-1} B^\top, \tag{2.14}$$

$$H_0 = Q - C^\top R_\gamma^{-1} C - \gamma I - \gamma C^\top R_\gamma^{-1} B^\top A - \gamma A^\top B R_\gamma^{-1} C + \gamma A^\top A - \gamma^2 A^\top B R_\gamma^{-1} B^\top A. \tag{2.15}$$

Since $R_\gamma \in \mathbb{R}^{m\times m}$ is symmetric indefinite, the matrix $W_B \equiv R_\gamma^{-1} B^\top$ can be computed by block LDL$^\top$ factorization with any pivoting strategy, for instance, the Bunch-Kaufman partial pivoting strategy, see e.g., [10, Chapter 11]. Suppose this algorithm yields the computed factorization $P R_\gamma P^\top \approx \widehat{L}\widehat{D}\widehat{L}^\top$, where P is a permutation matrix and $\widehat{D}$ has diagonal blocks of dimension 1 or 2. From Theorem 11.3 of [10], we obtain

$$\begin{aligned} fl(W_B) &= W_B + E_1, \\ |E_1| &\le p_1(m) u \left[|R_\gamma^{-1}|(|R_\gamma| + P^\top |\widehat{L}||\widehat{D}||\widehat{L}^\top| P)|fl(W_B)| \right] + O(u^2), \end{aligned} \tag{2.16}$$

where $p_1(m)$ is a linear polynomial. Since it can be shown that the matrix $|\widehat{L}||\widehat{D}||\widehat{L}^\top|$ satisfies the bound [9]

$$\| \, |\widehat{L}||\widehat{D}||\widehat{L}^\top| \, \|_M \le 36 m \rho_m \|R_\gamma\|_M, \tag{2.17}$$

where $\|R_\gamma\|_M \equiv \max_{i,j} |(R_\gamma)_{ij}|$ and ρ_m is the growth factor.

Therefore, it follows from (2.16) and (2.17) that the forward error E_1 satisfies

$$\begin{aligned} fl(W_B) &= W_B + E_1, \\ \|E_1\|_\infty &\le p_1(m) u \left[\|R_\gamma^{-1}\|_\infty \|R_\gamma\|_\infty \|fl(W_B)\|_\infty + \|R_\gamma^{-1}\|_\infty \| \, |\widehat{L}||\widehat{D}||\widehat{L}^\top| \, \|_\infty \|fl(W_B)\|_\infty \right] + O(u^2) \\ &\le p_1(m) u \left[\|R_\gamma^{-1}\|_\infty \|R_\gamma\|_\infty \|fl(W_B)\|_\infty + m\|R_\gamma^{-1}\|_\infty \| \, |\widehat{L}||\widehat{D}||\widehat{L}^\top| \, \|_M \|fl(W_B)\|_\infty \right] + O(u^2) \\ &\le p_1(m)(1 + 36 m^2 \rho_m) \rho_m u (\|R_\gamma^{-1}\|_\infty \|R_\gamma\|_\infty \|fl(W_B)\|_\infty) + O(u^2) \\ &\le p(m) \rho_m u (\|R_\gamma^{-1}\|_\infty \|R_\gamma\|_\infty \|fl(W_B)\|_\infty) + O(u^2), \end{aligned} \tag{2.18}$$

where $p(m)$ is a cubic polynomial. Similarly, the forward error bound in evaluating $W_C \equiv R_\gamma^{-1} C$ is given by

$$\begin{aligned} fl(W_C) &= W_C + E_2, \\ \|E_2\|_\infty &\le p(m) \rho_m u (\|R_\gamma^{-1}\|_\infty \|R_\gamma\|_\infty \|fl(W_C)\|_\infty) + O(u^2). \end{aligned} \tag{2.19}$$

Furthermore, it can be derived from (2.18) and (2.19) that

$$\begin{aligned} fl(\gamma B R_\gamma^{-1} B^\top A) &= \gamma B R_\gamma^{-1} B^\top A + E_3, \\ \|E_3\|_\infty &\leqslant p(m) \rho_m u (\gamma \|B\|_\infty \|R_\gamma^{-1}\|_\infty \|R_\gamma\|_\infty \|fl(W_B)\|_\infty) + O(u^2), \end{aligned} \tag{2.20}$$

and

$$\begin{aligned} fl(B R_\gamma^{-1} C) &= B R_\gamma^{-1} C + E_4, \\ \|E_4\|_\infty &\leqslant p(m) \rho_m u (\|B\|_\infty \|R_\gamma^{-1}\|_\infty \|R_\gamma\|_\infty \|fl(W_C)\|_\infty) \\ &\quad + m u (\|B\|_\infty \|R_\gamma^{-1}\|_\infty \|C\|_\infty) + O(u^2). \end{aligned} \tag{2.21}$$

