NONSYMMETRIC ALGEBRAIC RICCATI EQUATIONS AND WIENER-HOPF FACTORIZATION FOR M-MATRICES*

CHUN-HUA GUO†

Abstract. We consider the nonsymmetric algebraic Riccati equation for which the four coefficient matrices form an M-matrix. Nonsymmetric algebraic Riccati equations of this type appear in applied probability and transport theory. The minimal nonnegative solution of these equations can be found by Newton's method and basic fixed-point iterations. The study of these equations is also closely related to the so-called Wiener-Hopf factorization for M-matrices. We explain how the minimal nonnegative solution can be found by the Schur method and compare the Schur method with Newton's method and some basic fixed-point iterations. The development in this paper parallels that for symmetric algebraic Riccati equations arising in linear quadratic control.

Key words. nonsymmetric algebraic Riccati equations, M-matrices, Wiener-Hopf factorization, minimal nonnegative solution, Schur method, Newton's method, fixed-point iterations

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

1. Introduction. Symmetric algebraic Riccati equations have been the topic of extensive research. The theory, applications, and numerical solution of these equations are the subject of the monographs [20] and [24]. The algebraic Riccati equation that has received most attention comes from linear quadratic control. It has the form

$$XDX - XA - A^TX - C = 0, \tag{1.1}$$

where $A, C, D \in \mathbb{R}^{n\times n}$; C, D are symmetric positive semidefinite; the pair (A, D) is stabilizable, i.e., there is a $K \in \mathbb{R}^{n\times n}$ such that $A - BK$ is stable (a square matrix is stable if all its eigenvalues are in the open left half-plane); and the pair (C, A) is detectable, i.e., (A^T, C^T) is stabilizable. It is well known that (1.1) has a unique symmetric positive semidefinite solution X and the matrix $A - DX$ is stable (see [20], for example). This solution is the one required in applications and can be found numerically by iterative methods [3, 7, 10, 12, 13, 19] and subspace methods [4, 6, 22, 27, 32, 33].

In this paper, we consider the nonsymmetric algebraic Riccati equation

$$\mathcal{R}(X) = XCX - XD - AX + B = 0, \tag{1.2}$$

where A, B, C, D are real matrices of sizes $m \times m, m \times n, n \times m, n \times n$, respectively. Equation (1.2) in its general form has been studied in [8, 26, 30], for example. All the solutions of (1.2) can be found, in theory, by finding all the Jordan chains of the matrix

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

(see Theorem 7.1.2 of [20]). However, as pointed out in [22], it would be more appropriate to use Schur vectors instead of Jordan vectors.

*This work was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

†Department of Mathematics and Statistics, University of Regina, Regina, SK S4S 0A2, Canada (chguo@math.uregina.ca).

To get a nice theory for the equation (1.2), we need to add some conditions on the matrices A, B, C, and D, much the same as done for the symmetric equation (1.1).

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. 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 $s \geq \rho(B)$, where $\rho(\cdot)$ is the spectral radius. It is called a singular M-matrix if $s = \rho(B)$; it is called a nonsingular M-matrix if $s > \rho(B)$. Note that only nonsingular M-matrices defined here are called M-matrices in [14]. The slight change of definitions is made here for future convenience. The spectrum of a square matrix A will be denoted by $\sigma(A)$. The open left half-plane, the open right half-plane, the closed left half-plane and the closed right half-plane will be denoted by $\mathbb{C}_<$, $\mathbb{C}_>$, $\mathbb{C}_\leq$ and $\mathbb{C}_\geq$, respectively.

In [14], iterative methods are studied for the numerical solution of (1.2) with the condition

$$B > 0, \quad C > 0, \quad I \otimes A + D^T \otimes I \text{ is a nonsingular } M\text{-matrix}, \tag{1.4}$$

where $\otimes$ is the Kronecker product (for basic properties of the Kronecker product, see [21], for example). It is shown there 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.

