ON THE DOUBLING ALGORITHM FOR A (SHIFTED) NONSYMMETRIC ALGEBRAIC RICCATI EQUATION

CHUN-HUA GUO*, BRUNO IANNAZZO†, AND BEATRICE MEINI‡

Abstract. Nonsymmetric algebraic Riccati equations for which the four coefficient matrices form an irreducible M-matrix M are considered. The emphasis is on the case where M is an irreducible singular M-matrix, which arises in the study of Markov models. The doubling algorithm is considered for finding the minimal nonnegative solution, the one of practical interest. The algorithm has been recently studied by others for the case where M is a nonsingular M-matrix. A shift technique is proposed to transform the original Riccati equation into a new Riccati equation for which the four coefficient matrices form a nonsingular matrix. The convergence of the doubling algorithm is accelerated when it is applied to the shifted Riccati equation.

Key words. nonsymmetric algebraic Riccati equation, minimal nonnegative solution, doubling algorithm, convergence acceleration, shift technique

AMS subject classifications. 15A24, 15A48, 65F30, 65H10

1. Introduction.

We consider the nonsymmetric algebraic Riccati equation (or NARE)

$$XCX - XD - AX + B = 0, \tag{1.1}$$

where A, B, C, D are real matrices of sizes $m \times m, m \times n, n \times m, n \times n$, respectively, and we assume throughout that

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix} \tag{1.2}$$

is a nonsingular M-matrix or an irreducible singular M-matrix. As usual in algebraic Riccati equations theory one associates with the equation (1.1) the matrix

$$H = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix}. \tag{1.3}$$

Some relevant definitions are given below.

For any matrices $A, B \in \mathbb{R}^{m\times n}$, we write $A \geq B (A > B)$ if $a_{ij} \geq b_{ij} (a_{ij} > b_{ij})$ for all i, j. A real square matrix A is called a Z-matrix if all its off-diagonal elements are nonpositive. Any Z-matrix A can be written as $sI - B$ with $B \geq 0$. A Z-matrix A is called an M-matrix if $s \geq \rho(B)$, where $\rho(\cdot)$ is the spectral radius; it is called a singular M-matrix if $s = \rho(B)$ and a nonsingular M-matrix if $s > \rho(B)$. Given a square matrix A, we will denote by $\sigma(A)$ the set of the eigenvalues of A.

The NARE (1.1) has applications in transport theory and Markov models [17, 22, 23]. The solution of practical interest is the minimal nonnegative solution. The equation has attracted much attention recently [2, 4, 5, 7, 8, 10, 12, 13, 14, 18, 19, 21].

*Department of Mathematics and Statistics, University of Regina, Regina, SK S4S 0A2, Canada (chguo@math.uregina.ca). This author's research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

†Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, 56127 Pisa, Italy (iannazzo@mail.dm.unipi.it).

‡Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, 56127 Pisa, Italy (meini@dm.unipi.it).

Some properties of the NARE (1.1) are summarized below. See [7] and [8] for more details.

THEOREM 1.1. *The equation* (1.1) *has a minimal nonnegative solution* X. *If* M *is irreducible, then* $X > 0$ *and* $A-XC$ *and* $D-CX$ *are irreducible* M*-matrices. If* M *is a nonsingular* M*-matrix, then* $A-XC$ *and* $D-CX$ *are nonsingular* M*-matrices.*

We will also need the dual equation of (1.1)

$$YBY - YA - DY + C = 0. \tag{1.4}$$

This equation has the same type as (1.1): the matrix

$$\begin{bmatrix} A & -B \\ -C & D \end{bmatrix}$$

is a nonsingular M-matrix or an irreducible singular M-matrix if and only if the matrix M is so. The minimal nonnegative solution of (1.4) is denoted by Y.

A number of numerical methods have been studied for finding the minimal solution X. Recently, a doubling algorithm is studied in [14] and is shown to be efficient. The doubling algorithm itself is not new; it was studied in [1], for example. However, the presentation in [14] provides some new information about the algorithm, which makes its analysis easier for the NARE (1.1).

In [14], the discussion is limited to the case where M is a nonsingular M-matrix. In the application of the NARE in Markov chains, however, the most important case is the one where M is an irreducible singular M-matrix with zero row sums. So in this paper, we will assume that M is an irreducible (singular or nonsingular) M-matrix, with the emphasis on the singular case.

We show the applicability and convergence properties of the structure preserving doubling algorithm of [14] when M is singular. In particular we show that the algorithm has quadratic convergence when 0 is a simple eigenvalue of H. From the numerical experiments performed so far, the doubling algorithm shows a linear convergence of rate 1/2 if 0 has algebraic multiplicity equal to 2.

We introduce an alternative approach to treat the singular case based on a shift technique. The shift consists in performing a rank-one correction of the matrix H which moves one zero eigenvalue to a suitable nonzero real number. We construct a new Riccati equation associated with the shifted H, which has the same solution X of the original one, while the coefficients of the new Riccati equation form a nonsingular matrix if 0 is a simple eigenvalue of H.

We analyze the structure preserving doubling algorithm for the new Riccati equation and show that its convergence is faster (when no breakdown is encountered) than the convergence of the same algorithm applied to the original equation. In particular, when 0 is a double eigenvalue of H, the doubling algorithm applied to the new equation is shown to have quadratic convergence.

Numerical results show the effectiveness of the shift technique.

The paper is organized as follows. In Section 2 we recall some properties of the Riccati equations and nonnegative matrices. In Sections 3 and 4 we show that the structure preserving doubling algorithm of [14] can be applied also to the case where M is irreducible singular and show convergence results. In Section 5 we present the shift technique. In Section 6 we analyze the doubling algorithm applied to the new Riccati equation. In Section 7 we show some numerical experiments.

2. Preliminaries. When M is an irreducible singular M-matrix, by the Perron–Frobenius theory 0 is a simple eigenvalue and there are positive vectors u and v such that

$$u^T M = 0, \quad Mv = 0, \tag{2.1}$$

and the vectors u and v are each unique up to a scalar multiple.

For any solution S of the Riccati equation (1.1), the matrix H of (1.3) satisfies

$$H \begin{bmatrix} I \\ S \end{bmatrix} = \begin{bmatrix} I \\ S \end{bmatrix} R,$$

where $R = D - CS$. The eigenvalues of the matrix R are a subset of the eigenvalues of H.

Since $H = JM$, where $J = \begin{bmatrix} I_n & 0 \\ 0 & -I_m \end{bmatrix}$, then H has a one dimensional kernel and $u^T J$ and v are the left and right eigenvectors corresponding to the eigenvalue 0.

Writing $u^T = (u_1^T \; u_2^T)$ and $v^T = (v_1^T \; v_2^T)$, with $u_1, v_1 \in \mathbb{R}^n$ and $u_2, v_2 \in \mathbb{R}^m$, one can define $\mu = u_1^T v_1 - u_2^T v_2$.

The number μ determines some properties of the equation. Depending on the sign of μ and following a Markov chain terminology, one can classify the Riccati equations associated with an irreducible singular M-matrix in three categories: a Riccati equation will be called

(a) *positive recurrent* if $\mu > 0$;
(b) *null recurrent* if $\mu = 0$;
(c) *transient* if $\mu < 0$.

The *close to null recurrent* case, i.e. the case $\mu \approx 0$, deserves a particular attention, since it corresponds to an ill-conditioned zero eigenvalue for the matrix H, in fact if u and v are normalized such that $\|u\|_2 = \|v\|_2 = 1$, then $1/|\mu|$ is the condition number of the zero eigenvalue for the matrix H (see [6]).

In fluid queues problems, the vector v is known being the vector of ones, which will be denoted by e. In general v and u can be computed by performing a LU factorization of the matrix M and solving two triangular linear systems.

The next results concern μ, and are proved or follow easily from results shown in [7, 8, 11].

