ON THE ITERATIVE SOLUTION OF A CLASS OF NONSYMMETRIC ALGEBRAIC RICCATI EQUATIONS*

CHUN-HUA GUO† AND ALAN J. LAUB‡

Abstract. We consider the iterative solution of a class of nonsymmetric algebraic Riccati equations, which includes a class of algebraic Riccati equations arising in transport theory. For any equation in this class, Newton's method and a class of basic fixed-point iterations can be used to find its minimal positive solution whenever it has a positive solution. The properties of these iterative methods are studied and some practical issues are addressed. An algorithm is then proposed to find the minimal positive solution efficiently. Numerical results are also given.

Key words. nonsymmetric algebraic Riccati equations, M-matrices, Newton's method, fixed-point iterations, minimal positive solution, convergence rate

AMS subject classifications. 15A24, 65F10, 82C70

1. Introduction. In transport theory, we encounter nonsymmetric algebraic Riccati equations of the form

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

(see [10]), where $A, B, C, D \in \mathbb{R}^{n\times n}$ have the following structures:

$$A = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_n) - eq^T, \tag{1.2}$$

$$B = ee^T, \tag{1.3}$$

$$C = qq^T, \tag{1.4}$$

and

$$D = \mathrm{diag}(d_1, d_2, \ldots, d_n) - qe^T. \tag{1.5}$$

In the above,

$$\delta_i = \frac{1}{cw_i(1+\alpha)}, \quad d_i = \frac{1}{cw_i(1-\alpha)}, \tag{1.6}$$

and

$$e = (1, 1, \ldots, 1)^T, \quad q = (q_1, q_2, \ldots, q_n)^T \text{ with } q_i = \frac{c_i}{2w_i}, \tag{1.7}$$

where $0 < c \le 1,\ 0 \le \alpha < 1$, and

$$0 < w_n < \cdots < w_2 < w_1 < 1,$$

$$\sum_{i=1}^{n} c_i = 1,\ c_i > 0,\ i = 1, 2, \ldots, n.$$

*This research was supported in part by National Science Foundation grant ECS-9633326.

†Department of Computer Science, University of California, Davis, One Shields Avenue, Davis, CA 95616-8562 (chguo@math.uregina.ca). Current address: Department of Mathematics and Statistics, University of Regina, Regina, SK S4S 0A2, Canada.

‡College of Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616-5294 (laub@ucdavis.edu).

For descriptions on how these equations arise in transport theory, see [10] and references cited therein. Here we only note that the constants c and α have physical meanings and the constants c_i and w_i appear in a numerical quadrature formula of the form $\int_0^1 f(w)dw \approx \sum_{i=1}^n c_i f(w_i)$.

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. We can then define positive matrices, nonnegative matrices, etc. The existence of positive solutions of (1.1) has been shown in [9] and [10]. However, only the minimal positive solution is physically meaningful.

The minimal positive solution of (1.1) can be found by basic fixed-point iterations (see [9], for example). It is mentioned in [10] that the convergence of these fixed-point iterations can be very slow when $c \approx 1$ and $\alpha \approx 0$. In [10], the minimal positive solution of (1.1) is constructed explicitly. The solution formula needs all the zeros of a certain secular equation. To get a good approximation of the minimal positive solution, the secular equation must be solved very accurately. We note that Newton's method is not always valid as a correction method when $c \approx 1$ and $\alpha \approx 0$. This point will be made clear in later discussions.

General nonsymmetric algebraic Riccati equations of the form

$$\mathcal{R}(X) = XCX - XD - AX + B = 0, \tag{1.8}$$

where A, B, C, D are real matrices of sizes $m \times m, m \times n, n \times m, n \times n$, respectively, have also been studied in the literature. See [18], for example. All the solutions of (1.8) can be found, in theory, by finding all the Jordan chains of the matrix

$$H = \begin{pmatrix} D & -C \\ B & -A \end{pmatrix} \tag{1.9}$$

(see Theorem 7.1.2 of [14]). Iterative methods have also been studied for the solution of (1.8). For example, a convergence result for Newton's method is given in [4] under a certain condition on the matrices A, B, C, and D.

Iterative methods with good convergence properties are not available for (1.8) in its full generality. However, for a certain class of these equations, a fairly complete theory can be established for Newton's method and a class of basic fixed-point iterations. Our paper is devoted to the study of these iterative methods.

We start with some definitions. A real square matrix A is called a Z-matrix if all its off-diagonal elements are nonpositive. It is clear that any Z-matrix A can be written as $sI - B$ with $B \geq 0$. A Z-matrix A is called an M-matrix if $s > \rho(B)$, where $\rho(\cdot)$ is the spectral radius. It is called a singular M-matrix if $s = \rho(B)$. Note that A is an M-matrix if and only if A^T is so. Note also that a singular M-matrix is indeed singular ($\rho(B)$ is an eigenvalue of B by the theory of nonnegative matrices; see [21], for example).

The following result is well known (see [2] and [5], for example).

THEOREM 1.1. *For a Z-matrix A, the following are equivalent:*

1. *A is an M-matrix.*
2. *$A^{-1} \geq 0$.*
3. *$Av > 0$ for some vector $v > 0$.*
4. *All eigenvalues of A have positive real parts.*

The next result is also standard (see [17], for example).

THEOREM 1.2. *Let $A \in \mathbb{R}^{n\times n}$ be an M-matrix. If the elements of $B \in \mathbb{R}^{n\times n}$ satisfy the relations*

$$b_{ii} \geq a_{ii}, \quad a_{ij} \leq b_{ij} \leq 0, \quad i \neq j,\ 1 \leq i, j \leq n,$$

then B is also an M-matrix.

In this paper we consider nonsymmetric algebraic Riccati equations (1.8) with the following conditions:

$$B > 0, \quad C > 0, \quad I \otimes A + D^T \otimes I \text{ is an } M\text{-matrix}, \tag{1.10}$$

where $\otimes$ is the Kronecker product (for basic properties of the Kronecker product, see [15], for example).

Remark 1.1. It is clear that $I \otimes A + D^T \otimes I$ is a Z-matrix if and only if both A and D are Z-matrices. Since any eigenvalue of $I \otimes A + D^T \otimes I$ is the sum of an eigenvalue of A and an eigenvalue of D (see [15], for example), it follows from the equivalence of 1. and 4. in Theorem 1.1 that $I \otimes A + D^T \otimes I$ is an M-matrix when A, D are both M-matrices. That the converse is not true is shown by $A = I$ and $D = 0$.

The matrices A and D in (1.1) are both M-matrices by Theorem 1.1 since $Aw > 0$ and $D^T w > 0$ for $w = (w_1, w_2, \ldots, w_n)^T$. Therefore, (1.1) with A, B, C, D defined by (1.2)–(1.7) is a special case of (1.8) with the conditions in (1.10).

From now on, when we speak of (1.8), we always assume that the conditions in (1.10) are satisfied.