The condition (1.4) is motivated by a nonsymmetric algebraic Riccati equation arising in transport theory. That equation has the form (1.2) with $m = n$ and the matrices $A, B, C, D \in \mathbb{R}^{n\times n}$ have the following structures:

$$A = \frac{1}{\beta(1+\alpha)}W^{-1} - eq^T,\ B = ee^T,\ C = qq^T,\ D = \frac{1}{\beta(1-\alpha)}W^{-1} - qe^T. \tag{1.5}$$

In the above, $0 \leq \alpha < 1,\ 0 < \beta \leq 1$, and

$$e = (1, 1, \ldots, 1)^T, \quad q = \frac{1}{2}W^{-1}c,$$

where $W = \mathrm{diag}(w_1, w_2, \ldots, w_n),\ c = (c_1, c_2, \ldots, c_n)^T > 0$ with

$$0 < w_n < \cdots < w_2 < w_1 < 1, \quad c^T e = 1.$$

It is shown in [14] that $I \otimes A + D^T \otimes I$ is a nonsingular M-matrix for this equation. For descriptions on how the equation arises in transport theory, see [17] and references cited therein. The existence of positive solutions of this equation has been shown in [16] and [17]. However, only the minimal positive solution is physically meaningful. Numerical methods for finding the minimal solution have also been discussed in [16] and [17].

A more interesting equation of the form (1.2) has recently come to our attention. The equation arises from the Wiener-Hopf factorization of Markov chains [1, 23, 28, 29, 35]. Let Q be the Q-matrix associated with an irreducible continuous-time finite Markov chain $(X_t)_{t\geq 0}$ (a Q-matrix has nonnegative off-diagonal elements and nonpositive row sums; $\exp(tQ)$ is the transition matrix function of the Markov chain). We need to find a quadruple (Π_1, Q_1, Π_2, Q_2) such that

$$\begin{pmatrix} A & B \\ -C & -D \end{pmatrix}\begin{pmatrix} I & \Pi_2 \\ \Pi_1 & I \end{pmatrix} = \begin{pmatrix} I & \Pi_2 \\ \Pi_1 & I \end{pmatrix}\begin{pmatrix} Q_1 & 0 \\ 0 & -Q_2 \end{pmatrix}, \tag{1.6}$$

where

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is a partitioning of Q with A, D being square matrices, and Q_1, Q_2 are Q-matrices. It turns out that the matrices Π_1 and Π_2 of practical interest are the minimal nonnegative solutions of the nonsymmetric algebraic Riccati equations $ZBZ+ZA+DZ+C = 0$ and $ZCZ+ZD+AZ+B = 0$, respectively (see [35]). The relation in (1.6) has been called a Wiener-Hopf factorization of the leftmost matrix in (1.6). The factorization (1.6) makes perfect sense for *any* Q-matrix, with or without probabilistic significance. Barlow, Rogers, and Williams [1] noted that they did not know how to establish the factorization without appealing to probability theory (and ultimately to martingale theory). Rogers [28] studied the factorization in more detail, using again probabilistic results and interpretations.

Note that $-Q$ is an M-matrix for any Q-matrix Q. Thus, the Riccati equations arising from the study of Markov chains are essentially special cases of the Riccati equation (1.2) with condition (1.4). However, the strict positiveness of B and C could be restrictive. We will thus relax the condition (1.4) to conditions

$$B, C \geq 0, \quad I \otimes A + D^T \otimes I \text{ is a nonsingular } M\text{-matrix}, \tag{1.7}$$

and

$$B, C \neq 0, \quad (I \otimes A + D^T \otimes I)^{-1} \mathrm{vec} B > 0, \tag{1.8}$$

where the vec operator stacks the columns of a matrix into one long vector. For some of our discussions, condition (1.7) alone will be sufficient.

The theory of M-matrices will play an important role in our discussions. The following result is well known (see [5] and [9], for example).

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

(1) *A is a nonsingular M-matrix.*
(2) *$A^{-1} \geq 0$.*
(3) *$Av > 0$ for some vector $v > 0$.*
(4) *$\sigma(A) \subset \mathbb{C}_>$.*

The next result follows from the equivalence of (1) and (3) in Theorem 1.1 and can be found in [25], for example.

THEOREM 1.2. *Let $A \in \mathbb{R}^{n\times n}$ be a nonsingular 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 a nonsingular M-matrix.

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 [21], for example), it follows from the equivalence of (1) and (4) in Theorem 1.1 that $I \otimes A + D^T \otimes I$ is a nonsingular M-matrix when A, D are both nonsingular M-matrices.

2. Iterative methods. Newton's method and a class of basic fixed-point iterations are studied in [14] for the numerical solution of (1.2) under condition (1.4). In this section, we represent the main results in [14] under weaker conditions. These

results will be needed in later discussions. For Newton's method, we need (1.7) and (1.8). For basic fixed-point iterations, condition (1.8) is not necessary.