THEOREM 2.1. *Let M be an irreducible singular M-matrix, and let X and Y be the minimal nonnegative solutions of* (1.1) *and* (1.4)*, respectively. Then the following properties hold:*

(a) *if $\mu > 0$, then $Xv_1 = v_2$ and $Yv_2 < v_1$;*
(b) *if $\mu = 0$, then $Xv_1 = v_2$ and $Yv_2 = v_1$;*
(c) *if $\mu < 0$, then $Xv_1 < v_2$ and $Yv_2 = v_1$.*

THEOREM 2.2. *Let M be an irreducible M-matrix, and let $\lambda_1, \ldots, \lambda_{m+n}$ be the eigenvalues of $H = \mathrm{diag}(I_n, -I_m)M$ ordered by nonincreasing real part. Then λ_n and λ_{n+1} are real and*

$$\mathrm{Re}\lambda_{n+m} \le \cdots \le \mathrm{Re}\lambda_{n+2} < \lambda_{n+1} \le 0 \le \lambda_n < \mathrm{Re}\lambda_{n-1} \le \cdots \le \mathrm{Re}\lambda_1.$$

The minimal nonnegative solution X of the equation (1.1) *and Y of the dual equation* (1.4) *are such that the $\sigma(D - CX) = \{\lambda_1, \ldots, \lambda_n\}$ and $\sigma(A - XC) = \sigma(A - BY) = \{-\lambda_{n+1}, \ldots, -\lambda_{n+m}\}$.*

If $\mu > 0$ *then* $\lambda_n = 0, \lambda_{n+1} < 0$*; if* $\mu = 0$ *then* $\lambda_n = \lambda_{n+1} = 0$ *and there exists only one linearly independent eigenvector for the eigenvalue 0; if* $\mu < 0$ *then* $\lambda_n > 0, \lambda_{n+1} = 0$.

In the sequel, we will need some basic results about M-matrices. The first result can be found in [3], for example.

THEOREM 2.3. *For a* Z*-matrix* A*, the following are equivalent:*

(a) A *is a nonsingular* M*-matrix.*

(b) $A^{-1} \geq 0$.

(c) $Av > 0$ *for some vector* $v > 0$.

(d) *All eigenvalues of* A *have positive real parts.*

The equivalence of (a) and (c) in Theorem 2.3 implies the next result.

LEMMA 2.4. *Let* A *be a nonsingular* M*-matrix. If* $B \geq A$ *is a* Z*-matrix, then* B *is also a nonsingular* M*-matrix.*

Most of the statements in the following result are also well known.

LEMMA 2.5. *Let* M *be a nonsingular* M*-matrix or an irreducible singular* M*-matrix. Partition* M *as*

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

where M_{11} *and* M_{22} *are square matrices. Then* M_{11} *and* M_{22} *are nonsingular* M*-matrices. The Schur complement of* M_{11} (*or* M_{22}) *in* M *is also an* M*-matrix* (*singular or nonsingular according to* M)*. Moreover, the Schur complement is irreducible if* M *is irreducible.*

REMARK 2.6. The last statement in Lemma 2.5 follows from Theorem 2.3 of [20], where the irreducibility of the Schur complement is proved for any irreducible singular M-matrix of the form $I - P$ with P stochastic. For a general irreducible M-matrix M, we have $M = s(I - B)$ for some scalar $s > 0$ and some irreducible $B \geq 0$ with $\rho(B) \leq 1$. Note that if we replace B with a stochastic matrix with the same nonzero pattern, there will be no change of the nonzero pattern in the Schur complement. In other words, the irreducibility will not change.

3. The doubling algorithm. In this section we review the structure preserving doubling algorithm (SDA) for the NARE (1.1) and show that the algorithm is well-defined when M is an irreducible singular M-matrix. When M is a nonsingular M-matrix, the algorithm has already been shown to be well defined in [14], although the selection of a parameter in the algorithm is slightly more restrictive in [14].

For the minimal nonnegative solution X of the NARE (1.1), we have

$$H \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} R, \tag{3.1}$$

where H is defined in (1.3) and $R = D - CX$.

Using the Cayley transform

$$\mathcal{C}_\gamma : z \to \frac{z - \gamma}{z + \gamma}, \tag{3.2}$$

with a scalar $\gamma > 0$, we can transform (3.1) into

$$(H - \gamma I) \begin{bmatrix} I \\ X \end{bmatrix} = (H + \gamma I) \begin{bmatrix} I \\ X \end{bmatrix} R_\gamma, \tag{3.3}$$

where

$$R_\gamma = \mathcal{C}_\gamma(R) = (R + \gamma I)^{-1}(R - \gamma I).$$

Note that $R + \gamma I$ is nonsingular since R is an M-matrix by Theorem 1.1. For any $\gamma > 0$, the matrix $M_\gamma = M + \gamma I$ is a nonsingular M-matrix. So

$$A_\gamma = A + \gamma I, \quad D_\gamma = D + \gamma I$$

are nonsingular M-matrices. Let

$$W_\gamma = A_\gamma - BD_\gamma^{-1}C, \quad V_\gamma = D_\gamma - CA_\gamma^{-1}B \tag{3.4}$$

be the Schur complements of D_γ and A_γ, respectively, in M_γ. They are both nonsingular M-matrices by Lemma 2.5. It is shown in [14] that (3.3) can be reduced to

$$K \begin{bmatrix} I \\ X \end{bmatrix} = L \begin{bmatrix} I \\ X \end{bmatrix} R_\gamma, \tag{3.5}$$

by premultiplying both sides of (3.3) with a proper nonsingular matrix, where

$$K = \begin{bmatrix} E_\gamma & 0 \\ -H_\gamma & I \end{bmatrix}, \quad L = \begin{bmatrix} I & -G_\gamma \\ 0 & F_\gamma \end{bmatrix},$$

with

$$\begin{aligned} E_\gamma &= I - 2\gamma V_\gamma^{-1}, \quad F_\gamma = I - 2\gamma W_\gamma^{-1}, \\ G_\gamma &= 2\gamma D_\gamma^{-1} C W_\gamma^{-1}, \quad H_\gamma = 2\gamma W_\gamma^{-1} B D_\gamma^{-1}. \end{aligned} \tag{3.6}$$

Similarly, for the minimal nonnegative solution Y of the NARE (1.4), we have

$$(H - \gamma I) \begin{bmatrix} Y \\ I \end{bmatrix} S_\gamma = (H + \gamma I) \begin{bmatrix} Y \\ I \end{bmatrix} \tag{3.7}$$

and then

$$K \begin{bmatrix} Y \\ I \end{bmatrix} S_\gamma = L \begin{bmatrix} Y \\ I \end{bmatrix}, \tag{3.8}$$

where $S_\gamma = (S + \gamma I)^{-1}(S - \gamma I)$ with $S = A - BY$ being an M-matrix.

The doubling algorithm presented in [14] is the following, where the sequences $\{H_k\}$ and $\{G_k\}$ are going to approximate X and Y, respectively.

ALGORITHM 3.1.

$$\begin{aligned} &E_0 = E_\gamma, \quad F_0 = F_\gamma, \quad G_0 = G_\gamma, \quad H_0 = H_\gamma, \\ &E_{k+1} = E_k(I - G_k H_k)^{-1} E_k, \\ &F_{k+1} = F_k(I - H_k G_k)^{-1} F_k, \\ &G_{k+1} = G_k + E_k(I - G_k H_k)^{-1} G_k F_k, \\ &H_{k+1} = H_k + F_k(I - H_k G_k)^{-1} H_k E_k. \end{aligned}$$

In this section we show that the algorithm is well-defined. The convergence behavior of the algorithm will be studied in the next section.

THEOREM 3.2. *Let M be an irreducible M-matrix and*

$$\gamma \geq \max\{\max_{1\leq i\leq m} a_{ii}, \max_{1\leq i\leq n} d_{ii}\},$$

where a_{ii} and d_{ii} are the diagonal elements of A and D, respectively. Then E_γ, F_γ, R_γ, $S_\gamma < 0$. Moreover, $0 \leq G_\gamma < Y, 0 \leq H_\gamma < X, G_\gamma, H_\gamma \neq 0$, $I - G_\gamma H_\gamma$ and $I - H_\gamma G_\gamma$ are nonsingular M-matrices.