Therefore, from (2.18), (2.20) and (2.21), we can deduce that the forward error bounds in evaluating A_0 and G_0 in (2.13)–(2.14) are

$$\begin{aligned} fl(A_0) &= A_0 + E_5, \\ \|E_5\|_\infty &\leqslant p(m)\rho_m u\,[\,(\gamma\|A\|_\infty + 1)\|B\|_\infty \kappa_\infty(R_\gamma)\|fl(W_B)\|_\infty\,] \\ &\quad + (2n+3)u(\gamma\|B\|_\infty\|R_\gamma^{-1}\|_\infty\|B\|_1\|A\|_\infty) + (m+1)u(\|B\|_\infty\|R_\gamma^{-1}\|_\infty\|C\|_\infty) \\ &\quad + 2u\|A\|_\infty + O(u^2), \end{aligned} \tag{2.22}$$

and

$$\begin{aligned} fl(G_0) &= G_0 + E_6, \\ \|E_6\|_\infty &\leqslant p(m)\rho_m u(\|B\|_\infty \kappa_\infty(R_\gamma)\|fl(W_B)\|_\infty) \\ &\quad + mu(\|B\|_\infty\|R_\gamma^{-1}\|_\infty\|B\|_1) + O(u^2), \end{aligned} \tag{2.23}$$

where $\kappa_\infty(R_\gamma) = \|R_\gamma^{-1}\|_\infty\|R_\gamma\|_\infty$ is the condition number of R_γ.

Finally, from (2.18) and (2.19), the forward error bound in evaluating the matrix H_0 in 2.15 is given by

$$\begin{aligned} fl(H_0) &= H_0 + E_7, \\ \|E_7\|_\infty &\leqslant p(m)\rho_m u(\,\|C\|_1\kappa_\infty(R_\gamma)\|fl(W_C)\|_\infty + \gamma\|C\|_1\kappa_\infty(R_\gamma)\|fl(W_B)\|_\infty\|A\|_\infty \\ &\quad + \gamma\|A\|_1\|B\|_\infty\kappa_\infty(R_\gamma)\|fl(W_C)\|_\infty + \gamma^2\|A\|_1\|B\|_\infty\kappa_\infty(R_\gamma)\|fl(W_B)\|_\infty\|A\|_\infty\,) \\ &\quad + (m+6)u(\|C\|_1\|R_\gamma^{-1}\|_\infty\|C\|_\infty) + 6u(\|Q\|_\infty + \gamma) + (n+3)u(\gamma\|A\|_1\|A\|_\infty) \\ &\quad + (n+m+5)u(\gamma\|C\|_1\|R_\gamma^{-1}\|_\infty\|B\|_1\|A\|_\infty) \\ &\quad + (n+m+4)u(\gamma\|A\|_1\|B\|_\infty\|R_\gamma^{-1}\|_\infty\|C\|_\infty) \\ &\quad + (3n+2)u(\gamma^2\|A\|_1\|B\|_\infty\|R_\gamma^{-1}\|_\infty\|B\|_1\|A\|_\infty) + O(u^2). \end{aligned} \tag{2.24}$$

In order to control the forward error bounds of A_0, G_0 and H_0, and the conditioning of $I + G_0H_0$, we consider the following min-max optimization problem, to determine an optimal value $\hat{\gamma} > 0$:

$$\min_{\gamma>0} F(\gamma) \equiv \max_{i=1,2,3}\{f_i(\gamma)\}, \tag{2.25}$$

where the functions $f_1(\gamma) = \kappa_\infty(R_\gamma)$, $f_2(\gamma) = \gamma^2\kappa_\infty(R_\gamma)$ and $f_3(\gamma) = \mathrm{cond}(I + G_0H_0)$, respectively.

Since the condition number $\kappa_\infty(R_\gamma)$ is bounded as $\gamma \to \infty$, it follows that $F(\gamma)$ becomes unbounded as $\gamma \to \infty$. Extensive numerical experiments on randomly generated matrices indicate that $F(\gamma)$ is a strictly convex function in the neighborhood of the optimal $\hat{\gamma}$ where the global minimum of $F(\gamma)$ occurs. For illustration, we report a sample of graphs of $F(\gamma)$ in Figure 2.1.

We can apply the Fibonacci search method to compute an approximate value of $\hat{\gamma}$, see, e.g., [3, p. 272]. Our experience indicates that three to five iterations of Fibonacci search are adequate to obtain a suboptimal yet acceptable approximation to $\hat{\gamma}$.

3. Convergence of SDA. In [13], necessary and sufficient conditions are given for the existence of the stabilizing solution $X_s = X_s^\top \in \mathbb{R}^{n\times n}$ to DARE (1.1).

Definition 3.1. *[13] A regular ESP is called disconjugate if it has no generalized eigenvalues on the unit circle and V_1 is invertible in (1.8).*

It has been proven in [13] that the DARE (1.1) has a unique stabilizing solution X_s if and only if the ESP is disconjugate. In this section we shall first characterize the quadratic convergence