We first consider the application of Newton's method to (1.2). 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

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

The Newton method for the solution of (1.2) is

$$X_{i+1} = X_i - (\mathcal{R}'_{X_i})^{-1}\mathcal{R}(X_i), \quad i = 0, 1, \ldots, \tag{2.2}$$

given that the maps $\mathcal{R}'_{X_i}$ are all invertible. In view of (2.1), the iteration (2.2) is equivalent to

$$(A - X_iC)X_{i+1} + X_{i+1}(D - CX_i) = B - X_iCX_i, \quad i = 0, 1, \ldots. \tag{2.3}$$

THEOREM 2.1. *Consider the equation* (1.2) *with conditions* (1.7) *and* (1.8). *If there is a positive matrix* X *such that* $\mathcal{R}(X) \le 0$, *then* (1.2) *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* (1.2). *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*

$$I \otimes (A - SC) + (D - CS)^T \otimes I \tag{2.4}$$

is an M*-matrix.*

The proof of the above theorem is exactly the same as that of [14, Theorem 2.1]. We do not have an analogous result for nonnegative solutions if the condition (1.8) is dropped. Note that, under the conditions of Theorem 2.1, any nonnegative matrix satisfying $\mathcal{R}(X) \le 0$ must be positive (see the remark following Theorem 2.3).

Concerning the convergence rate of Newton's method, we have the following result. The proof is again the same as in [14].

THEOREM 2.2. *Let the sequence* $\{X_i\}$ *be as in Theorem* 2.1. *If the matrix* (2.4) *is a nonsingular* M*-matrix, then* $\{X_i\}$ *converges to* S *quadratically. If* (2.4) *is an irreducible singular* M*-matrix, then* $\{X_i\}$ *converges to* S *either quadratically or linearly with rate* $1/2$.

We believe that quadratic convergence is impossible in the singular case, but we have no proof for this.

We now consider a class of fixed-point iterations for (1.2) under condition (1.7) only. If we write $A = A_1 - A_2$, $D = D_1 - D_2$, then (1.2) 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 a nonsingular M-matrix by Theorem 1.2. We then have a class of fixed-point iterations

$$X_{k+1} = \mathcal{L}^{-1}(X_kCX_k + X_kD_2 + A_2X_k + B), \tag{2.5}$$

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 a nonsingular M-matrix, the operator $\mathcal{L}$ is invertible and $\mathcal{L}^{-1}(X) \ge 0$ for $X \ge 0$.

THEOREM 2.3. *Consider the equation* (1.2) *with condition* (1.7). *For the fixed-point iterations* (2.5) *and* $X_0 = 0$, *we have* $X_k \le X_{k+1}$ *for any* $k \ge 0$. *If* $\mathcal{R}(X) \le 0$ *for some nonnegative matrix* X, *then we also have* $X_k \le X$ *for any* $k \ge 0$. *Moreover,* $\{X_k\}$ *converges to the minimal nonnegative solution of* (1.2).

Proof. It is easy to prove by induction that $X_k \le X_{k+1}$ for any $k \ge 0$. When $\mathcal{R}(X) \le 0$ for some nonnegative matrix X, we can prove by induction that $X_k \le X$ for any $k \ge 0$. The limit X^* of $\{X_k\}$ is then a solution of $\mathcal{R}(X) = 0$ and must be the minimal nonnegative solution, since $X^* \le X$ for any nonnegative matrix such that $\mathcal{R}(X) \le 0$. □

Remark 2.1. If condition (1.8) is also satisfied, then the matrix X_1 produced by (2.5) with $A_1 = A$ and $D_1 = D$ is positive. This is because $\mathrm{vec}X_1 = (I \otimes A + D^T \otimes I)^{-1}\mathrm{vec}B$. Thus, for any nonnegative matrix X such that $\mathcal{R}(X) \le 0$, we have $X \ge X_1 > 0$.

The next comparison result follows easily from Theorem 2.3.

THEOREM 2.4. *Consider the equation* (1.2) *with condition* (1.7) *and let* S *be the minimal nonnegative solution of* (1.2). *If any element of* B *or* C *decreases but remains nonnegative, or if any diagonal element of* $I \otimes A + D^T \otimes I$ *increases, or if any off-diagonal elment of* $I \otimes A + D^T \otimes I$ *increases but remains nonpositive, then the equation so obtained also has a minimal nonnegative solution* $\tilde{S}$. *Moreover,* $\tilde{S} \le S$.

Proof. Let the new equation be

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

It is clear that $\tilde{R}(S) \le 0$. Since $I \otimes \tilde{A} + \tilde{D}^T \otimes I$ is still a nonsingular M-matrix by Theorem 1.2, the conclusions follow from Theorem 2.3. □

The following result is concerned with the convergence rates of the fixed-point iterations. It is a slight modification of Theorem 3.2 in [14]. The proof given there is valid without change.