Proof. We have

$$E_\gamma = I - 2\gamma V_\gamma^{-1} = V_\gamma^{-1}(V_\gamma - 2\gamma I).$$

Since V_γ is the Schur complement of A_γ in the irreducible nonsingular M-matrix M_γ, it is also an irreducible nonsingular M-matrix by Lemma 2.5. So $V_\gamma^{-1} > 0$ (see [3]). Since $\gamma \geq \max_{1\leq i\leq n} d_{ii}$, $V_\gamma - 2\gamma I = -\gamma I + D - CA_\gamma^{-1}B \leq 0$. Since $V_\gamma - 2\gamma I$ is irreducible, it has no zero columns. It then follows that $E_\gamma < 0$.

Since $R = D - CX$ is an irreducible M-matrix by Theorem 1.1, $R + \gamma I$ is an irreducible nonsingular M-matrix and $(R + \gamma I)^{-1} > 0$, for any $\gamma > 0$. For $\gamma \geq \max_{1\leq i\leq n} d_{ii}$, $R - \gamma I = D - \gamma I - CX \leq 0$. Since $R - \gamma I$ is irreducible and thus has no zero columns, it follows that $R_\gamma = (R + \gamma I)^{-1}(R - \gamma I) < 0$.

Similarly, using $\gamma \geq \max_{1\leq i\leq m} a_{ii}$, we can prove that $F_\gamma < 0$, $S_\gamma < 0$.

It is clear that $G_\gamma, H_\gamma \geq 0$. Since M is irreducible, $B, C \neq 0$. It follows that $H_\gamma, G_\gamma \neq 0$. It is shown in [14] that $X - H_\gamma = F_\gamma X R_\gamma$. Since $F_\gamma X R_\gamma > 0$, we have $0 \leq H_\gamma < X$. Similarly, we have $0 \leq G_\gamma < Y$. So $0 \leq G_\gamma H_\gamma < YX$. By Theorem 2.1 we have $YXv_1 \leq v_1$. Thus $\rho(G_\gamma H_\gamma) < \rho(YX) \leq 1$ by the Perron–Frobenius theory. Therefore, $I - G_\gamma H_\gamma$ is a nonsingular M-matrix by Theorem 2.3. Similarly, $I - H_\gamma G_\gamma$ is a nonsingular M-matrix. ∎

THEOREM 3.3. *Let M be an irreducible M-matrix. Then for $k \geq 1$, $E_k, F_k > 0$, $H_{k-1} < H_k < X$, $G_{k-1} < G_k < Y$ and $I - H_k G_k, I - G_k H_k$ are nonsingular M-matrices.*

Proof. For any nonnegative matrices U, V, W such that UVW is defined, if $U, W > 0$ and $V \neq 0$ then $UVW > 0$. Since $E_0, F_0 < 0$ and $I - G_0H_0, I - H_0G_0$ are nonsingular M-matrices, we have

$$E_1, F_1 > 0, \quad H_1 > H_0, \quad G_1 > G_0.$$

For the doubling algorithm, it is shown in [14] that

$$X - H_1 = F_1 X R_\gamma^2, \quad Y - G_1 = E_1 Y S_\gamma^2.$$

Thus, $H_1 < X$ and $G_1 < Y$. Then, as in the proof of Theorem 3.2, $I - G_1H_1$ and $I - H_1G_1$ are nonsingular M-matrices. The statements in the theorem are now easily proved by induction. ∎

4. Convergence of the doubling algorithm. For the doubling algorithm, we have (see [14])

$$X - H_k = F_k X R_\gamma^{2^k}, \quad Y - G_k = E_k Y S_\gamma^{2^k},$$

$$E_k = (I - G_k X) R_\gamma^{2^k} \leq R_\gamma^{2^k}, \quad F_k = (I - H_k Y) S_\gamma^{2^k} \leq S_\gamma^{2^k} \tag{4.1}$$

for each $k \geq 1$. So

$$X - H_k = (I - H_k Y) S_\gamma^{2^k} X R_\gamma^{2^k} \leq S_\gamma^{2^k} X R_\gamma^{2^k}, \tag{4.2}$$

$$Y - G_k = (I - G_k X) R_\gamma^{2^k} Y S_\gamma^{2^k} \leq R_\gamma^{2^k} Y S_\gamma^{2^k}. \tag{4.3}$$

When M is an irreducible nonsingular M-matrix, we have $\rho(R_\gamma) < 1$ and $\rho(S_\gamma) < 1$. It follows that $\{H_k\}$ converges to X, $\{G_k\}$ converges to Y, $\{E_k\}$ and $\{F_k\}$ converge to 0, all quadratically. This result is shown in [14] under the assumption that $\gamma > \max\{\max a_{ii}, \max d_{ii}\}$, but without the irreducibility assumption. Here we would like to allow $\gamma = \max\{\max a_{ii}, \max d_{ii}\}$, since this γ will be shown to be optimal in some sense. From (4.2) and (4.3), we also have

$$\limsup_{k\to\infty} \sqrt[2^k]{\|H_k - X\|} \leq \rho(R_\gamma)\rho(S_\gamma), \tag{4.4}$$

$$\limsup_{k\to\infty} \sqrt[2^k]{\|G_k - Y\|} \leq \rho(R_\gamma)\rho(S_\gamma). \tag{4.5}$$

THEOREM 4.1. *Let M be an irreducible singular M-matrix.*

(a) *If $\mu > 0$, then $\{H_k\}$ ($\{G_k\}$) converges to X (Y) quadratically with*

$$\limsup_{k\to\infty} \sqrt[2^k]{\|H_k - X\|} \leq \rho(S_\gamma) < 1, \quad \limsup_{k\to\infty} \sqrt[2^k]{\|G_k - Y\|} \leq \rho(S_\gamma),$$

$\{F_k\}$ converges to 0 quadratically with

$$\limsup_{k\to\infty} \sqrt[2^k]{\|F_k\|} \leq \rho(S_\gamma),$$

and $\{E_k\}$ is bounded.

(b) *If $\mu < 0$, then $\{H_k\}$ ($\{G_k\}$) converges to X (Y) quadratically with*

$$\limsup_{k\to\infty} \sqrt[2^k]{\|H_k - X\|} \leq \rho(R_\gamma) < 1, \quad \limsup_{k\to\infty} \sqrt[2^k]{\|G_k - Y\|} \leq \rho(R_\gamma),$$

$\{E_k\}$ converges to 0 quadratically with

$$\limsup_{k\to\infty} \sqrt[2^k]{\|E_k\|} \leq \rho(R_\gamma),$$

and $\{F_k\}$ is bounded.

(c) *If $\mu = 0$, then $\{H_k\}$ converges to X, $\{G_k\}$ converges to Y and $\{E_k\}, \{F_k\}$ are bounded.*

Proof. When $\mu > 0$, $\rho(S_\gamma) < 1$ and $\rho(R_\gamma) = 1$. Moreover, -1 is a simple eigenvalue of R_γ and there are no other eigenvalues on the unit circle. When $\mu < 0$, $\rho(R_\gamma) < 1$ and $\rho(S_\gamma) = 1$. Moreover, -1 is a simple eigenvalue of S_γ and there are no other eigenvalues on the unit circle. The statements in (a) and (b) are then valid in view of (4.1), (4.2) and (4.3).

When $\mu = 0$, $\rho(R_\gamma) = 1$, -1 is a simple eigenvalue of R_γ and there are no other eigenvalues on the unit circle. Also, $\rho(S_\gamma) = 1$, -1 is a simple eigenvalue of S_γ and there are no other eigenvalues on the unit circle. The boundedness of $\{E_k\}$ and $\{F_k\}$ then follows immediately. However, from (4.2) and (4.3), we cannot see the convergence of $\{H_k\}$ and $\{G_k\}$ to X and Y, respectively. So we will take a different approach.