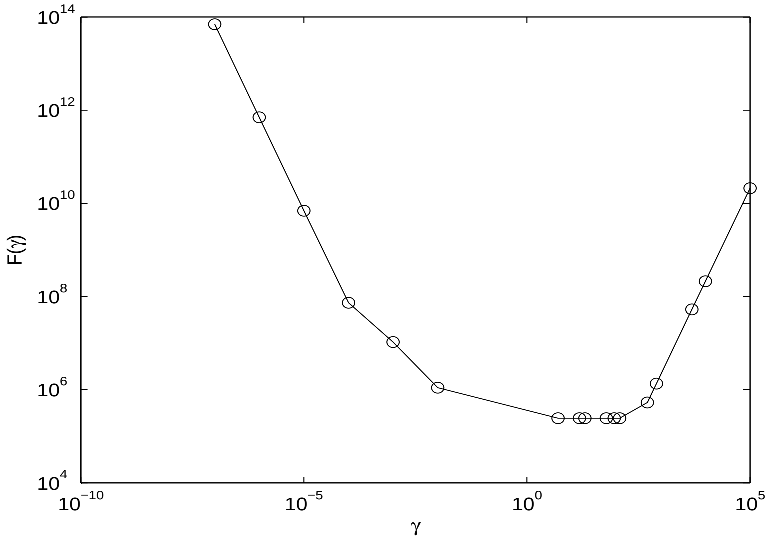

FIG. 2.1. *The graph of* $F(\gamma)$.

of SDA under the same conditions. With the appropriate selection of Y, it is easily seen from (2.7) that the standard symplectic pencil $\mathcal{M}_0 = \lambda\mathcal{L}_0$ in (2.6) is also disconjugate if the ESP is disconjugate. Therefore, we deduce that the simplified DARE (2.8) has a unique stabilizing solution $\widehat{X}_s$. For simplicity, we only consider complex matrices in the following convergence theorems. The proofs for real symplectic pencils can be modified slightly from the complex cases. Suppose that there exist nonsingular U, W such that

$$U\mathcal{M}_0W = \begin{bmatrix} J_s & 0 \\ 0 & I \end{bmatrix}, \quad U\mathcal{L}_0W = \begin{bmatrix} I & 0 \\ 0 & J_s \end{bmatrix}, \tag{3.1}$$

where $\rho(J_s) < 1$. If we denote

$$W = \begin{bmatrix} W_1 & W_3 \\ W_2 & W_4 \end{bmatrix}, \tag{3.2}$$

where $W_i \in \mathbb{C}^{n\times n}$ for $i = 1, 2, 3, 4$, the quadratic convergence of the SDA algorithm can be tackled from Theorem 1 of [6].

THEOREM 3.1. *Suppose that the ESP is disconjugate. If W_1 and W_4 in (3.2) are invertible, then the sequences $\{A_k,\ H_k,\ G_k\}$ computed by the SDA algorithm satisfy:*

(i) $\|A_k\| = O(\|J_s^{2^k}\|) \to 0$ *as* $k \to \infty$,

(ii) $H_k \to \widehat{X}_s$, *where* $\widehat{X}_s$ *is a stabilizing solution of DARE (2.8)*

$$\widehat{X} = A_0^\top \widehat{X}(I + G_0\widehat{X})^{-1}A_0 + H_0,$$

(iii) $G_k \to \widehat{X}_d$, *where* $\widehat{X}_d$ *solves the dual DARE*

$$\widehat{Y} = A_0\widehat{Y}(I + H_0\widehat{Y})^{-1}A_0^\top + G_0.$$

Moreover, the convergence rate in (i)–(iii) above is $O\left(|\lambda_n|^{2^k}\right)$, *where* $|\lambda_1| \leqslant \cdots \leqslant |\lambda_n| < 1 < |\lambda_n|^{-1} \leqslant \cdots \leqslant |\lambda_1|^{-1}$ *with* $\lambda_i,\ \lambda_i^{-1}$ *being the eigenvalues of* $\mathcal{M}_0 - \lambda\mathcal{L}_0$ *(including* 0 *and* ∞*).*

REMARK 3.1. Under the same conditions of Theorem 3.1 and $\widehat{X}_s$ is the unique stabilizing solution of DARE (2.8), it follows that the symmetric matrix $X_s = \widehat{X}_s + Y$ must be the unique maximal, stabilizing solution of DARE (1.1).

On the other hand, when the ESP satisfies the condition **(A)**, we shall prove the linear convergence of SDA with ratio less than $\frac{1}{2}$. Denote the Jordan block of size p corresponding to a unimodular eigenvalue $\omega \equiv e^{\imath\theta}$ by

$$J_{\omega,p} = \begin{bmatrix} \omega & 1 & 0 & \cdots & 0 \\ 0 & \omega & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \omega \end{bmatrix}_{p\times p}. \tag{3.3}$$

In particular, for the unimodular eigenvalues $\omega_j = e^{\imath\theta_j}$ of the matrix pair $(\mathcal{M}_0, \mathcal{L}_0)$ with $p = 1$, we have

$$J_{\omega_j,1} = [\omega_j] \tag{3.4}$$

for $j = 1, \ldots, n-\ell$. From symplectic Kronecker's Theorem for $(\mathcal{M}, \mathcal{L})$ (see [15]), there exist a symplectic matrix $\mathcal{Z}$ (i.e., $\mathcal{Z}^\top \mathcal{J} \mathcal{Z} = \mathcal{J}$) and a nonsingular $\mathcal{Q}$ such that

$$\mathcal{Q}\mathcal{M}_0\mathcal{Z} = \begin{bmatrix} J_s \oplus J_1 & 0_\ell \oplus \Gamma_1 \\ 0_n & I_\ell \oplus J_1^{-H} \end{bmatrix} \equiv \mathcal{J}_{\mathcal{M}}, \tag{3.5a}$$