2. Newton's method. We now consider the application of Newton's method to the Riccati equation (1.8). For any matrix norm $\mathbb{R}^{m\times n}$ is a Banach space, and the Riccati function $\mathcal{R}$ is a mapping from $\mathbb{R}^{m\times n}$ into itself. The first Fréchet derivative of $\mathcal{R}$ at a matrix X is a linear map $\mathcal{R}_X^{'} : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ given by

$$\mathcal{R}'_X(Z) = -\big((A - XC)Z + Z(D - CX)\big). \tag{2.1}$$

Also, the second derivative at X, $\mathcal{R}_X^{''} : \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$, is given by

$$\mathcal{R}_X^{''}(Z_1, Z_2) = Z_1 C Z_2 + Z_2 C Z_1. \tag{2.2}$$

The Newton method for the solution of (1.8) is

$$X_{i+1} = X_i - (\mathcal{R}_{X_i}^{'})^{-1}\mathcal{R}(X_i),\ i = 0, 1, \ldots, \tag{2.3}$$

given that the maps $\mathcal{R}_{X_i}^{'}$ are all invertible. In view of (2.1), the iteration (2.3) is equivalent to

$$(A - X_i C)X_{i+1} + X_{i+1}(D - CX_i) = B - X_i C X_i,\ i = 0, 1, \ldots. \tag{2.4}$$

THEOREM 2.1. *If there is a positive matrix X such that $\mathcal{R}(X) \le 0$, then* (1.8) *has a positive solution S such that $S \le X$ for every positive matrix X for which $\mathcal{R}(X) \le 0$. In particular, S is the minimal positive solution of* (1.8). *For the Newton iteration* (2.3) *with $X_0 = 0$, the sequence $\{X_i\}$ is well defined, $X_0 < X_1 < \cdots$, and* $\lim X_i = S$. *Furthermore, the matrix S is such that*

$$M_S = I \otimes (A - SC) + (D - CS)^T \otimes I$$

is either an M-matrix or a singular M-matrix.

Proof. Let X be any positive matrix such that

$$XCX - XD - AX + B \le 0. \tag{2.5}$$

For the Newton iteration (2.4) with $X_0 = 0$, we have

$$AX_1 + X_1D = B.$$

This equation is equivalent to

$$(I \otimes A + D^T \otimes I)\mathrm{vec}X_1 = \mathrm{vec}B, \tag{2.6}$$

where the vec operator stacks the columns of a matrix into one long vector (see [14, p. 99], for example). Since $I \otimes A + D^T \otimes I$ is an M-matrix by assumption, we get from (2.6) that $\mathrm{vec}X_1 > 0$, i.e., $X_1 > 0$. Therefore, the statement

$$X_k < X_{k+1},\ X_k < X,\ I \otimes (A - X_kC) + (D - CX_k)^T \otimes I \text{ is an } M\text{-matrix} \tag{2.7}$$

is true for $k = 0$.

We now assume that (2.7) is true for $k = i \geq 0$. By (2.4) and (2.5) we have

$$\begin{aligned} &(A - X_iC)(X_{i+1} - X) + (X_{i+1} - X)(D - CX_i) \\ &= B - X_iCX_i - AX + X_iCX - XD + XCX_i \\ &\leq -(X - X_i)C(X - X_i). \end{aligned} \tag{2.8}$$

Since $X_i < X$ and $I \otimes (A - X_iC) + (D - CX_i)^T \otimes I$ is an M-matrix, it follows from (2.8) that $X_{i+1} < X$. By (2.4)

$$\begin{aligned} &(A - X_{i+1}C)X_{i+1} + X_{i+1}(D - CX_{i+1}) \\ &= \big(A - X_iC - (X_{i+1} - X_i)C\big)X_{i+1} + X_{i+1}\big(D - CX_i - C(X_{i+1} - X_i)\big) \\ &= B - (X_{i+1} - X_i)C(X_{i+1} - X_i) - X_{i+1}CX_{i+1}. \end{aligned} \tag{2.9}$$

It follows from (2.9) and (2.5) that

$$\begin{aligned} &(A - X_{i+1}C)(X_{i+1} - X) + (X_{i+1} - X)(D - CX_{i+1}) \\ &\leq -(X_{i+1} - X_i)C(X_{i+1} - X_i) - (X_{i+1} - X)C(X_{i+1} - X) < 0. \end{aligned}$$

Therefore,

$$\big(I \otimes (A - X_{i+1}C) + (D - CX_{i+1})^T \otimes I\big)\mathrm{vec}(X - X_{i+1}) > 0.$$

Thus $I \otimes (A - X_{i+1}C) + (D - CX_{i+1})^T \otimes I$ is an M-matrix by Theorem 1.1. By (2.9) and (2.4)

$$\begin{aligned} &(A - X_{i+1}C)(X_{i+1} - X_{i+2}) + (X_{i+1} - X_{i+2})(D - CX_{i+1}) \\ &= -(X_{i+1} - X_i)C(X_{i+1} - X_i) < 0. \end{aligned}$$

Therefore, $X_{i+1} < X_{i+2}$. We have thus proved that (2.7) is true for $k = i + 1$. Hence, by the principle of mathematical induction, (2.7) is true for all $k \geq 0$. The Newton sequence is now well defined, monotonically increasing, and bounded above. Let $\lim_{k\to\infty} X_k = S$. Then S is a solution of (1.8) by (2.4). Since $S \leq X$ for any X such that $\mathcal{R}(X) \leq 0$, S is the minimal positive solution of (1.8). For all $i \geq 0$, we can write $I \otimes (A - X_iC) + (D - CX_i)^T \otimes I = rI - T_i$ with $T_i \geq 0$ and $\rho(T_i) < r$. Now, $M_S = rI - T$ with $T = \lim_{i\to\infty} T_i$. Since $\rho(T) \leq r$, the matrix M_S is either an M-matrix or a singular M-matrix. □

Remark 2.1. The above result is similar in nature to Theorem 9.1.1 of [14]. The result is also somewhat related to a monotone convergence result on Newton's method for convex operators in partially ordered spaces, as described in Theorem 5.1 of [20]. In order to apply that theorem, we need to know that there is a positive matrix X such that $\mathcal{R}(X) \le 0$ and $I \otimes (A - XC) + (D - CX)^T \otimes I$ is an M-matrix. When this is true, that theorem implies that the Newton sequence with $X_0 = 0$ is well defined, $X_0 \le X_1 \le \cdots$, and $\lim X_k = X^* \le X$ is a solution of $\mathcal{R}(X) = 0$. With the hindsight from Theorem 2.1, such a positive matrix X does not exist if $I \otimes (A - SC) + (D - CS)^T \otimes I$ is a singular M-matrix for the minimal positive solution S. In fact, the existence of such an X would imply $S \le X$ by Theorem 2.1, which would in turn imply that $I \otimes (A - SC) + (D - CS)^T \otimes I$ is an M-matrix by Theorem 1.2.