For the minimal solution X of (1.1), we have from (3.5)

$$E_\gamma = (I - G_\gamma X)R_\gamma, \quad X - H_\gamma = F_\gamma X R_\gamma. \tag{4.6}$$

Since $0 \le G_\gamma X < YX$, $I - G_\gamma X$ is a nonsingular M-matrix as in the proof of Theorem 3.2. Eliminating R_γ in (4.6) gives

$$X = F_\gamma X (I - G_\gamma X)^{-1} E_\gamma + H_\gamma. \tag{4.7}$$

We now consider the basic fixed-point iteration for (4.7):

$$X_{k+1} = F_\gamma X_k (I - G_\gamma X_k)^{-1} E_\gamma + H_\gamma, \quad X_0 = 0. \tag{4.8}$$

It is easily proved by induction that $I - G_\gamma X_k$ is a nonsingular M-matrix and $X_k \le X_{k+1} \le X$ for all $k \ge 0$. Therefore, $\lim X_k = \widehat{X}$ with $0 \le \widehat{X} \le X$. Since $I - G_\gamma X$ is a nonsingular M-matrix, so is $I - G_\gamma \widehat{X}$. Thus we have

$$\widehat{X} = F_\gamma \widehat{X} (I - G_\gamma \widehat{X})^{-1} E_\gamma + H_\gamma. \tag{4.9}$$

We are going to show $\widehat{X} = X$. Let $\widehat{R}_\gamma = (I - G_\gamma \widehat{X})^{-1} E_\gamma$. Then

$$E_\gamma = (I - G_\gamma \widehat{X})\widehat{R}_\gamma, \quad \widehat{X} - H_\gamma = F_\gamma \widehat{X} \widehat{R}_\gamma.$$

So instead of (3.5) we have

$$K \begin{bmatrix} I \\ \widehat{X} \end{bmatrix} = L \begin{bmatrix} I \\ \widehat{X} \end{bmatrix} \widehat{R}_\gamma, \tag{4.10}$$

which can be transformed back to

$$(H - \gamma I) \begin{bmatrix} I \\ \widehat{X} \end{bmatrix} = (H + \gamma I) \begin{bmatrix} I \\ \widehat{X} \end{bmatrix} \widehat{R}_\gamma. \tag{4.11}$$

If $G_\gamma \widehat{X} = G_\gamma X$, then $\widehat{R}_\gamma = R_\gamma$ and $I - \widehat{R}_\gamma$ is nonsingular. If $G_\gamma \widehat{X} \ne G_\gamma X$, then we have $0 < (I - G_\gamma \widehat{X})^{-1}(-E_\gamma) \le (I - G_\gamma X)^{-1}(-E_\gamma)$ and $(I - G_\gamma \widehat{X})^{-1}(-E_\gamma) \ne (I - G_\gamma X)^{-1}(-E_\gamma)$ since $E_\gamma < 0$. It follows from the Perron–Frobenius theory that $\rho((I - G_\gamma \widehat{X})^{-1}(-E_\gamma)) < \rho((I - G_\gamma X)^{-1}(-E_\gamma))$. Thus $\rho(\widehat{R}_\gamma) < \rho(R_\gamma) = 1$ and again $I - \widehat{R}_\gamma$ is nonsingular. Now (4.11) can be rewritten as

$$H \begin{bmatrix} I \\ \widehat{X} \end{bmatrix} = \begin{bmatrix} I \\ \widehat{X} \end{bmatrix} \gamma (I + \widehat{R}_\gamma)(I - \widehat{R}_\gamma)^{-1}. \tag{4.12}$$

Thus $\widehat{X}$ is also a nonnegative solution of (1.1). Since X is minimal we have $\widehat{X} = X$ and so $\lim X_k = X$. Now, the sequence $\{H_k\}$ produced by the doubling algorithm is such that $H_k = X_{2^k}$ (see [1]). Therefore, $\lim H_k = X$, as required. The proof of $\lim G_k = Y$ is similar. □

Remark 4.2. When $\mu = 0$, the convergence of the doubling algorithm is not quadratic in general since the convergence of $\{X_k\}$ in (4.8) is sublinear in general. But the relation $H_k = X_{2^k}$ itself says that the convergence of the doubling algorithm is much faster than the basic fixed point iteration. The convergence in this case has been observed to be linear with rate 1/2.

In the doubling algorithm, we have the freedom to choose the parameter γ. In view of (4.4) and (4.5), the next result says $\gamma = \max\{\max a_{ii}, \max d_{ii}\}$ is optimal, in some sense, for the doubling algorithm.

THEOREM 4.3. *For $\gamma \ge \max\{\max a_{ii}, \max d_{ii}\}$, $\rho(R_\gamma)$ and $\rho(S_\gamma)$ are nondecreasing functions of γ.*

Proof. Since $R = D - CX$ is an irreducible M-matrix, it can be written in the form $sI - N$, where $N \ge 0$ is irreducible. It follows from the Perron–Frobenius theorem that there is a positive vector v such that $Rv = \lambda_n v$. Now

$$-R_\gamma v = (\gamma I + R)^{-1}(\gamma I - R)v = (\gamma + \lambda_n)^{-1}(\gamma - \lambda_n)v.$$

Since $-R_\gamma > 0$, it follows from the Perron–Frobenius theory that $\rho(R_\gamma) = \rho(-R_\gamma) = (\gamma + \lambda_n)^{-1}(\gamma - \lambda_n)$, which is a nondecreasing function of γ. Similarly, $\rho(S_\gamma)$ is a nondecreasing function of γ. ∎

5. A shift technique. In this section we assume that M is an irreducible singular M-matrix. The vectors u and v are as in (2.1), and we have three cases: $\mu > 0, \mu = 0$ and $\mu < 0$. The next result shows that the case $\mu < 0$ is easily reduced to the case $\mu > 0$.

LEMMA 5.1. *The matrix X is the minimal nonnegative solution of* (1.1) *if and only if $Z = X^T$ is the minmal nonnegative solution of the equation*

$$ZC^TZ - ZA^T - D^TZ + B^T = 0. \tag{5.1}$$

The equation (1.1) *is transient if and only if the equation* (5.1) *is positive recurrent.*

Proof. The first statement is easily shown by taking transpose on both sides of the equation. The M-matrix corresponding to (5.1) is

$$M_t = \begin{bmatrix} A^T & -C^T \\ -B^T & D^T \end{bmatrix}.$$

Since

$$\begin{bmatrix} v_2^T & v_1^T \end{bmatrix} M_t = 0, \quad M_t \begin{bmatrix} u_2 \\ u_1 \end{bmatrix} = 0,$$

the second statement follows readily. ∎

REMARK 5.2. When $\mu \le 0$, from the above proof and Theorem 2.1 we know that the minimal nonnegative solution X of (1.1) is such that $X^Tu_2 = u_1$, or in other words, $u_2^TX = u_1^T$.

From now on, we assume that $\mu \ge 0$.

Our shift technique will be based on the following result (see also [16]).

LEMMA 5.3. *Let T be a singular matrix and $Tw = 0$ for a nonzero vector w. Assume that r is a vector with $r^Tw = 1$ and η is a scalar. Then the eigenvalues of the matrix*

$$\widehat{T} = T + \eta w r^T$$

are those of T except that one zero eigenvalue of T is replaced by η.

Proof. We may easily verify that $\widehat{T} - \lambda I = (T - \lambda I)(I - \lambda^{-1}\eta w r^T)$ for any complex number λ different from zero. Taking determinants one has that, for any $\lambda \ne 0$,

$$\det(\widehat{T} - \lambda I) = \det(T - \lambda I)\frac{\lambda - \eta}{\lambda}.$$

By using a continuity argument, since $\lambda = 0$ is a zero of $\det(T - \lambda I)$, the above equation holds for any λ. This completes the proof. □

We now construct a rank-one modification of the matrix H in (1.3):

$$\widehat{H} = H + \eta v p^T, \tag{5.2}$$

where $\eta > 0$ is a scalar and $p \geq 0$ is a vector with $p^T v = 1$. Since $\widehat{H}$ is a singular matrix with $Hv = 0$, we know from Lemma 5.3 that the eigenvalues of $\widehat{H}$ are those of H except that one zero eigenvalue of H is replaced by η.