$$\mathcal{Q}\mathcal{L}_0\mathcal{Z} = \begin{bmatrix} I_\ell \oplus I_\mu & 0_n \\ 0_n & J_s^H \oplus I_\mu \end{bmatrix} \equiv \mathcal{J}_{\mathcal{L}}, \tag{3.5b}$$

where J_s is the stable Jordan block of size ℓ (i.e., $\rho(J_s) < 1$),

$$J_1 = J_{\omega_1,1} \oplus \cdots \oplus J_{\omega_\mu,1}, \tag{3.6}$$

$$\Gamma_1 = I_\mu, \tag{3.7}$$

with $\mu = n - \ell$ and $\oplus$ denotes the direct sum of matrices. Note that $J_1^{-H} = J_1$, and the matrices $\mathcal{J}_{\mathcal{M}}$ and $\mathcal{J}_{\mathcal{L}}$ in (3.5) commute with each other. Therefore, from (3.5), one can derive

$$\mathcal{M}_0\mathcal{Z}\mathcal{J}_{\mathcal{L}} = \mathcal{Q}^{-1}\mathcal{J}_{\mathcal{L}}\mathcal{J}_{\mathcal{M}} = \mathcal{L}_0\mathcal{Z}\mathcal{J}_{\mathcal{M}}. \tag{3.8}$$

From (3.5) it follows that span$\{\mathcal{Z}(:, 1:n)\}$ forms the unique stable Lagrangian deflating subspace of $(\mathcal{M}, \mathcal{L})$ corresponding to $J_s \oplus J_1$. Let $\{(\mathcal{M}_k, \mathcal{L}_k)\}_{k=0}^\infty$ be the sequence of symplectic pairs generated by Algorithm 2.1, or $\{(\mathcal{M}_k, \mathcal{L}_k)\}_{k=0}^\infty$ be the sequence of symplectic pairs in SSF with

$$\mathcal{M}_k = \begin{bmatrix} A_k & 0 \\ -H_k & I \end{bmatrix}, \quad \mathcal{L}_k = \begin{bmatrix} I & G_k \\ 0 & A_k^\top \end{bmatrix} \tag{3.9}$$

generated by Algorithm 2.1. It follows from (3.8) as well as (2.10)–(2.11) that

$$\begin{aligned} \mathcal{M}_1\mathcal{Z}\mathcal{J}_{\mathcal{L}}^2 &= \mathcal{M}_*\mathcal{M}_0\mathcal{Z}\mathcal{J}_{\mathcal{L}}^2 = \mathcal{M}_*\mathcal{L}_0\mathcal{Z}\mathcal{J}_{\mathcal{M}}\mathcal{J}_{\mathcal{L}} = \mathcal{L}_*\mathcal{M}_0\mathcal{Z}\mathcal{J}_{\mathcal{L}}\mathcal{J}_{\mathcal{M}} \\ &= \mathcal{L}_*\mathcal{L}_0\mathcal{Z}\mathcal{J}_{\mathcal{M}}^2 = \mathcal{L}_1\mathcal{Z}\mathcal{J}_{\mathcal{M}}^2. \end{aligned} \tag{3.10}$$

Inductively, we have

$$\mathcal{M}_k\mathcal{Z}\mathcal{J}_{\mathcal{L}}^{2^k} = \mathcal{L}_k\mathcal{Z}\mathcal{J}_{\mathcal{M}}^{2^k} \tag{3.11}$$

for $k = 1, 2, \ldots$. From the definitions of $\mathcal{J}_{\mathcal{M}}$ and $\mathcal{J}_{\mathcal{L}}$ in (3.5a) and (3.5b), respectively, it can be deduced that (3.11) can be rewritten as

$$\mathcal{M}_k \mathcal{Z} \begin{bmatrix} I_n & 0_n \\ 0_n & (J_s^H)^{2^k} \oplus I_\mu \end{bmatrix} = \mathcal{L}_k \mathcal{Z} \begin{bmatrix} J_s^{2^k} \oplus J_1^{2^k} & 0_\ell \oplus \Gamma_k \\ 0_n & I_\ell \oplus J_1^{2^k} \end{bmatrix}, \tag{3.12}$$

where

$$\Gamma_k = 2^{k-1} J_1^{2^{k-1}-1} \tag{3.13}$$

for $k = 1, 2, \ldots$. From (3.13), we immediately obtain the following Lemma.

Lemma 3.1. *Let J_1 and Γ_k be defined in* (3.6) *and* (3.13), *respectively. Then Γ_k is invertible and satisfies*

$$\|\Gamma_k^{-1}\| = O(2^{-k}), \quad \|J_1^{2^k}\| = O(1). \tag{3.14}$$

Based on the Lemma 3.1, we shall prove the linear convergence of SDA under the condition **(A)**. We now partition $\mathcal{Z}$ in (3.8) by

$$\mathcal{Z} = \begin{bmatrix} Z_1 & Z_3 \\ Z_2 & Z_4 \end{bmatrix}, \tag{3.15}$$

where $Z_i \in \mathbb{C}^{n \times n}$, for $i = 1, 2, 3, 4$.