Remark 2.2. Even if A and D are both M-matrices, it is not necessarily true that $A - SC$ and $D - CS$ are both M-matrices or singular M-matrices. This is shown by the scalar case with $B = C = 1$, $D = 1/2$, and $A = 3/2$. For this example, $S = 1$, $A - SC = 1/2$, and $D - CS = -1/2$. This example also shows that the matrix M_S in Theorem 2.1 can indeed be a singular M-matrix.

The following comparison result is an immediate consequence of Theorem 2.1.

COROLLARY 2.2. *Let S be the minimal solution of* (1.8). *If any element of B or C decreases but remains positive, or if any diagonal element of $I \otimes A + D^T \otimes I$ increases, or if any off-diagonal element of $I \otimes A + D^T \otimes I$ increases but remains nonpositive, then the equation so obtained also has a minimal positive solution $\tilde{S}$. Moreover, $\tilde{S} \le S$.*

Proof. Let the new equation be

$$\tilde{\mathcal{R}}(X) = X\tilde{C}X - X\tilde{D} - \tilde{A}X + \tilde{B} = 0.$$

It is clear that $\tilde{\mathcal{R}}(S) \le 0$. Since $I \otimes \tilde{A} + \tilde{D}^T \otimes I$ is still an M-matrix by Theorem 1.2, the conclusions follow from Theorem 2.1. □

Remark 2.3. As an easy consequence of the above corollary, we can conclude that the minimal positive solution of (1.1) increases in c. In [10], it is also concluded that the minimal solution decreases in α. This conclusion is not a consequence of the above corollary and is, in fact, not valid.

Example 2.1. Consider the Riccati equation (1.1) with $n = 2$ and

$$c_1 = c_2 = 1/2, \quad w_1 = 3/4, \quad w_2 = 1/4, \quad c = 1/2.$$

If $\alpha = 0.1$, then the minimal solution (to four digits without rounding) is

$$\begin{pmatrix} 0.2758 & 0.1196 \\ 0.1344 & 0.0766 \end{pmatrix}.$$

If $\alpha = 0.2$, then the minimal solution (to four digits without rounding) is

$$\begin{pmatrix} 0.2639 & 0.1087 \\ 0.1372 & 0.0746 \end{pmatrix}.$$

This example shows the minimal solution does not necessarily decrease in α.

As to the convergence rate of Newton's method, the following result is immediate.

THEOREM 2.3. *If the matrix M_S in Theorem* 2.1 *is an M-matrix, then for $X_0 = 0$ the Newton sequence $\{X_k\}$ converges to S quadratically.*

Proof. If M_S is an M-matrix, then the Fréchet derivative $\mathcal{R}_S^{'}$ is an invertible map. Since $\mathcal{R}$ is a smooth function, the convergence of the Newton sequence must be quadratic (see [11] and [19], for example). ∎

If the matrix M_S is a singular M-matrix, the map $\mathcal{R}_S^{'}$ is not invertible and the convergence of Newton's method is more complicated. The convergence behavior of Newton's method in this case will be clarified by following the strategy used in [8] for symmetric algebraic Riccati equations and using a theorem on Newton's method at singular points (see [3, Theorem 1.2] and [12, Theorem 1.1], for example).

LEMMA 2.4. *If M_S is a singular M-matrix, then 0 is a simple eigenvalue of M_S. Let $\mathcal{N} = \mathrm{Ker}(\mathcal{R}_S^{'})$ and $\mathcal{M} = \mathrm{Im}(\mathcal{R}_S^{'})$. Then $\mathcal{N}$ is one-dimensional, $\mathbb{R}^{m\times n} = \mathcal{N} \oplus \mathcal{M}$, and the map $\mathcal{B} : \mathcal{N} \to \mathcal{N}$ given by*

$$\mathcal{B}(N) = P_{\mathcal{N}} \mathcal{R}_S^{''}(N_0, N)$$

is invertible for nonzero $N_0 \in \mathcal{N}$, where $P_{\mathcal{N}}$ is the projection on the null space $\mathcal{N}$ parallel to the range $\mathcal{M}$.

Proof. We write $M_S = rI - T$ with $T \geq 0$ and $\rho(T) = r > 0$. Since T is clearly irreducible, we know by the Perron-Frobenius Theorem (see [21]) that $\rho(T)$ is a simple eigenvalue of T with a positive eigenvector. Thus, we can find mn orthonormal vectors $u_1, u_2, \ldots, u_{mn}$ such that $u_1 > 0$ and

$$U^{-1} M_S U = \begin{pmatrix} 0 & 0 \\ 0 & M_{22} \end{pmatrix}, \tag{2.10}$$

where $U = (u_1\ u_2\ \cdots u_{mn})$ and M_{22} is an $(mn-1)\times(mn-1)$ nonsingular matrix. Now, $\mathcal{R}_S^{'}(N) = -(A - SC)N - N(D - CS) = 0$ if and only if $M_S \mathrm{vec}N = 0$. In view of (2.10), $M_S \mathrm{vec}N = 0$ if and only if $\mathrm{vec}N = U(a, 0, \ldots, 0)^T = a\, u_1$ for some $a \in \mathbb{R}$, in which case we write $N = a\,\mathrm{unvec}u_1$ (i.e., the unvec operator is the inverse of the vec operator). Thus $\mathcal{N} = \{a\,\mathrm{unvec}u_1 \mid a \in \mathbb{R}\}$. Similarly, $\mathcal{M} = \{b_2 \mathrm{unvec}u_2 + \cdots + b_{mn}\mathrm{unvec}u_{mn} \mid b_2, \ldots, b_{mn} \in \mathbb{R}\}$. Therefore, $\mathcal{N}$ is one-dimensional and $\mathbb{R}^{m\times n} = \mathcal{N} \oplus \mathcal{M}$. To prove the map $\mathcal{B}$ is invertible, we only need to show $P_{\mathcal{N}}(\mathrm{unvec}u_1 C \mathrm{unvec}u_1) \neq 0$ (see (2.2)). Since $u_1 > 0$ and $\mathrm{vec}(\mathrm{unvec}u_1 C \mathrm{unvec}u_1) = k_1u_1 + k_2u_2 + \cdots + k_{mn}u_{mn}$ for some real numbers $k_1, k_2, \ldots, k_{mn}$, we have

$$k_1 = u_1^T \mathrm{vec}(\mathrm{unvec}u_1 C \mathrm{unvec}u_1) > 0.$$

Thus, $P_{\mathcal{N}}(\mathrm{unvec}u_1 C \mathrm{unvec}u_1) = k_1 \mathrm{unvec}u_1 \neq 0$, as required. ∎

LEMMA 2.5. *For any fixed $\theta > 0$, let*

$$Q = \{i :\ \|P_{\mathcal{M}}(X_i - S)\| > \theta \|P_{\mathcal{N}}(X_i - S)\|\}.$$

Then there exist an integer i_0 and a constant $\eta > 0$ such that $\|X_{i+1} - S\| \leq \eta \|X_i - S\|^2$ for all i in Q for which $i \geq i_0$.