THEOREM 2.5. *Consider the equation* (1.2) *with condition* (1.7) *and let* S *be the minimal nonnegative solution of* (1.2). *For the fixed-point iterations* (2.5) *with* $X_0 = 0$, *we have*

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

Equality holds if S *is positive.*

COROLLARY 2.6. *For the equation* (1.2) *with condition* (1.7), *if the minimal nonnegative solution* S *of* (1.2) *is positive, then the matrix* (2.4) *is an* M*-matrix.*

Proof. Let A_1 and D_1 be the diagonal part of A and D, respectively. By Theorem 2.5, we have $\rho\big((I\otimes A_1+D_1^T\otimes I)^{-1}(I\otimes(A_2+SC)+(D_2+CS)^T\otimes I)\big) \le 1$. Therefore, for any $\epsilon > 0$, $\rho\big((I\otimes(A_1+\epsilon A_1)+(D_1+\epsilon D_1)^T\otimes I)^{-1}(I\otimes(A_2+SC)+(D_2+CS)^T\otimes I)\big) < 1$. Thus, $\epsilon(I \otimes A_1 + D_1^T \otimes I) + I \otimes (A - SC) + (D - CS)^T \otimes I$ is a nonsingular M-matrix (see [5], for example). It follows that (2.4) is an M-matrix. □

As in [14], we have the following result about the spectral radius in Theorem 2.5.

THEOREM 2.7. *Consider the equation* (1.2) *with condition* (1.7) *and let* S *be the minimal nonnegative solution of* (1.2). *If* (2.4) *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 (2.4) *is a nonsingular* 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*

$$\rho\big((I \otimes \tilde{A}_1 + \tilde{D}_1^T \otimes I)^{-1}(I \otimes (\tilde{A}_2 + SC) + (\tilde{D}_2 + CS)^T \otimes I)\big)$$

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

Therefore, the convergence of these iterations is linear if (2.4) is a nonsingular M-matrix. When (2.4) is a singular M-matrix, the convergence is sublinear. Within this class of iterative methods, three iterations deserve special attention. 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 $D_1 = D$. It will be called FP3.

3. A sufficient condition for the existence of nonnegative solutions. In the last section, the existence of a nonnegative solution of (1.2) is guaranteed under the assumption that there is a nonnegative matrix X such that $\mathcal{R}(X) \le 0$. The usefulness of this kind of assumption was evident in the ease we had in proving Theorem 2.4. However, if the equation (1.2) does not have a nonnegative solution, the search for a nonnegative matrix X such that $\mathcal{R}(X) \le 0$ will necessarily be fruitless. In this section, we will give a sufficient condition of a different kind for the existence of nonnegative solutions of the equation (1.2). This condition is suggested by the Wiener-Hopf factorization of Markov chains.

THEOREM 3.1. *If the matrix*

$$K = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix} \tag{3.1}$$

is a nonsingular M-matrix, then the equation (1.2) *has a nonnegative solution S such that $D-CS$ is a nonsingular M-matrix. If* (3.1) *is an irreducible singular M-matrix, then* (1.2) *has a nonnegative solution S such that $D-CS$ is an M-matrix.*

Proof. If (3.1) is a nonsingular M-matrix, then $T = \operatorname{diag}(D, A)$ is also a nonsingular M-matrix by Theorem 1.2. If (3.1) is an irreducible singular M-matrix, then T is a nonsingular M-matrix by the Perron-Frobenius theory (see [5] or [34]). Thus, in either case, A and D are nonsingular M-matrices. Therefore, condition (1.7) is satisfied. We take $X_0 = 0$ and use FP1:

$$A_1X_{i+1} + X_{i+1}D_1 = X_iCX_i + X_iD_2 + A_2X_i + B, \quad i = 0, 1, \ldots. \tag{3.2}$$

By Theorem 2.3, $X_i \le X_{i+1}$ for any $i \ge 0$.