Theorem 3.2. *Suppose that the ESP in* (1.5) *satisfies the condition* **(A)** *and the DARE (1.1) has an almost stabilizing solution X_s with property* **(P)**. *Let $Z_{2b} = Z_2(:, 1:\mu)$ and $Z_{4a} = Z_4(1:\ell, :)$. If the matrix $[Z_{4a}\ Z_{2b}]$ is invertible, then Z_1 is invertible, $\widehat{X_s} = Z_2 Z_1^{-1}$ is an almost stabilizing solution of DARE* (2.8), *and the sequences $\{A_k, G_k, H_k\}$ generated by Algorithm 2.1 satisfy*

(1) $\limsup_{k\to\infty} \sqrt[k]{\|A_k\|} \le \frac{1}{2}$.

(2) $\limsup_{k\to\infty} \dfrac{\|H_{k+1} - \widehat{X_s}\|}{\|H_k - \widehat{X_s}\|} \le \frac{1}{2}$, *i.e., $H_k \to \widehat{X_s}$ linearly with rate less than $\frac{1}{2}$. Moreover, $X_s = \widehat{X_s} + Y$.*

Proof. By **(A)** and assumptions of Lemma 1.1, the space span$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ and span$\begin{bmatrix} I \\ X \end{bmatrix}$ forms the same unique stable Lagrangian of $(\mathcal{M}_0, \mathcal{L}_0)$ corresponding to $J_s \oplus J_1$, then Z_1 is invertible, and $\widehat{X_s} = Z_2 Z_1^{-1}$ solves the DARE (2.8)[19].

On the other hand, substituting $(\mathcal{M}_k, \mathcal{L}_k)$ of (3.9) and $\mathcal{Z}$ of (3.15) into (3.12), and comparing both sides we obtain

$$A_k Z_1 = (Z_1 + G_k Z_2)\left(J_s^{2^k} \oplus J_1^{2^k}\right), \tag{3.16a}$$

$$\begin{aligned} A_k Z_3 \left((J_s^H)^{2^k} \oplus I_\mu\right) = {} & (Z_1 + G_k Z_2)(0_\ell \oplus \Gamma_k) \\ & + (Z_3 + G_k Z_4)\left(I_\ell \oplus J_1^{2^k}\right), \end{aligned} \tag{3.16b}$$

$$-H_k Z_1 + Z_2 = A_k^\top Z_2 \left(J_s^{2^k} \oplus J_1^{2^k}\right), \tag{3.16c}$$

$$(-H_k Z_3 + Z_4)\left((J_s^H)^{2^k} \oplus I_\mu\right) = A_k^\top Z_2 (0_\ell \oplus \Gamma_k) + A_k^\top Z_4 \left(I_\ell \oplus J_1^{2^k}\right). \tag{3.16d}$$

Substituting (3.16c) into (3.16d), we have

$$(-\widehat{X_S}Z_3+A_k^\top Z_2(J_s^{2^k}\oplus J_1^{2^k})Z_3+Z_4)((J_s^H)^{2^k}\oplus I_\mu) = A_k^\top Z_2\,(0_\ell\oplus\Gamma_k)+A_k^\top Z_4\left(I_\ell\oplus J_1^{2^k}\right). \quad (3.17)$$

Postmultiplying (3.17) by $(I_\ell\oplus\Gamma_k^{-1})$, we get

$$\begin{aligned} &{A_k}^\top\left[Z_2(0_\ell\oplus I_\mu)+Z_4(I_\ell\oplus J_1^{2^k}\Gamma_k^{-1})-Z_2(J_s^{2^k}\oplus J_1^{2^k}))Z_3((J_s^H)^{2^k}\oplus I_\mu)\Gamma_k^{-1}\right]\\ &=(-\widehat{X_S}Z_3+Z_4)((J_s^H)^{2^k}\oplus\Gamma_k^{-1}). \end{aligned} \quad (3.18)$$

By Lemma 3.1 and assumptions in Theorem 3.2, for sufficient large k, the matrix $Z_2\,(0_\ell\oplus I_\mu)+Z_4\left(I_\ell\oplus J_1^{2^k}\Gamma_k^{-1}\right)$ would be invertible. Then the sequence $\{A_k\}$ satisfies

$$\limsup_{k\to\infty}\sqrt[k]{\|A_k\|}\le\limsup_{k\to\infty}\sqrt[k]{O(1)2^{-k}}=\frac{1}{2} \quad (3.19)$$

From (3.16c), we get

$$\limsup_{k\to\infty}\frac{\|H_{k+1}-\widehat{X_S}\|}{\|H_k-\widehat{X_S}\|}\le\limsup_{k\to\infty}\sqrt[k]{\|H_k-\widehat{X_S}\|}=\limsup_{k\to\infty}\sqrt[k]{\|A_k^\top Z_2(J_s^{2^k}\oplus J_1^{2^k})\|}\le\frac{1}{2}.$$

■

COROLLARY 3.2. *Assume that (A,B) is d-stabilizable and that the same conditions as in Theorem 3.2 hold. If the DARE* (1.1) *has a maximal solution X_+, then it must coincide with the almost stabilizing solution X_s computed by SDA.*

Proof. Form Theorem 1.2, it can be seen that the maximal solution X_+ satisfies $R+B^\top X_+B>0$ and $\rho(A_F)\le 1$. In addition, since (A,B) is d-stabilizable, the assumptions of Lemma 1.1 can be guaranteed. Therefore, the subspaces span $\begin{bmatrix} I \\ X_+-Y\end{bmatrix}$ and span $\begin{bmatrix} I \\ \widehat{X_s}\end{bmatrix}$ are unique stable Lagrangian subspaces of matrix pair $(\mathcal{M}_0,\mathcal{L}_0)$ corresponding to the spectrum of $J_s\oplus J_1$. Hence, this completes the proof. ■

REMARK 3.2. If the d-stabilizability of the pair (A,B) is replaced by a weaker condition (1.7), then the conclusion of Corollary 3.2 still holds.