Proof. The proof is analogous to that of [8, Theorem 2.2], although the algebraic Riccati equations considered in that paper are different from the Riccati equations being considered here. ∎

COROLLARY 2.6. *Assume that , for given $\theta > 0$, $\|P_{\mathcal{M}}(X_i - S)\| > \theta \|P_{\mathcal{N}}(X_i - S)\|$ for all i large enough. Then $X_i \to S$ quadratically.*

We are now ready to clarify the convergence behavior of Newton's method when the matrix M_S is a singular M-matrix.

THEOREM 2.7. *If M_S is a singular M-matrix and the convergence of the Newton sequence $\{X_i\}$ in Theorem* 2.1 *is not quadratic, then $\|(\mathcal{R}'_{X_i})^{-1}\| \le \beta\|X_i - S\|^{-1}$ for all $i \ge 1$ and some constant $\beta > 0$. Moreover,*

$$\lim_{i\to\infty} \frac{\|X_{i+1} - S\|}{\|X_i - S\|} = \frac{1}{2}, \quad \lim_{i\to\infty} \frac{\|P_{\mathcal{M}}(X_i - S)\|}{\|P_{\mathcal{N}}(X_i - S)\|^2} = 0.$$

Proof. The result follows from Theorem 2.1, Lemma 2.4, Corollary 2.6, and [12, Theorem 1.1]. □

3. A class of fixed-point iterations. If we write

$$A = A_1 - A_2, \quad D = D_1 - D_2,$$

(1.8) becomes

$$A_1X + XD_1 = XCX + XD_2 + A_2X + B.$$

We use only those splittings of A and D such that $A_2, D_2 \ge 0$, and A_1 and D_1 are Z-matrices. In these situations, the matrix $I \otimes A_1 + D_1^T \otimes I$ is an M-matrix by Theorem 1.2. We then have a class of fixed-point iterations

$$X_{k+1} = \mathcal{L}^{-1}(X_kCX_k + X_kD_2 + A_2X_k + B), \tag{3.1}$$

where the linear operator $\mathcal{L}$ is given by

$$\mathcal{L}(X) = A_1X + XD_1.$$

Since $I \otimes A_1 + D_1^T \otimes I$ is an M-matrix, the operator $\mathcal{L}$ is invertible and $\mathcal{L}^{-1}(X) > 0$ for $X > 0$.

THEOREM 3.1. *If $\mathcal{R}(X) \le 0$ for some positive matrix X, then for the fixed-point iterations* (3.1) *and $X_0 = 0$, we have for any $k \ge 1$*

$$X_0 < X_1 < \cdots < X_k < X. \tag{3.2}$$

Moreover, $\lim_{k\to\infty} X_k = S$.

Proof. The order relation (3.2) can easily be proved by induction. The limit X^* is then a solution of $\mathcal{R}(X) = 0$ and must be the minimal positive solution S, since $X^* \le X$ for any positive matrix X such that $\mathcal{R}(X) \le 0$. □

Remark 3.1. The comparison result on the minimal positive solution (Corollary 2.2) also follows from the above simple result.

The following result is concerned with the convergence rates of these fixed-point iterations.

THEOREM 3.2. *For the fixed-point iterations* (3.1) *with $X_0 = 0$, we have*

$$\limsup_{k\to\infty} \sqrt[k]{\|X_k - S\|} = \rho\big((I \otimes A_1 + D_1^T \otimes I)^{-1}(I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)\big).$$

Proof. By a theorem on general fixed-point iterations (see [13, p. 21], for example), we have

$$\limsup_{k\to\infty} \sqrt[k]{\|X_k - S\|} \le \rho(\mathcal{G}'_S), \tag{3.3}$$

where $\mathcal{G}_S^{'}$ is the Fréchet derivative at S of the map $\mathcal{G}$ given by

$$\mathcal{G}(X) = \mathcal{L}^{-1}(XCX + XD_2 + A_2X + B).$$

It is easily found that $\mathcal{G}_S^{'}$ is given by

$$\mathcal{G}_S^{'}(H) = \mathcal{L}^{-1}\big((A_2 + SC)H + H(D_2 + CS)\big).$$

We now show that, in fact, equality holds in (3.3). We may assume the norm in (3.3) is the Frobenius norm.

Let $E_k = S - X_k$. We have $E_{k+1} = P_k(E_k)$, where the operator P_k is given by

$$P_k(H) = \mathcal{L}^{-1}\big((A_2 + SC)H + H(D_2 + CX_k)\big). \tag{3.4}$$

Note that $\lim_{k\to\infty} P_k = \mathcal{G}_S^{'}$. Thus, for any $\epsilon > 0$, we can find an integer l such that

$$\rho(P_l) \geq \rho(\mathcal{G}_S^{'}) - \epsilon.$$

Now, since $0 = X_0 < X_1 < \cdots$, we have

$$\begin{aligned}\limsup_{k\to\infty} \sqrt[k]{\|X_k - S\|} &= \limsup_{k\to\infty} \sqrt[k]{\|P_{k-1}\cdots P_l P_{l-1}\cdots P_0(S)\|} \\ &\geq \limsup_{k\to\infty} \sqrt[k]{\|(P_l)^{k-l}(P_0)^l(S)\|}.\end{aligned}$$

Since $(P_0)^l(S) > 0$, we have $(P_0)^l(S) > c_l E$, where $c_l > 0$ is a constant and E is the matrix with all its elements equal to one. Also, $\|(P_l)^{k-l}\| = \|(P_l)^{k-l}(S_{l,k})\|$, where $S_{l,k} \in \mathbb{R}^{m\times n}$ is such that $\|S_{l,k}\| = 1$ and $S_{l,k} \geq 0$. Now,

$$\begin{aligned}\limsup_{k\to\infty} \sqrt[k]{\|X_k - S\|} &\geq \limsup_{k\to\infty} \sqrt[k]{\|c_l(P_l)^{k-l}(E)\|} \\ &\geq \limsup_{k\to\infty} \sqrt[k]{c_l\|(P_l)^{k-l}(S_{l,k})\|} \\ &= \limsup_{k\to\infty} \sqrt[k]{\|(P_l)^{k-l}\|} \\ &= \rho(P_l) \geq \rho(\mathcal{G}_S^{'}) - \epsilon.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have

$$\limsup_{k\to\infty} \sqrt[k]{\|X_k - S\|} = \rho(\mathcal{G}_S^{'}).$$

A number λ is an eigenvalue of $\mathcal{G}_S^{'}$ if and only if for some $H \neq 0$,

$$\mathcal{L}^{-1}\big((A_2 + SC)H + H(D_2 + CS)\big) = \lambda H,$$

which is the same as

$$(A_2 + SC)H + H(D_2 + CS) = \lambda(A_1H + HD_1).$$

or

$$(I \otimes A_1 + D_1^T \otimes I)^{-1}(I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)\mathrm{vec}H = \lambda \mathrm{vec}H.$$

Thus,

$$\rho(\mathcal{G}_S^{'}) = \rho\big((I \otimes A_1 + D_1^T \otimes I)^{-1}(I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)\big).$$

This completes the proof. ∎

We can say something more about the spectral radius in Theorem 3.2.