We write $p^T = (p_1^T, p_2^T)$ and

$$\widehat{H} = \begin{bmatrix} \widehat{D} & -\widehat{C} \\ \widehat{B} & -\widehat{A} \end{bmatrix}, \quad \widehat{M} = \begin{bmatrix} \widehat{D} & -\widehat{C} \\ -\widehat{B} & \widehat{A} \end{bmatrix},$$

where

$$\widehat{D} = D + \eta v_1 p_1^T, \quad \widehat{C} = C - \eta v_1 p_2^T,$$

$$\widehat{B} = B + \eta v_2 p_1^T, \quad \widehat{A} = A - \eta v_2 p_2^T.$$

Corresponding to $\widehat{M}$ we define the new NARE

$$Z\widehat{C}Z - Z\widehat{D} - \widehat{A}Z + \widehat{B} = 0. \tag{5.3}$$

We have the following important property about the NARE (5.3).

THEOREM 5.4. *If $\mu \geq 0$, then $Z = X$ is a solution of the NARE (5.3) and $\sigma(\widehat{D} - \widehat{C}X) = \{\lambda_1, \ldots, \lambda_{n-1}, \eta\}$, where X is the minimal nonnegative solution of the original NARE (1.1).*

Proof. Observe that

$$X\widehat{C}X - X\widehat{D} - \widehat{A}X + \widehat{B} = XCX - XD - AX + B - \eta(Xv_1 - v_2)(p_2^T X + p_1^T).$$

Since X is a solution of (1.1) and $Xv_1 = v_2$ by Theorem 2.1, X is also a solution of the shifted equation (5.3). We have $\widehat{D} - \widehat{C}X = D - CX + \eta v_1(p_1^T + p_2^T X)$. Since $(D - CX)v_1 = Dv_1 - Cv_2 = 0$ and $(p_1^T + p_2^T X)v_1 = p^T v = 1$, the eigenvalues of $\widehat{D} - \widehat{C}X$ are $\lambda_1, \ldots, \lambda_{n-1}, \eta$ by Theorem 2.2 and Lemma 5.3. □

In what follows we will show that the dual equation of (5.3) has a solution $\widehat{Y}$ such that the eigenvalues of $-(\widehat{A} - \widehat{B}\widehat{Y})$ are the remaining eigenvalues of $\widehat{H}$: $\lambda_{n+1}, \ldots, \lambda_{n+m}$.

LEMMA 5.5. *The eigenvalues of the matrix*

$$W = \begin{bmatrix} \eta p_1^T(v_1 - Yv_2) & \eta(p_1^T Y + p_2^T) \\ (B + \eta v_2 p_1^T)(v_1 - Yv_2) & -(A - BY - \eta v_2(p_1^T Y + p_2^T)) \end{bmatrix}$$

are $\eta, \lambda_{n+1}, \ldots, \lambda_{n+m}$.

Proof. We have

$$W = W_0 + \eta \begin{bmatrix} 1 \\ v_2 \end{bmatrix} \begin{bmatrix} p_1^T(v_1 - Yv_2) & p_1^T Y + p_2^T \end{bmatrix},$$

where

$$W_0 = \begin{bmatrix} 0 & 0 \\ B(v_1 - Yv_2) & -(A - BY) \end{bmatrix}.$$

The eigenvalues of W_0 are $0, \lambda_{n+1}, \ldots, \lambda_{n+m}$ by Theorem 2.2. Since

$$W_0 \begin{bmatrix} 1 \\ v_2 \end{bmatrix} = 0, \quad \begin{bmatrix} p_1^T(v_1 - Yv_2) & p_1^T Y + p_2^T \end{bmatrix} \begin{bmatrix} 1 \\ v_2 \end{bmatrix} = p^T v = 1,$$

the eigenvalues of W are $\eta, \lambda_{n+1}, \ldots, \lambda_{n+m}$ by Lemma 5.3. □

LEMMA 5.6. *If $\mu > 0$, then there is a positive vector f such that $(1, f^T)$ is a left eigenvector of W corresponding to the eigenvalue η.*

Proof. When $\mu > 0$, we have $Yv_2 < v_1$ by Theorem 2.1. Since $A - BY$ is an irreducible M-matrix, $A - BY - \eta v_2(p_1^T Y + p_2^T)$ is an irreducible Z-matrix. Since $\eta(p_1^T Y + p_2^T) \neq 0$ and $(B + \eta v_2 p_1^T)(v_1 - Yv_2) \neq 0$, the matrix W is irreducible by a simple graph argument. It is clear that W can be written in the form $N - sI$, where $N \geq 0$ is irreducible. The result then follows from the Perron–Frobenius theorem. □

LEMMA 5.7. *If $\mu > 0$, then the matrix $\hat{Y} = Y + (Yv_2 - v_1)f^T$ is a solution of the dual equation of* (5.3).

Proof. Let $\mathcal{R}(Z) = ZBZ - ZA - DZ + C$ and $\widehat{\mathcal{R}}(Z) = Z\widehat{B}Z - Z\widehat{A} - \widehat{D}Z + \widehat{C}$. We are to show $\widehat{\mathcal{R}}(\widehat{Y}) = 0$. Since $\mathcal{R}(Y) = 0$, we have

$$\widehat{\mathcal{R}}(\widehat{Y}) = (\widehat{\mathcal{R}}(\widehat{Y}) - \mathcal{R}(\widehat{Y})) + (\mathcal{R}(\widehat{Y}) - \mathcal{R}(Y)).$$

A straightforward computation shows

$$\begin{aligned} \widehat{\mathcal{R}}(\widehat{Y}) - \mathcal{R}(\widehat{Y}) &= \eta(\widehat{Y}v_2 - v_1)(p_1^T \widehat{Y} + p_2^T) \\ &= \eta(Yv_2 - v_1)(1 + f^T v_2)(p_1^T Y + p_2^T + p_1^T(Yv_2 - v_1)f^T). \end{aligned}$$

Also, we have

$$\mathcal{R}(\widehat{Y}) - \mathcal{R}(Y) = (Yv_2 - v_1)(f^T B(Yv_2 - v_1)f^T - f^T(A - BY)).$$

where we have used the fact that

$$\begin{aligned} (D - YB)(Yv_2 - v_1) &= -Dv_1 + YBv_1 + (DY - YBY)v_2 \\ &= -Dv_1 + YBv_1 + (C - YA)v_2 = 0. \end{aligned}$$

Thus, to show $\widehat{\mathcal{R}}(\widehat{Y}) = 0$, we only need to check

$$\begin{aligned} &f^T(B + \eta v_2 p_1^T)(Yv_2 - v_1)f^T + \eta p_1^T(Yv_2 - v_1)f^T \\ &\quad - f^T(A - BY - \eta v_2(p_1^T Y + p_2^T)) + \eta(p_1^T Y + p_2^T) = 0, \end{aligned}$$

which is true by the choice of f in Lemma 5.6. □

We now show that the solution $\widehat{Y}$ has the desired spectral property.