4. Numerical examples. The aim of this section is to illustrate the superior performance of the SDA algorithm, as compared to the Newton's Method [8]. The flop counts for each iteration in SDA and NTM is $\frac{23}{3}n^3$ and $30n^3$, respectively. We test some numerical examples satisfying hypothesis (**A**) by SDA and Newton's methods NTM. The convergence of NTM is guaranteed under the same conditions, and the rate of convergence is linear with ratio $\frac{1}{2}$.

Note that NTM can be used to solve DARE with a more general case when R is singular, but we must assume that there is a symmetric solution X_+ of the inequality $\mathcal{R}(X_+)\ge 0$ for which $R+B^\top X_+B>0$. In SDA, G and H are only required to be symmetric. In this release condition, the existence of sequence $\{A_k,G_k,H_k\}$ should be guarantee. As mentioned before, the approximate solution computed by SDA algorithm is an almost stabilizing solution X_s of DARE (1.1) when all eigenvalues of A_F are in closed unit disk while Newton's method converges to the maximal solution X_+ of DARE (1.1). In Corollary 3.2, we prove that these two solutions are coincident which is observed in our numerical experiments.

We reports the numbers of iterations by "ITs", the CPU time by "CPU" for two algorithms, and the "Err" in SDA and NTM is defined by $\|H_{k+1}-H_k\|$ and $\|X_{k+1}-X_k\|$, respectively. We list five examples in this section. Example 4.1, 4.2 and 4.3 are identical to examples of Benner

	SDA	NTM
DNRes	2.89×10^{-16}	1.6×10^{-16}
ITs	6	6
CPU	0.016	0.12
Err	1.1×10^{-10}	1.4×10^{-10}

TABLE 4.1
Results for Example 4.1.

et. al. (1999) which was presented originally in Ionescu (1992), Guo (1998), Sun (1998) et. al.. The forth example has been observed the convergence rate of SDA when the close loop matrix A_F has eigenvalues on the unit circle and semisimple. In the last example , we are increasing the dimension n of these matrices, and list the CPU time ratio of SDA and NTM with n increases.

For the residual of DARE, we use the "normalized" residual (DNRes) formula

$$\text{DNRes} \equiv \frac{\| - \widetilde{X} + A^\top \widetilde{X} A + Q - (C + B^\top \widetilde{X} A)^\top (R + B^\top \widetilde{X} B)^{-1} (C + B^\top \widetilde{X} A)\|}{\|\widetilde{X}\| + \|A^\top \widetilde{X} A\| + \|Q\| + \|(C + B^\top \widetilde{X} A)^\top (R + B^\top \widetilde{X} B)^{-1} (C + B^\top \widetilde{X} A)\|}.$$

proposed in [4], where $\widetilde{X}$ is an approximate solution to DARE.

All computations were performed in MATLAB/version 7.0 on a PC with a Intel Pentium-IV 3.2 GHZ processor and 2.5 GB main memory, using IEEE double-precision.

EXAMPLE 4.1. [13] For the following numerical data with singular A and R

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 1 \\ -1 & 7 \end{bmatrix},$$

$$Q = \begin{bmatrix} -\frac{4}{11} & -\frac{4}{11} \\ -\frac{4}{11} & \frac{7}{11} \end{bmatrix}, \quad R = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}.$$

Note that the rank of R is 1, i.e., R is singular. The symplectic pencil $(\mathcal{M}, \mathcal{L})$ has no eigenvalue on the unit circle. Since at each NTM iteration, algorithm 2.2 must be solve a Stein equation which is expensive [2]. From Table 4.1 we see that SDA and NTM are both converge quadratically and the CPU time of SDA is fast than of NTM.

EXAMPLE 4.2. [8] We consider the DARE with $n = m = 2$ defined by

$$A = \begin{bmatrix} 0 & -1 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad C = 0, \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}.$$

We note that R is singular, and the symplectic pencil $(\mathcal{M}, \mathcal{L})$ has eigenvalues $\{0, 1, 1, \infty, \infty, \infty\}$. It can be easily seen $X = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ which is the only solution of the DARE. The close-loop $A + BF$ has eigenvalues $\{0, 1\}$ and the elementary divisors of $\mathcal{M} - \lambda\mathcal{L}$ corresponding to the eigenvalue $\{1\}$ are of degree two. We can see that the convergence of SDA and NTM are both linear with rate $\frac{1}{2}$. The numerical results are recorded in Table 4.2.