THEOREM 3.3. *If M_S is a singular M-matrix, then*

$$\rho\big((I \otimes A_1 + D_1^T \otimes I)^{-1}(I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)\big) = 1.$$

If M_S is an M-matrix, and $A = \tilde{A}_1 - \tilde{A}_2$, $D = \tilde{D}_1 - \tilde{D}_2$ are such that $0 \le \tilde{A}_2 \le A_2$ and $0 \le \tilde{D}_2 \le D_2$, then

$$\begin{aligned}&\rho\big((I \otimes \tilde{A}_1 + \tilde{D}_1^T \otimes I)^{-1}(I \otimes (\tilde{A}_2 + SC) + (\tilde{D}_2 + CS)^T \otimes I)\big)\\ &\le \rho\big((I \otimes A_1 + D_1^T \otimes I)^{-1}(I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)\big) < 1.\end{aligned}$$

Proof. Since

$$M_S = (I \otimes A_1 + D_1^T \otimes I) - (I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)$$

and

$$M_S = (I \otimes \tilde{A}_1 + \tilde{D}_1^T \otimes I) - (I \otimes (\tilde{A}_2 + SC) + (\tilde{D}_2 + CS)^T \otimes I)$$

are regular splittings [21] of M_S, the second conclusion follows from the standard results in [21]. If M_S is a singular M-matrix, then $M_S v = 0$ for some $v \neq 0$. Thus,

$$(I \otimes A_1 + D_1^T \otimes I)^{-1}(I \otimes (A_2 + SC) + (D_2 + CS)^T \otimes I)v = v,$$

and the first conclusion follows. ∎

Therefore, the convergence of these iterations is linear if M_S is an M-matrix. When M_S is a singular M-matrix, the convergence is sublinear. Within this class of iterative methods, three iterations are worthy of special mention. The first one is obtained when we take A_1 and D_1 to be the diagonal part of A and D, respectively. This is the simplest iteration in the class and will be called FP1. The second one is obtained when we take A_1 to be the lower triangular part of A and take D_1 to be the upper triangular part of D. This iteration will be called FP2. The last one is obtained when we take $A_1 = A$ and $D_1 = D$. This is the fastest iteration in this class (see second part of Theorem 3.3) and will be called FP3.

4. Some practical issues and an overall algorithm. If (1.8) has a positive solution, the minimal positive solution can thus be found by the Newton iteration or some basic fixed-point iterations. Starting with the zero matrix, each of these iterations produces a monotonically increasing sequence, the limit of which is the minimal positive solution S. The matrix M_S associated with S is either an M-matrix or a singular M-matrix. When M_S is an M-matrix, the convergence of Newton's method is quadratic and the convergence of the basic fixed-point iterations is linear. When M_S is a singular M-matrix, the convergence of Newton's method is at least linear and the convergence of the basic fixed-point iterations is sublinear. Therefore, Newton's method is always much faster than the other methods in terms of iteration counts. It must be noted, however, that the computational work involved in one step

of Newton's method is much higher than that involved in one step of a basic fixed-point iteration. For the Newton iteration (2.3), the equation $-\mathcal{R}'_{X_k}(H) = \mathcal{R}(X_k)$, i.e., $(A - X_kC)H + H(D - CX_k) = \mathcal{R}(X_k)$, can be solved by the algorithms described in [1] and [6]. If we use the Bartels-Stewart algorithm [1] to solve the Sylvester equation, the computational work for each Newton iteration is about $62n^3$ flops when $m = n$ (see [7] for the definition of a "flop"). By comparison, FP1 and FP2 need about $8n^3$ flops for each iteration. For FP3 we can use the Bartels-Stewart algorithm for the first iteration. It needs about $54n^3$ flops. For each subsequent iteration, it needs about $14n^3$ flops.

For the basic fixed-point iteration (3.1), the error reduction at the $(k+1)th$ step is determined by the operator P_k in (3.4). Since $0 = X_0 < X_1 < \cdots$, we can see that the error reduction is more significant initially. For Newton's method, of course, the error reduction is much more significant at a late stage of iteration unless the matrix M_S is nearly singular. It is therefore advisable to start with some basic fixed-point iteration and switch to Newton's method after the residual error has been reduced to a certain level. From Theorem 2.1 we know that Newton's method, starting with the zero matrix, produces a monotonically increasing sequence. Now, with the initial guess produced by some basic fixed-point iteration, will the Newton sequence still be monotonic?

PROPOSITION 4.1. *Assume that $\mathcal{R}(X) \leq 0$ for some positive matrix X. If $\{X_k\}_{k=1}^{k_0}$ is produced by basic fixed-point iteration* (3.1) *with $X_0 = 0$ and $\{X_k\}_{k=k_0+1}^{\infty}$ is produced by Newton's method with X_{k_0} as an initial guess, then*

$$0 < X_1 < X_2 < \cdots < X_{k_0} < X_{k_0+1} < \cdots,$$

and $\lim_{k\to\infty} X_k = S$, the minimal positive solution.

Proof. We already know from Theorem 3.1 that $0 < X_1 < X_2 < \cdots < X_{k_0} < S$. Now, for $1 \leq k \leq k_0$, we have

$$\begin{aligned}\mathcal{R}(X_k) &= X_kCX_k + X_kD_2 + A_2X_k + B - A_1X_k - X_kD_1 \\ &= X_kCX_k - X_{k-1}CX_{k-1} + (X_k - X_{k-1})D_2 + A_2(X_k - X_{k-1}) > 0.\end{aligned}$$

Since X_{k_0+1} is obtained from X_{k_0} by Newton's method, $-\mathcal{R}'_{X_{k_0}}(X_{k_0+1} - X_{k_0}) = \mathcal{R}(X_{k_0})$. Thus, $(A-X_{k_0}C)(X_{k_0+1}-X_{k_0})+(X_{k_0+1}-X_{k_0})(D-CX_{k_0}) > 0$. Since $X_{k_0} < S$ and $I \otimes (A-SC)+(D-CS)^T \otimes I$ is either an M-matrix or a singular M-matrix, it follows from the Perron-Frobenius Theorem that $I \otimes (A - X_{k_0}C) + (D - CX_{k_0})^T \otimes I$ is an M-matrix. Therefore, $X_{k_0+1} > X_{k_0}$. Once this is proved, it follows as in the proof of Theorem 2.1 that $X_{k_0} < X_{k_0+1} < \cdots$, and $\lim_{k\to\infty} X_k = S$. □

Remark 4.1. We may apply the above strategy without knowing whether (1.8) has a positive solution. If we find that $X_k < X_{k+1}$ is not true for some $k \geq k_0$, then we can conclude that (1.8) does not have a positive solution. Note however that $X_k < X_{k+1}$ is true for all $0 \leq k < k_0$, even if (1.8) has no positive solutions. This is another difference between Newton's method and the basic fixed-point iterations.