THEOREM 5.8. *If $\mu > 0$, then the solution $\widehat{Y}$ of the dual equation of* (5.3) *given in Lemma* 5.7 *is such that $\sigma(\widehat{A} - \widehat{B}\widehat{Y}) = \{-\lambda_{n+1}, \ldots, -\lambda_{n+m}\}$.*

Proof. Notice that

$$\widehat{A} - \widehat{B}\widehat{Y} = A - BY - \eta v_2(p_1^T Y + p_2^T) - (B + \eta v_2 p_1^T)(Yv_2 - v_1)f^T.$$

For the matrix $W = \left[\begin{smallmatrix} W_{1,1} & W_{1,2} \\ W_{2,1} & W_{2,2} \end{smallmatrix}\right]$ in Lemma 5.5, we have

$$\begin{bmatrix} 1 & f^T \\ 0 & I \end{bmatrix} W \begin{bmatrix} 1 & f^T \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} 1 & f^T \\ 0 & I \end{bmatrix} W \begin{bmatrix} 1 & -f^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} \eta & 0 \\ W_{21} & -(\widehat{A} - \widehat{B}\widehat{Y}) \end{bmatrix},$$

where we have used Lemma 5.6. Therefore, the eigenvalues of $\widehat{A} - \widehat{B}\widehat{Y}$ are $-\lambda_{n+1}$, …, $-\lambda_{n+m}$ by Lemma 5.5. ∎

The case $\mu = 0$ has to be treated separately since $Yv_2 = v_1$ in this case and thus the matrix W in Lemma 5.5 does not have a left eigenvector of the form $(1, f^T)$ corresponding to the eigenvalue η. In this case, we need to assume $p_1 > 0$. Actually, it is advisable in general to use a vector p with $p_1 > 0$ also in the case $\mu > 0$, since this choice of p guarantees that the matrix $\widehat{Y}$ is bounded independent of the nearness to null recurrence (as can be seen in the proof of the following theorem). In the next section, however, we will use a vector p without this assumption for a special class of the NARE (1.1). There, the boundedness of $\widehat{Y}$ independent of the nearness to null recurrence will be guaranteed in another way.

THEOREM 5.9. *If $\mu = 0$ and $p_1 > 0$, then the dual equation of* (5.3) *has a solution $\widehat{Y}$ such that $\sigma(\widehat{A} - \widehat{B}\widehat{Y}) = \{-\lambda_{n+1}, \ldots, -\lambda_{n+m}\}$.*

Proof. We use a continuity argument similar to the one used in [9] when a shift technique in [15] is used for null recurrent quasi-birth-death problems. We introduce the irreducible singular M-matrix

$$M(k) = \begin{bmatrix} D(k) & -C(k) \\ -B(k) & A(k) \end{bmatrix} = \begin{bmatrix} D & -C \\ -(1+\frac{1}{k})B & (1+\frac{1}{k})A \end{bmatrix}$$

($k = 1, 2, \ldots$). The left and right eigenvectors of $M(k)$ corresponding to the zero eigenvalue are given by

$$u_1(k) = u_1, \quad u_2(k) = (1 + \frac{1}{k})^{-1} u_2, \quad v_1(k) = v_1, \quad v_2(k) = v_2.$$

Thus, the NARE corresponding to $M(k)$

$$ZCZ - ZD - A(k)Z + B(k) = 0 \qquad (5.4)$$

is positive recurrent since $u_1(k)^T v_1(k) > u_2(k)^T v_2(k)$. Let $Y(k)$ be the minimal nonnegative solution of the dual equation of (5.4). Then $Y(k)v_2 < v_1$ and in particular the sequence $\{Y(k)\}$ is bounded. When M is replaced by $M(k)$, we have a matrix $W(k)$ corresponding to the matrix W in Lemma 5.5. Let $(1, f(k)^T)$ be the left eigenvector of $W(k)$ corresponding to the eigenvalue η. Now, $\widehat{Y}(k) = Y(k) + (Y(k)v_2 - v_1)f(k)^T$ is a solution of the dual equation of

$$Z\widehat{C}Z - Z\widehat{D} - \widehat{A}(k)Z + \widehat{B}(k) = 0,$$

where $\widehat{A}(k) = A(k) - \eta v_2 p_2^T$ and $\widehat{B}(k) = B(k) + \eta v_2 p_1^T$, and the eigenvalues of $\widehat{A}(k) - \widehat{B}(k)\widehat{Y}(k)$ are those of $A(k) - B(k)Y(k)$. We need to show that the sequence $\{\widehat{Y}(k)\}$ is bounded. Since $(1, f(k)^T)W(k) = \eta(1, f(k)^T)$, we have

$$\begin{aligned} \eta &= \eta p_1^T(v_1 - Y(k)v_2) + f(k)^T(B(k) + \eta v_2 p_1^T)(v_1 - Y(k)v_2) \\ &\geq f(k)^T(\eta v_2 p_1^T)(v_1 - Y(k)v_2) \end{aligned}$$

and thus

$$f(k)^T(v_2p_1^T)(v_1 - Y(k)v_2) = p_1^T(v_1 - Y(k)v_2)f(k)^Tv_2 \le 1.$$

Since $p_1, v_2 > 0$, $\{(v_1 - Y(k)v_2)f(k)^T\}$ is bounded and thus $\{\widehat{Y}(k)\}$ is bounded. Let $\widehat{Y}$ be any limit point of the sequence $\{\widehat{Y}(k)\}$. Then the eigenvalues of $\widehat{A} - \widehat{B}\widehat{Y}$ are those of $A - BY$ since $\lim Y(k) = Y$ by Theorem 3.3 of [11]. ◻

When $\mu = 0$, the matrix H has two zero eigenvalues. The above shift technique moves one zero eigenvalue to a positive number. We may use a *double-shift* to move the other zero eigenvalue to a negative number. Recall that $Hv = 0$, where $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, and $w^TH = 0$, where $w = \begin{bmatrix} u_1 \\ -u_2 \end{bmatrix}$. We define the matrix

$$\overline{H} = H + \eta vp^T + \xi qw^T, \tag{5.5}$$

where $\eta > 0$, $\xi < 0$, p and q are such that $p^Tv = q^Tw = 1$. Since v and w are orthogonal vectors, the double-shift moves one zero eigenvalue to η and the other to ξ. Indeed, the eigenvalues of $\widetilde{H} = H + \xi qw^T$ are those of $\widetilde{H}^T = H^T + \xi wq^T$, which are the eigenvalues of H except that one zero eigenvalue is replaced by ξ, by Lemma 5.3. Also, the eigenvalues of $\overline{H} = \widetilde{H} + \eta vp^T$ are the eigenvalues of $\widetilde{H}$ except that the remaining zero eigenvalue is replaced by η, by Lemma 5.3 again.

From $\overline{H}$ we may define a new Riccati equation

$$Z\overline{C}Z - Z\overline{D} - \overline{A}Z + \overline{B} = 0. \tag{5.6}$$

As before, the minimal nonnegative solution X of (1.1) is a solution of (5.6) such that $\sigma(\overline{D} - \overline{C}X) = \{\eta, \lambda_1, \dots, \lambda_{n-1}\}$. However, it seems very difficult to determine the existence of a solution $\overline{Y}$ of the dual equation of (5.6) such that $\sigma(\overline{A} - \overline{B}\,\overline{Y}) = \{-\xi, -\lambda_{n+2}, \dots, -\lambda_{n+m}\}$. We will not investigate the double-shift any further in this paper.

6. The doubling algorithm applied to the shifted equation. In this section we assume $\mu \ge 0$. We will show that the doubling algorithm applied to (5.3) converges faster (if no breakdown occurs) than the doubling algorithm applied to (1.1). The applicability of the SDA algorithm to the shifted equation for a general NARE is still work in progress, but we will prove that no breakdown occurs under suitable assumptions on the matrix $\widehat{M}$.

6.1. Convergence properties. By Theorems 5.4, 5.8 and 5.9, the matrices X and $\widehat{Y}$ are such that

$$\widehat{H}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix}(\widehat{D} - \widehat{C}X), \quad \widehat{H}\begin{bmatrix} \widehat{Y} \\ I \end{bmatrix} = \begin{bmatrix} \widehat{Y} \\ I \end{bmatrix}(-(\widehat{A} - \widehat{B}\widehat{Y})), \tag{6.1}$$

where $\sigma(\widehat{D} - \widehat{C}X) = \{\lambda_1, \dots, \lambda_{n-1}, \eta\}$, $\sigma(\widehat{A} - \widehat{B}\widehat{Y}) = \{-\lambda_{n+1} \dots, -\lambda_{n+m}\}$. Recall that we need to assume $p_1 > 0$ for the vector p used in the shift technique when $\mu = 0$, to get the second equation in (6.1).