EXAMPLE 4.3. [21] Consider the DARE (1.1) with

$$A = \begin{bmatrix} 0 & 10^{-1} & 0 \\ 0 & 0 & 10^{-1} \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},$$

	SDA	NTM
DNRes	1.2×10^{-16}	1.6×10^{-16}
ITs	24	26
CPU	0.031	0.14
Err	3.0×10^{-8}	1.4×10^{-8}

TABLE 4.2
Results for Example 4.2.

	SDA	NTM
DNRes	4.6×10^{-16}	3.9×10^{-16}
ITs	2	2
CPU	0.031	0.047
Err	0	0

TABLE 4.3
Results for Example 4.3.

and

$$Q = \begin{bmatrix} 10^5 & 0 & 0 \\ 0 & 10^3 & 0 \\ 0 & 0 & -10 \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = 0.$$

The exact solution of the DARE is $X = \mathrm{diag}(10^5, 10^3, 0)$. Since the symplectic pair $(\mathcal{M}, \mathcal{L})$ has no eigenvalues on the unit circle. Two methods are both converge quadratically. Here, we choose the initial matrix $L_0 = 0$ so that $A - BL_0 = A$ is d-stable. From table 4.3, it follows that the normalized residuals for X by SDA and NTM have both 16 significant digits.

EXAMPLE 4.4. Let r is a arbitrary number (using r = rand), and $R = \begin{bmatrix} 1 & r \end{bmatrix} \begin{bmatrix} 1 \\ r \end{bmatrix}$, $A = \begin{bmatrix} 2 + r^2 & 0 \\ 0 & 0 \end{bmatrix}$, $B = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $C = 0$, and $Q = I_2 - A^\top A + (C + B^\top XA)^\top (R + B^\top B)^{-1}(C + B^\top XA)$. It can be easily verified that almost stabilizing solution $X = I_2$, the matrix $A + BF = \begin{bmatrix} 1 & 0 \\ r & 0 \end{bmatrix}$ has eigenvalues $1, 0$, i.e., the assumption (**A**) is to be guaranteed. To using Newton's method, it needs to choose an initial matrix L_0 such that $A - BL_0$ is d-stable. It is easy to check $L_0 = A$ that a desire answer. We list absolute error and DNRes of after 20th iteration in SDA and NTM algorithms, from table 4.4, the convergence rate is linear with ratio $\frac{1}{2}$ can be observed.

EXAMPLE 4.5. In this example, we run the algorithms on some randomly generated examples with the dimension n varying from 50 to 300. We will construct $n \times n$ matrices A, B, C, Q and R such that the spectrum of $A + BF$ are all lies on the unit circle. In the first place, let U be an randomly unitary matrix and $A = 2U$. X is a symmetry positive definite matrix, and R is a symmetric positive semidefinite matrix with one eigenvalue 0, $n - 1$ eigenvalues between 0 and 1. Let $B = X^{-\frac{1}{2}}\mathrm{chol}(I - R)$, $C = \frac{1}{2}B^{-1}A - B^\top XA$, $Q = X - A^\top XA - A^\top B^{-2}A$. We check at once that the matrix $A + BF$ and the unitary matrix U is identical, which is clear from the design.

In figure 4.1 , we report the comparison of CPU times with respect to the SDA and NTM for $n = 50, 100, 150, 200, 250, 300$. We also list the normalized residuals (NRes) in table 4.5. From table 4.5 we see that the residuals of SDA are smaller than those of NTM up to 1-2 digits for all

ITs	Err(SDA)	Err(NTM)
20	1.26e-6	1.68e-6
21	6.29e-7	6.21e-7
22	3.15e-7	3.16e-7
23	1.58e-7	1.64e-7
24	8.01e-8	8.07e-8

TABLE 4.4
Results for Example 4.4.

n, which means that SDA computes more accurate solutions than NTM, generally. In figure 4.1, we see that the cpu time of SDA is about 10% to 30% of that of NTM.

TABLE 4.5
Results for Example 4.5.

Methods		NTM	SDA	Methods		NTM	SDA
$n = 50$	IT	17	19	n=200	IT	18	20
	CPU	0.41	0.17		CPU	14	2.1
	NRes	4.6e-12	2.3e-13		NRes	5.2e-14	9.6-14
$n = 100$	IT	18	19	n=250	IT	17	21
	CPU	2.3	0.33		CPU	29	4.9
	NRes	4.3e-12	6.1e-13		NRes	4.1e-12	5.6e-14
$n = 150$	IT	18	20	n=300	IT	17	20
	CPU	6.8	1.1		CPU	57	6.4
	NRes	5.4e-12	1.6e-13		NRes	2.3e-12	7.8e-14

5. Concluding remarks. In this paper, we propose the structured doubling algorithm for solving the stabilizing or almost stabilizing solution of DARE 1.1. In Theorem 3.1 and 3.2, we prove quadratic and global linear with ratio $\frac{1}{2}$ convergence for SDA algorithms, respectively. The convergence behavior is similar to that of Newton's method. We prove in Corollary 3.2 that the almost stabilizing solution computed by SDA is the same as the maximal solution computed by Newton's method. However, in each Newton's iteration, a Stein equation must be solved, which is rather expensive. Numerical examples show that our structured doubling algorithm is efficient, out-performing Newton's method.

Acknowledgments. We would like to thank Professor Eric King-Wah Chu from Monash University for many helpful suggestions and valuable comments.