Remark 4.2. The results in Section 2 are still valid when the Newton iteration is started with a matrix produced by a basic fixed-point iteration as in Proposition 4.1.

The convergence behavior of the iterative methods we have discussed depends on the matrix $M_S = I \otimes (A - SC) + (D - CS)^T \otimes I$, in which S is the minimal positive solution to be found. The matrix M_S is a singular M-matrix if and only if $\lambda_i + \mu_j = 0$ for some eigenvalue λ_i of $A - SC$ and some eigenvalue μ_j of $D - CS$. There is some connection between the eigenvalues of $A - SC$ (or $D - CS$) and the eigenvalues of the matrix H in (1.9). In fact, the following result is true.

PROPOSITION 4.2. *If X is any solution of* (1.8)*, then any eigenvalue of $D-CX$ is an eigenvalue of H and any eigenvalue of $A-XC$ is the negative of some eigenvalue of H.*

Proof. It is easy to verify that

$$\begin{pmatrix} I & 0 \\ X & I \end{pmatrix}^{-1} \begin{pmatrix} D & -C \\ B & -A \end{pmatrix} \begin{pmatrix} I & 0 \\ X & I \end{pmatrix} = \begin{pmatrix} D-CX & -C \\ 0 & -(A-XC) \end{pmatrix}.$$

The conclusions follow immediately. ∎

However, when we are going to use iterative methods to find the minimal positive solution, we would not bother to find all the eigenvalues of the matrix H. Even if we know all the eigenvalues of H, Proposition 4.2 is not adequate to determine all the eigenvalues of $A-SC$ and $D-CS$. For (1.1), we know from the results in [10] that M_S is a singular M-matrix if and only if $c=1$ and $\alpha=0$. This explains why Newton's method may not be valid as a correction method when $c\approx 1$ and $\alpha\approx 0$. For (1.8), whether the matrix M_S is a singular M-matrix (or nearly so) can be inferred from the speed of convergence of the iterative method we are using. For example, very slow convergence of a basic fixed-point iteration indicates that the matrix M_S is a singular M-matrix or nearly so. By Proposition 4.1 we can always use the Newton iteration when the convergence of the fixed-point iteration is unsatisfactory.

When the matrix M_S is singular and the convergence of Newton's method is not quadratic, we know from Theorem 2.7 that the convergence must be linear with rate 1/2 and the error will rapidly be dominated by the null space component. As is the case for symmetric algebraic Riccati equations (see Theorems 3.1 and 3.2 of [8]), very accurate approximation for the minimal positive solution can be obtained by computing $Y_{k+1}=X_k-2(\mathcal{R}'_{X_k})^{-1}\mathcal{R}(X_k)$ when $\|P_{\mathcal{M}}(X_k-S)\|\le\epsilon\|P_{\mathcal{N}}(X_k-S)\|$ and ϵ is very small. Note that $\|X_k-S\|$ need not be very small when ϵ is very small. In this case, the Sylvester equation $-\mathcal{R}'_{X_k}(H)=\mathcal{R}(X_k)$ is not nearly singular and can be solved by the Bartels-Stewart algorithm very accurately. Without this double Newton step, Newton's method will take many more iterations. Even linear convergence with rate 1/2 can fail to be realized due to a nearly singular Jacobian at a late stage. Therefore, when we apply Newton's method, we can *try* a double Newton step first. If the approximation obtained fails to satisfy a given stopping criterion, then we use the original Newton iteration instead and try a double Newton step with the new iterate, i.e., we have an algorithm similar to Algorithm 3.3 of [8] for symmetric algebraic Riccati equations. Although the added cost of trying the double Newton step is minor, the strategy can be used in a wiser way. That is, we can try the double Newton step only when there are indications that we are solving a problem with $\mathcal{R}'_S$ singular (or nearly singular) and that the error is already essentially in the null space (or approximate null space). The next result shows how we can get such indications.

PROPOSITION 4.3. *Assume that $\mathcal{R}'_S$ is singular and $\{X_k\}_{k=k_0}^{\infty}$ is the Newton sequence in Proposition* 4.1*. If $X_k-S\in\mathcal{N}(k\ge k_0)$, then*

$$X_{k+1}-S=\frac{1}{2}(X_k-S),\quad \mathcal{R}(X_{k+1})=\frac{1}{4}\mathcal{R}(X_k).$$

Furthermore,

$$\lim_{r_k\to 0}\frac{\|X_{k+1}-S\|}{\|X_k-S\|}=\frac{1}{2},\quad \lim_{r_k\to 0}\frac{\mathcal{R}(X_k)}{\|X_k-S\|^2}=C_0, \tag{4.1}$$

where

$$r_k = \frac{\|P_{\mathcal{M}}(X_k - S)\|}{\|P_{\mathcal{N}}(X_k - S)\|^2}$$

and C_0 is a constant positive matrix. In particular,

$$\lim_{r_k \to 0} \frac{\|\mathcal{R}(X_{k+1})\|}{\|\mathcal{R}(X_k)\|} = \frac{1}{4}.$$

Proof. As in Theorem 3.1 of [8], we have $X_{k+1} - S = \frac{1}{2}(X_k - S) \in \mathcal{N}$ when $X_k - S \in \mathcal{N}$. Thus, in view of (2.2),

$$\begin{aligned}\mathcal{R}(X_{k+1}) &= \mathcal{R}(S) + \mathcal{R}_S^{'}(X_{k+1} - S) + \frac{1}{2}\mathcal{R}_S^{''}(X_{k+1} - S, X_{k+1} - S) \\ &= (X_{k+1} - S)C(X_{k+1} - S) = \frac{1}{4}(X_k - S)C(X_k - S) = \frac{1}{4}\mathcal{R}(X_k).\end{aligned}$$