We apply the Cayley transform with $\gamma > 0$ to each of the equations in (6.1), thus obtaining

$$\begin{aligned} (\widehat{H} - \gamma I)\begin{bmatrix} I \\ X \end{bmatrix} &= (\widehat{H} + \gamma I)\begin{bmatrix} I \\ X \end{bmatrix}\mathcal{C}_\gamma(\widehat{D} - \widehat{C}X), \\ (\widehat{H} - \gamma I)\begin{bmatrix} \widehat{Y} \\ I \end{bmatrix}\mathcal{C}_\gamma(\widehat{A} - \widehat{B}\widehat{Y}) &= (\widehat{H} + \gamma I)\begin{bmatrix} \widehat{Y} \\ I \end{bmatrix}. \end{aligned} \tag{6.2}$$

We then proceed as in Section 3, with the equations (3.3) and (3.7) replaced by the equations in (6.2), and obtain a sequence of matrices $\{\widehat{H}_k\}$ by the doubling algorithm (assuming that no breakdown occurs). The sequence $\{\widehat{H}_k\}$ is going to approximate X, the minimal nonnegative solution of (1.1).

As in Section 4, we can prove that

$$\limsup_{k\to\infty} \sqrt[2^k]{\|\widehat{H}_k - X\|} \le \rho(\mathcal{C}_\gamma(\widehat{D} - \widehat{C}X))\rho(\mathcal{C}_\gamma(\widehat{A} - \widehat{B}\widehat{Y})).$$

For easy comparison, we recall that the sequence $\{H_k\}$ generated in Section 3 is such that

$$\limsup_{k\to\infty} \sqrt[2^k]{\|H_k - X\|} \le \rho(\mathcal{C}_\gamma(D - CX))\rho(\mathcal{C}_\gamma(A - BY)).$$

Note that

$$\begin{aligned}
&\sigma(\mathcal{C}_\gamma(\widehat{D} - \widehat{C}X)) = \{\mathcal{C}_\gamma(\lambda_1), \ldots, \mathcal{C}_\gamma(\lambda_{n-1}), \mathcal{C}_\gamma(\eta)\},\\
&\sigma(\mathcal{C}_\gamma(D - CX)) = \{\mathcal{C}_\gamma(\lambda_1), \ldots, \mathcal{C}_\gamma(\lambda_{n-1}), \mathcal{C}_\gamma(0)\},\\
&\sigma(\mathcal{C}_\gamma(\widehat{A} - \widehat{B}\widehat{Y})) = \sigma(\mathcal{C}_\gamma(A - BY)) = \{\mathcal{C}_\gamma(-\lambda_{n+1}), \ldots, \mathcal{C}_\gamma(-\lambda_{m+n})\}.
\end{aligned}$$

By Theorem 2.2 and the property of the Cayley transform, we have $\rho(\mathcal{C}_\gamma(\widehat{D} - \widehat{C}X)) < \rho(\mathcal{C}_\gamma(D - CX)) = 1$ and $\rho(\mathcal{C}_\gamma(\widehat{A} - \widehat{B}\widehat{Y})) = \rho(\mathcal{C}_\gamma(A - BY)) \le 1$. Therefore the convergence of the doubling algorithm applied to the shifted equation is faster than the convergence of the same algorithm applied to the original equation. As we will show in the numerical experiments, the number of steps necessary for convergence can decrease dramatically by using the shift technique. In particular, when $\mu = 0$, the SDA algorithm applied to the shifted equation still has quadratic convergence. According to the results in [11], the shift equation is also better conditioned than the original equation.

At this point it must be specified which is the best choice for the parameter η in terms of the speed of convergence. In fact, the fastest convergence is expected when $\rho(\mathcal{C}_\gamma(\widehat{D} - \widehat{C}X))$ is minimal, i.e., when $|\mathcal{C}_\gamma(\eta)| \le \max\{|\mathcal{C}_\gamma(\lambda_i)|, i = 1, \ldots, n-1\}$. This happens, for instance, if $\eta = \gamma$ (i.e. $\mathcal{C}_\gamma(\eta) = 0$).

6.2. Applicability for a special class of NARE. We can prove the applicability of the SDA algorithm to the shifted equation for a special class of the NARE (1.1) and for proper choices of η, p and γ.

The special class consists of equations for which

(a) either C has at least one positive column or D has at least one column with no zero entries.

(b) either A is irreducible or B has a nonzero row.

For ease of notation, we assume without loss of generality [11] that $v = e$, the vector of ones. We define $r_1 \in \mathbb{R}^n$ and $r_2 \in \mathbb{R}^m$ by

$$(r_1)_j = \min_{1\le i\le n, i\ne j} |d_{ij}|, \quad j = 1, \ldots, n, \quad (r_2)_j = \min_{1\le i\le n} c_{ij}, \quad j = 1, \ldots, m,$$

and define r by $r^T = (r_1^T, r_2^T)$. We have $r \ne 0$ for each equation in the special class, and we use the shift (5.2) with $\eta = r^T e$ and $p = r/r^T e$.

We are going to show that the matrix $\widehat{M}$ is such that

$$\widehat{B} \ge 0, \quad \widehat{C} \ge 0, \quad I \otimes \widehat{A} + \widehat{D}^T \otimes I \text{ is a nonsingular } M\text{-matrix}, \tag{6.3}$$

where $\otimes$ is the Kronecker product. It is clear that $\widehat{B} \geq 0$, $\widehat{C} \geq 0$. We only need to show that the Z-matrix $I \otimes \widehat{A} + \widehat{D}^T \otimes I$ is a nonsingular M-matrix. Note that

$$(D - r_2^T eI)e - (C - er_2^T)e = De - Ce = 0.$$

So $D - r_2^T eI$ is an M-matrix. Now,

$$\begin{aligned} I \otimes \widehat{A} + \widehat{D}^T \otimes I &\geq I \otimes (A - er_2^T) + (D - r_2^T eI)^T \otimes I + r_2^T eI \otimes I \\ &= I \otimes (A - er_2^T + r_2^T eI) + (D - r_2^T eI)^T \otimes I. \end{aligned}$$

Since $(A - er_2^T + r_2^T eI)e = Ae = Be \geq 0$, $A - er_2^T + r_2^T eI$ is a nonsingular M-matrix by the condition (b) for the special class. Therefore, $I \otimes \widehat{A} + \widehat{D}^T \otimes I$ is also a nonsingular M-matrix.

Proceeding as in Section 3, it is not difficult to show that the doubling algorithm can be applied to the NARE corresponding to $\widehat{M}$ with $\gamma \geq \max\{\max a_{ii}, \max d_{ii}\} + \|r\|_\infty$. Indeed, the matrices to be inverted in the algorithm are all nonsingular M-matrices, as in Section 3.

By our definition of the vector p, we no longer have $p_1 > 0$ unless all off-diagonal elements of D are negative. Therefore, we need a different proof for the existence of $\widehat{Y}$ with the property in (6.1).