REFERENCES

[1] Z. BAI, J. DEMMEL AND M. GU, *An inverse free oarallel spectral divide and conquer algorithm for nonsymmetric eigenproblem*, Numer. Math. 76(3) (1997), pp. 279–308.
[2] R. H. BARTELS AND G. W. STEWART, *Solution of the matrix equation $AX + XB = C$*, Comm. ACM., 15 (1972), pp. 820–826.
[3] M. S. BAZARAA, H. D. SHERALL AND C. M. SHETTY, *Nonlinear Programming*, John Wiley, 1993.
[4] P. BENNER AND R. BYERS, *Evaluating Products of matrix pencils and collapsing matrix products*, Num. Lin. Alg. Appl., 8 (2001), pp. 357–380.
[5] P. BENNER, A. J. LAUB AND V. MEHRMANN, *A collection of benchmark examples for the numerical solution of algebraic Riccati equations II: discrete-time case*, Techical Report SPC 95-23, Fakultät für Mathematik, Technische Universität Chemnite-Zwickau, FRG, 1995.

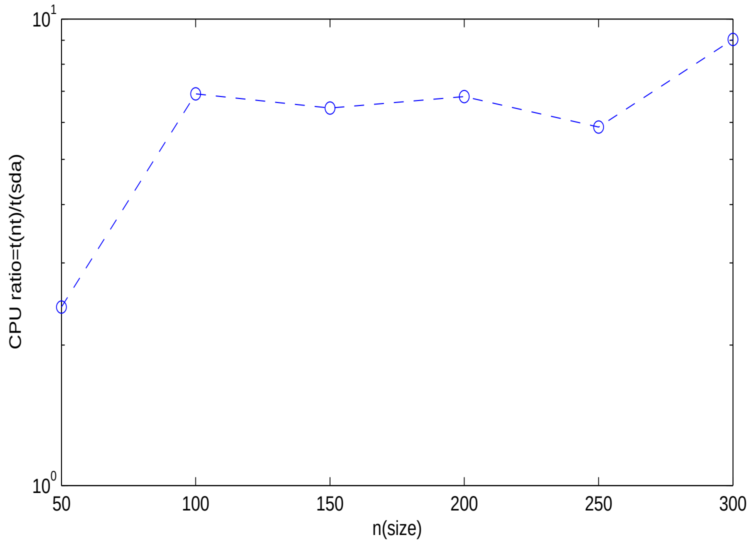

Fig. 4.1. *The comparison of CPU time of Example 4.5.*

[6] E. K.-W. Chu, H.-Y. Fan and W.-W. Lin, *Structure-preserving algorithms for periodic discrete-time algebraic Riccati equations*, Int. J. Control, 77 (2004), pp. 767–788.

[7] A. Emami-Naeini and G. Franklin, *Comments on the numerical solution of the discrete-time algebraic Riccati equation*, IEEE Trans. Automat. Control, AC-25 (1980), pp. 1015–1016.

[8] C.-H. Guo, *Newton's method for discrete algebraic Riccati equations with the colsed-loop matrix has eigenvalues on the unit circle*, SIAM J. Matrix Anal. Appl. 20 (1998), pp. 279–294.

[9] N. J. Higham, *Stability of the diagonal pivoting method with partial pivoting*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 52–65.

[10] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, 2002.

[11] T.-W. Hwang, E. K.-W. Chu, and W.-W. Lin, *A generalized structure-preserving doubling algorithm for generalized discrete-time algebraic Riccati equations*, Int. J. Control 78 (2005), pp. 1063–1075

[12] T.-W. Hwang and W.-W. Lin , *Structure-preserving algorithms for Hermitian solutions of algebraic Riccati equations*, Math. Comp., revised.

[13] V. Ionescu and M. Weiss, *On comupting the stabilizing solution of the discrete-time Riccati equation*, Lin. Alg. Appl., 174 (1992), pp. 229–238.

[14] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995.

[15] W.-W. Lin, V. Mehrmann and H. Xu, *Canonical forms for Hamiltonian and Sympletic matrices and pencils*, Lin. Alg. Appl., 302-303 (1999), pp. 469–533.

[16] W.-W. Lin and S.-F. Xu, *Convergence analysis of structure-preserving doubling algorithms for Riccati-type matrix equations*, SIAM Matrix Anal. Appl., 28(1) (2006), pp. 26-39.

[17] A. N. Malyshev, *Parallzl algorithm for solvint some spetral problems of linear algebra*, Lin. Alg. Appl. 188-189 (1993), pp. 489–520.

[18] MathWorks, *MATLAB user's guide (for windows version)*, The Math Works, Inc., 2002.

[19] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem*, Springer-Verlag, 1991.

[20] C. Paige and C. Van Loan , *A Schur decomposition for Hamilitonian matrices*, Lin. Alg. Appl., 41(1981), pp. 11–32.

[21] J. G. Sun, *Sensitivity of the discrete-time algebra Riccati equation*, Lin. Alg. Appl., 255-276 (1998), pp. 595–615.

[22] P. Van Dooren, *A generalized eigenvalue approach for solving Riccati equations*, SIAM J. Sci. Comput., 2 (1981), pp. 121–135.