If r_k is sufficiently small, we have as in Theorem 3.2 of [8]

$$\|X_k - 2(\mathcal{R}_{X_k}^{'})^{-1}\mathcal{R}(X_k) - S\| \le \gamma \frac{\|P_{\mathcal{M}}(X_k - S)\|}{\|P_{\mathcal{N}}(X_k - S)\|} \tag{4.2}$$

for some constant γ. Since the left-hand side of (4.2) can be written as $\|2(X_{k+1} - S) - (X_k - S)\|$, the first limit in (4.1) follows easily. Now, let $\mathcal{N} = \text{span}\{N_0\}$ with $N_0 > 0$ and $\|N_0\| = 1$. Since

$$\begin{aligned}\mathcal{R}(X_k) &= \mathcal{R}(S) + \mathcal{R}_S^{'}(X_k - S) + \frac{1}{2}\mathcal{R}_S^{''}(X_k - S, X_k - S) \\ &= \mathcal{R}_S^{'}(P_{\mathcal{M}}(X_k - S)) + (X_k - S)C(X_k - S),\end{aligned}$$

we get easily that

$$\lim_{r_k \to 0} \frac{\mathcal{R}(X_k)}{\|X_k - S\|^2} = N_0 C N_0 > 0.$$

The proof is thus complete. □

When $\mathcal{R}_S^{'}$ is singular, we know from Theorem 2.7 that $\lim_{k\to\infty} r_k = 0$ unless the convergence of Newton's method is quadratic. The above proposition tells us that we may choose to try the double Newton step with a current Newton iterate X_k only when $\|\mathcal{R}(X_k)\|/\|\mathcal{R}(X_{k-1})\| \approx 1/4$.

We now propose the following algorithm for finding the minimal positive solution S of (1.8) whenever it has a positive solution. The algorithm may also detect that the equation actually does not have a positive solution. The choices of the splittings and parameters in step 1 of the algorithm can be made according to the guidelines provided immediately after the algorithm.

ALGORITHM 4.4.

1. *Choose splittings $A = A_1 - A_2$ and $D = D_1 - D_2$;*
   choose parameters $k_0, \epsilon, \eta_1, \eta_2, \eta_3 > 0$.
2. *Set $X_0 = 0$, $T(X_0) = B$, $r_0 = \|B\|_\infty$.*
3. *For $k = 1, 2, \ldots$, do:*
   solve $A_1 X_k + X_k D_1 = T(X_{k-1})$;
   compute $\mathcal{R}(X_k)$, $r_k = \|\mathcal{R}(X_k)\|_\infty$;
   if $r_k/r_0 < \eta_1$ or $k \ge k_0$, goto 4;
   compute $T(X_k) = T(X_{k-1}) + \mathcal{R}(X_k)$.

4. For $p = k, k+1, \ldots,$ do:
solve $-\mathcal{R}'_{X_k}(H) = \mathcal{R}(X_k)$ for $H = (h_{ij})$;
if $h_{ij} < -\eta_2\|H\|_\infty$ for some (i,j), then stop (no solution);
compute $X_{p+1} = X_p + H$, $\mathcal{R}(X_{p+1})$, $r_{p+1} = \|\mathcal{R}(X_{p+1})\|_\infty$;
if $r_{p+1}/r_0 < \epsilon$, then stop and $S \approx X_{p+1}$;
if $|\frac{r_{p+1}}{r_p} - \frac{1}{4}| < \eta_3$, then
compute $Z = X_p + 2H$, $r = \|\mathcal{R}(Z)\|_\infty$;
if $r/r_0 < \epsilon$, then stop and $S \approx Z$.

In the above algorithm, we can select a particular basic fixed-point iteration by choosing proper splittings of A and D. Normally we can use FP1 or FP2. Although FP2 is faster in general, FP1 may take advantage of the structures in a specific equation more easily. In the algorithm, ϵ is the required precision and is usually much smaller than η_1. The small number η_2 is introduced to numerically check if $H > 0$ has been violated. This number should be related to the unit roundoff. The small number η_3 is used to determine if the double Newton step should be tried. A smaller η_3 should be used for a smaller ϵ. If (1.8) does not have a positive solution, the criterion $r_k/r_0 < \eta_1$ in step 3 of the algorithm may never be satisfied. In the algorithm, we have let k_0 be the maximal number of fixed-point iterations allowed. The nonexistence of a positive solution can often be detected by Newton's method in step 4. When (1.8) has a positive solution, the algorithm will produce a finite sequence approaching the minimal positive solution. The sequence is obtained by a fixed-point iteration followed by ordinary Newton's method, with the exception that the last term in the sequence is possibly obtained by the double Newton step. It should be noted that the matrix Z produced by the double Newton step is not used in subsequent Newton iterations.

5. Numerical results. We first give a simple example to illustrate the performance of the iterative methods we have studied.

TABLE 5.1
Iteration counts for Example 5.1, $\alpha = 6.0$

ϵ	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}	10^{-12}
NM	3	4	4	5	5	5
FP1	11	22	33	44	54	65
FP2	10	19	29	38	48	57
FP3	7	15	23	31	38	46

TABLE 5.2
Iteration counts for Example 5.1, $\alpha = 4.27$

ϵ	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}	10^{-12}
NM	5	7	8	9	9	10
FP1	40	245	533	822	1112	1402
FP2	36	222	480	739	998	1257
FP3	29	182	396	611	827	1042

Example 5.1. Consider (1.8) with $m = n = 2$ and

$$A = \begin{pmatrix} \alpha & -2 \\ -1 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 5 & -1 \\ -1 & 4 \end{pmatrix}.$$

TABLE 5.3
Iteration counts for Example 5.1, $\alpha = 4.267191$

ϵ	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}	10^{-12}
NM	5	8	11	14	15	15
FP1	40	450	4477	25328	54350	83603
FP2	37	414	4119	23000	49020	75239
FP3	29	335	3339	18899	40559	62395

For $\alpha = 4.26$, we apply Newton's method with $X_0 = 0$ and find

$$X_6 = \begin{pmatrix} 0.3865 & 0.4048 \\ 0.3583 & 0.2943 \end{pmatrix}, \quad X_7 = \begin{pmatrix} 0.3713 & 0.3872 \\ 0.3490 & 0.2836 \end{pmatrix}.$$

Since $X_6 < X_7$ is not true, the equation has no positive solutions in this case. Experiments show that the equation has a positive solution for $\alpha = 4.267191$. Thus, it has positive solutions for all $\alpha \geq 4.267191$ by Corollary 2.2. In Tables 5.1–5.3, we have recorded, for three values of α, the number of iterations needed to have $\|\mathcal{R}(X_k)\|_\infty < \epsilon$ for Newton's method (NM) and the three basic fixed-point iterations. For all four methods, we use $X_0 = 0$. From the tables, we can see that the three basic fixed-point iterations have similar efficiency. For $\alpha = 6.0$, the basic fixed-point iterations are still adequate. For $\alpha = 4.27$ and $\alpha = 4.267191$, however, the advantage of Newton's method is very clear. In all three cases, the basic fixed-point iterations are useful for initial error reduction. We may consider using Newton's method after a certain number of fixed-point iterations. However, other features of Algorithm 4.4 have no role to play, since the existence of a positive solution is known for each α and quadratic convergence of Newton's method is visible even for $\alpha = 4.267191$.