We proceed as follows. First, we note that

$$Y\widehat{B}Y - Y\widehat{A} - \widehat{D}Y + \widehat{C} = \eta(Ye - e)(p_1^T Y + p_2^T) \leq 0,$$

By Theorem 2.3 of [7], there is a minimal $\widehat{Y} : 0 \leq \widehat{Y} \leq Y$ such that

$$\widehat{Y}\widehat{B}\widehat{Y} - \widehat{Y}\widehat{A} - \widehat{D}\widehat{Y} + \widehat{C} = 0.$$

Note that

$$\widehat{H} \begin{bmatrix} I & \widehat{Y} \\ X & I \end{bmatrix} = \begin{bmatrix} I & \widehat{Y} \\ X & I \end{bmatrix} \begin{bmatrix} \widehat{D} - \widehat{C}X & \\ & -(\widehat{A} - \widehat{B}\widehat{Y}) \end{bmatrix}.$$

When $\mu > 0$, we have $Xe = e$ and $\widehat{Y}e < e$ and the matrix

$$\begin{bmatrix} I & \widehat{Y} \\ X & I \end{bmatrix}$$

is nonsingular. So the eigenvalues of $\widehat{A} - \widehat{B}\widehat{Y}$ are $-\lambda_{n+1}, \ldots, -\lambda_{n+m}$. By a continuity argument, the eigenvalues of $\widehat{A} - \widehat{B}\widehat{Y}$ are also $-\lambda_{n+1}, \ldots, -\lambda_{n+m}$ when $\mu = 0$.

7. Numerical experiments. We compare the numerical behavior of the SDA algorithm applied to (1.1) and to the shifted equation (5.3), when $\mu \geq 0$. Recall that the case $\mu < 0$ is easily reduced to the case $\mu > 0$ through Lemma 5.1.

The numerical experiments are performed by using Matlab; the stop condition is $\min\{\|E_k\|_1, \|F_k\|_1\} < 10^{-15}$.

We take $\gamma = \max\{\max a_{ii}, \max d_{ii}\}$ for the Cayley transform, as suggested by Theorem 4.3. For the shift technique, we take $\eta = \gamma$ and $p = e/v^T e$, where e is the vector (of suitable size) with all components equal to 1.

TEST 7.1. [7] *Random choice of a singular M-matrix with $Me = 0$.* To construct M, we generate R, a 100×100 random matrix, and define $M = \mathrm{diag}(Re) - R$. The matrices A, B, C and D are 50×50.

We generate 5 different matrices M in this way, each with $\mu > 0$. In Table 7.1 we report the number of iterations and the relative residual, defined as

$$\text{res} = \frac{\|XCX - XD - AX + B\|_1}{\|XCX\|_1 + \|XD\|_1 + \|AX\|_1 + \|B\|_1}.$$

As one can see, the number of steps applied to the shifted equation is smaller, while the residual error remains roughly the same ($u \approx 2.2 \times 10^{-16}$ is the unit roundoff).

TABLE 7.1
SDA applied to original and shifted NARE

	SDA		SDA applied to shifted NARE	
	iter	res/err	iter	res/err
Test 1	12–13	res=$1.1u$–$2.1u$	5	res=$1.2u$–$1.7u$
Test 2	33	err=1.6×10^{-9}	5	err=2.2×10^{-16}
Test 3	18	err=3.5×10^{-13}	4	err=2.3×10^{-13}

TEST 7.2. [2, Example 1]*A null recurrent case.* Let

$$M = \begin{bmatrix} 0.003 & -0.001 & -0.001 & -0.001 \\ -0.001 & 0.003 & -0.001 & -0.001 \\ -0.001 & -0.001 & 0.003 & -0.001 \\ -0.001 & -0.001 & -0.001 & 0.003 \end{bmatrix}$$

where D is a 2×2 matrix. The minimal positive solution is $X = \frac{1}{2}E_{2,2}$, where $E_{m,n}$ is the $m \times n$ matrix having all entries equal to 1.

In this case the SDA algorithm shows linear convergence while the SDA applied to the shifted equation has quadratic convergence. Indeed, as reported in Table 7.1 the number of steps decreases dramatically. Since the solution is explicitly known, we have compared the absolute error, defined as the 1-norm of the difference between the exact and the computed solution, obtained with both methods. Observe that the solution computed without performing the shift is much less accurate than the one obtained by applying the shift. This phenomenon is to be expected in view of the theoretical results in [11].

TEST 7.3. [2, Example 3] *A positive recurrent Markov chain with non-square coefficients.* In this example $A = \text{diag}(0.018\, E_{2,1})$, $D = \text{diag}(180.002\, E_{18,1}) - 10\, E_{18,18}$, $B = 0.001\, E_{2,18}$ and $C = B^T$. The solution is known to be $\frac{1}{18}E_{2,18}$. The results are shown in Table 7.1, the reduction of the number of iterations for the shifted equation is significant.

REFERENCES

[1] B. D. O. ANDERSON, *Second-order convergent algorithms for the steady-state Riccati equation*, Int. J. Control, 28 (1978), pp. 295–306.

[2] N. G. BEAN, M. M. O'REILLY, AND P. G. TAYLOR, *Algorithms for return probabilities for stochastic fluid flows*, Stoch. Models, 21 (2005), pp. 149–184.

[3] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, PA, 1994.

[4] D. A. BINI, B. IANNAZZO, G. LATOUCHE, AND B. MEINI, *On the solution of algebraic Riccati equations arising in fluid queues*, Linear Algebra Appl., 413 (2006), no. 2-3, pp. 474–494.

[5] S. FITAL AND C.-H. GUO, *Convergence of the solution of a nonsymmetric matrix Riccati differential equation to its stable equilibrium solution*, J. Math. Anal. Appl., 318 (2006), pp. 648–657.

[6] G. H. Golub and C. F. Van Loan, *Matrix Computations, Second Edition*, Johns Hopkins Univ. Press, Baltimore, MD, 1989.

[7] C.-H. Guo, *Nonsymmetric algebraic Riccati equations and Wiener–Hopf factorization for M-matrices*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 225–242.

[8] C.-H. Guo, *A note on the minimal nonnegative solution of a nonsymmetric algebraic Riccati equation*, Linear Algebra Appl., 357 (2002), pp. 299–302.

[9] C.-H. Guo, *Comments on a shifted cyclic reduction algorithm for quasi-birth-death problems*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 1161–1166.

[10] C.-H. Guo, *Efficient methods for solving a nonsymmetric algebraic Riccati equation arising in stochastic fluid models*, J. Comput. Appl. Math., 192 (2006), pp. 353–373.

[11] C.-H. Guo and N. J. Higham, *Iterative solution of a nonsymmetric algebraic Riccati equation*, preprint, 2005.

[12] C.-H. Guo and A. J. Laub, *On the iterative solution of a class of nonsymmetric algebraic Riccati equations*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 376–391.

[13] X.-X. Guo and Z.-Z. Bai, *On the minimal nonnegative solution of nonsymmetric algebraic Riccati equation*, J. Comput. Math., 23 (2005), pp. 305–320.

[14] X.-X. Guo, W.-W. Lin and S.-F. Xu, *A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation*, Numer. Math., 103 (2006), pp. 393–412.

[15] C. He, B. Meini, and N. H. Rhee, *A shifted cyclic reduction algorithm for quasi-birth-death problems*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 673–691.

[16] R. A. Horn and S. Serra-Capizzano, *Canonical and standard forms for certain rank one perturbations and an application to the (complex) Google pageranking problem*, submitted for publication, 2006.

[17] J. Juang, *Existence of algebraic matrix Riccati equations arising in transport theory*, Linear Algebra Appl., 230 (1995), pp. 89–100.

[18] J. Juang and W.-W. Lin, *Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 228–243.

[19] L.-Z. Lu, *Solution form and simple iteration of a nonsymmetric algebraic Riccati equation arising in transport theory*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 679–685.

[20] C. D. Meyer, *Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems*, SIAM Rev., 31 (1989), pp. 240–272.

[21] V. Ramaswami, *Matrix analytic methods for stochastic fluid flows*, in: Proceedings of the 16th International Teletraffic Congress, Elsevier Science B. V., Edinburgh, 1999, pp. 1019–1030.

[22] L. C. G. Rogers, *Fluid models in queueing theory and Wiener–Hopf factorization of Markov chains*, Ann. Appl. Probab., 4 (1994), pp. 390–413.

[23] L. C. G. Rogers and Z. Shi, *Computing the invariant law of a fluid model*, J. Appl. Probab., 31 (1994), pp. 885–896.