Example 5.2. We now consider (1.1) for $n = 64$ and $n = 128$. The constants c_i and w_i are given by a numerical quadrature formula on the interval $[0, 1]$, which is obtained by dividing $[0, 1]$ into $n/4$ subintervals of equal length and applying Gauss-Legendre quadrature with 4 nodes to each subinterval.

We apply Algorithm 4.4 with the splittings of A and D being those corresponding to FP1, and take $k_0 = 200$, $\epsilon = 10^{-12}$, $\eta_1 = 10^{-3}$, $\eta_2 = 10^{-6}$, and $\eta_3 = 10^{-6}$. For this example it is actually unnecessary to introduce the parameter η_2, since the existence of positive solutions has been guaranteed by the theoretical results in [9] and [10]. We carry out the computation for $n = 64$ and $n = 128$. The parameter pair (α, c) is taken to be $(0.5, 0.5), (10^{-8}, 0.999999), (10^{-14}, 1)$, and $(0, 1)$. The results are recorded in Tables 5.4–5.5. For example, when $n = 64$ and $(\alpha, c) = (0, 1)$, the residual is reduced to 0.9916D-03r_0 after 170 FP1 iterations (r_0 is the initial residual). The residual is then reduced to 0.4937D-05r_0 after 4 Newton iterations. The fifth Newton iteration fails to achieve the required accuracy, but the double Newton step (DN) works (it reduces the residual to 0.1763D-13r_0). The double Newton step is also tried with the fourth Newton iteration, but without success. For this example, $\mathcal{R}_S'$ is singular when $(\alpha, c) = (0, 1)$.

6. Conclusions. We have discussed the iterative solution of a class of nonsymmetric algebraic Riccati equations, which includes a class of algebraic Riccati equations arising in transport theory. The coefficient matrices of any equation in this larger class have a special sign structure. Using this structure and the theory of M-matrices, we have shown that Newton's method and a class of basic fixed-point iterations can be used to find its minimal positive solution whenever it has a positive solution. We have also proposed an overall algorithm for the solution of the nonsymmetric algebraic

TABLE 5.4
Convergence history for Example 5.2, $n = 64$

(0.5, 0.5)	$(10^{-8}, 0.999999)$	$(10^{-14}, 1)$	(0, 1)
5 FP1	170 FP1	170 FP1	170 FP1
0.6844D-03	0.9889D-03	0.9916D-03	0.9916D-03
2 NM	7 NM	4 NM	4 NM
0.5464D-15	0.5832D-14	0.4937D-05	0.4937D-05
no DN tries	no DN tries	DN (second try)	DN (second try)
		0.1671D-13	0.1763D-13

TABLE 5.5
Convergence history for Example 5.2, $n = 128$

(0.5, 0.5)	$(10^{-8}, 0.999999)$	$(10^{-14}, 1)$	(0, 1)
5 FP1	170 FP1	170 FP1	170 FP1
0.6847D-03	0.9915D-03	0.9942D-03	0.9942D-03
2 NM	7 NM	4 NM	4 NM
0.1117D-14	0.5677D-14	0.4953D-05	0.4953D-05
no DN tries	no DN tries	DN (second try)	DN (second try)
		0.1606D-13	0.1650D-13

Riccati equation. The algorithm is basically a combination of Newton's method and a basic fixed-point iteration, but it has two additional features: the algorithm can detect that an equation actually does not have a positive solution; it can also detect and solve a singular or nearly singular problem efficiently. There are still, however, some unsolved problems about the nonsymmetric algebraic Riccati equation. For example, it is of interest to know what reasonable conditions on the coefficient matrices of the equation will ensure the existence of a positive solution. It is also of interest to determine if quadratic convergence is really possible for Newton's method in the singular case. For symmetric algebraic Riccati equations, subspace methods are frequently used (see [16], for example). It would be worthwhile to consider whether the minimal positive solution of the equation can also be found efficiently by subspace methods.

Acknowledgments. Chun-Hua Guo would like to thank Peter Lancaster for introducing him to the study of algebraic Riccati equations several years ago. He also gratefully acknowledges the support of an NSERC postdoctoral fellowship. Both authors thank the referees for their very helpful comments.

REFERENCES

[1] R. H. BARTELS AND G. W. STEWART, *Solution of the matrix equation* $AX + XB = C$, Comm. ACM, 15 (1972), pp. 820–826.
[2] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.
[3] D. W. DECKER AND C. T. KELLEY, *Newton's method at singular points* I, SIAM J. Numer. Anal., 17 (1980), pp. 66–70.
[4] J. W. DEMMEL, *Three methods for refining estimates of invariant subspaces*, Computing, 38 (1987), pp. 43–57.
[5] M. FIEDLER AND V. PTAK, *On matrices with non-positive off-diagonal elements and positive principal minors*, Czech. Math. J., 12 (1962), pp. 382–400.
[6] G. H. GOLUB, S. NASH, AND C. VAN LOAN, *A Hessenberg-Schur method for the problem* $AX + XB = C$, IEEE Trans. Autom. Control, 24 (1979), pp. 909–913.
[7] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations, 3rd ed.*, Johns Hopkins Univ. Press, Baltimore, MD, 1996.

[8] C.-H. Guo and P. Lancaster, *Analysis and modification of Newton's method for algebraic Riccati equations*, Math. Comp., 67 (1998), pp. 1089–1105.
[9] J. Juang, *Existence of algebraic matrix Riccati equations arising in transport theory*, Linear Algebra Appl., 230 (1995), pp. 89–100.
[10] J. Juang and W.-W. Lin, *Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 228–243.
[11] L. V. Kantorovich and G. P. Akilov, *Functional Analysis in Normed Spaces*, Pergamon, New York, 1964.
[12] C. T. Kelley, *A Shamanskii-like acceleration scheme for nonlinear equations at singular roots*, Math. Comp., 47 (1986), pp. 609–623.
[13] M. A. Krasnoselskii, G. M. Vainikko, P. P. Zabreiko, Ya. B. Rutitskii, and V. Ya. Stetsenko, *Approximate Solution of Operator Equations*, Wolters-Noordhoff Publishing, Groningen, 1972.
[14] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Oxford University Press, 1995.
[15] P. Lancaster and M. Tismenetsky, *The Theory of Matrices, 2nd ed.*, Academic Press, Orlando, FL, 1985.
[16] A. J. Laub, *A Schur method for solving algebraic Riccati equations*, IEEE Trans. Autom. Control, 24 (1979), pp. 913–921.
[17] J. A. Meijerink and H. A. van der Vorst, *An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix*, Math. Comp., 31 (1977), pp. 148–162.
[18] H.-B. Meyer, *The matrix equation $AZ+B-ZCZ-ZD=0$*, SIAM J. Appl. Math., 30 (1976), pp. 136–142.
[19] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
[20] J. S. Vandergraft, *Newton's method for convex operators in partially ordered spaces*, SIAM J. Numer. Anal., 4 (1967), pp. 406–432.
[21] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.