If (3.1) is a nonsingular M-matrix, we can find $v_1, v_2 > 0$ such that

$$D_1v_1 - D_2v_1 - Cv_2 = u_1 > 0, \quad A_1v_2 - A_2v_2 - Bv_1 = u_2 > 0. \tag{3.3}$$

We will show that $X_kv_1 \le v_2 - A_1^{-1}u_2$ for all $k \ge 0$. The inequality is true for $k = 0$ since $v_2 - A_1^{-1}u_2 = A_1^{-1}(A_2v_2 + Bv_1) \ge 0$ by the second equation in (3.3). Assume that $X_iv_1 \le v_2 - A_1^{-1}u_2\,(i \ge 0)$. Then, by (3.2) and (3.3),

$$\begin{aligned} A_1X_{i+1}v_1 + X_{i+1}D_1v_1 &= X_iCX_iv_1 + X_iD_2v_1 + A_2X_iv_1 + Bv_1 \\ &\le X_iCv_2 + X_iD_2v_1 + A_2v_2 + Bv_1 \\ &\le X_iD_1v_1 + A_1v_2 - u_2. \end{aligned}$$

Since $X_{i+1}D_1v_1 \ge X_iD_1v_1$, we have $A_1X_{i+1}v_1 \le A_1v_2 - u_2$. Therefore, $X_{i+1}v_1 \le v_2 - A_1^{-1}u_2$. Thus, we have proved by induction that $X_kv_1 \le v_2 - A_1^{-1}u_2$ for all $k \ge 0$. Now, the sequence $\{X_i\}$ is monotonically increasing and bounded above, and hence

has a limit. Let $S = \lim_{i\to\infty} X_i$. It is clear that S is a nonnegative solution of (1.2) and $Sv_1 \le v_2 - A_1^{-1}u_2 < v_2$. Thus, $(D - CS)v_1 \ge Dv_1 - Cv_2 = u_1 > 0$. Therefore, $D - CS$ is a nonsingular M-matrix by Theorem 1.1. If (3.1) is an irreducible singular M-matrix, there are $v_1, v_2 > 0$ (by the Perron-Frobenius theory) such that

$$D_1v_1 - D_2v_1 - Cv_2 = 0, \quad A_1v_2 - A_2v_2 - Bv_1 = 0.$$

We can prove as before that the sequence $\{X_i\}$ produced by FP1 is such that $X_iv_1 \le v_2$ for all $i \ge 0$. The limit S of the sequence is a nonnegative solution of (1.2) with $Sv_1 \le v_2$. Therefore, $(D - CS)v_1 \ge Dv_1 - Cv_2 = 0$. Thus, $D - CS + \epsilon I$ is a nonsingular M-matrix for any $\epsilon > 0$. So, $D - CS$ is an M-matrix. □

Remark 3.1. We know from Theorem 2.3 that the matrix S in the proof is the minimal nonnegative solution of (1.2). Note also that we have obtained in the proof some additional information about the minimal solution. It will be seen later (from Theorem 4.2) that the minimal solution S is the only solution X that makes $D - CX$ an M-matrix when the matrix (3.1) is a nonsingular M-matrix. If the matrix (3.1) is an irreducible singular M-matrix, then (1.2) may have more than one nonnegative solutions X such that $D - CX$ is an M-matrix. For example, for $A = B = 1$ and $C = D = 2$, the scalar equation (1.2) has two positive solutions $X = 1$ and $X = 1/2$. The first makes $D - CX$ a singular M-matrix. The second makes $D - CX$ a nonsingular M-matrix.

We have seen in the proof of Theorem 3.1 that condition (1.7) is satisfied when the matrix (3.1) is a nonsingular M-matrix or an irreducible singular M-matrix. It is clear that (1.8) is not necessarily true when (3.1) is a nonsingular M-matrix. If (3.1) is an irreducible singular M-matrix, we have $B, C \ne 0$. However,

$$(I \otimes A + D^T \otimes I)^{-1}\mathrm{vec}B > 0 \tag{3.4}$$

is not necessarily true.

Assume that (3.1) is an irreducible M-matrix. If (3.4) is true, then the minimal nonnegative solution S of (1.2) must be positive. However, a more practical method to verify the positivity of S is to apply FP1 with $X_0 = 0$ to equation (1.2).

Example 3.1. For equation (1.2) with

$$A = C = D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

the matrix (3.1) is an irreducible singular M-matrix, but (3.4) is not true. However, the minimal nonnegative solution S is still positive. In fact, if we apply FP1 with $X_0 = 0$ to the equation (1.2), we get $X_2 > 0$. Thus, $S \ge X_2 > 0$.

We also have the following sufficient condition for the positivity of S.

PROPOSITION 3.2. *If* (3.1) *is an* M*-matrix such that* $B, C \ne 0$ *and* A, D *are irreducible, then* (3.1) *is irreducible,* (3.4) *is true, and* $S > 0$.

Proof. The matrix (3.1) is irreducible by a graph argument (see Theorem 2.2.7 of [5]). As shown in the proof of Theorem 3.1, the matrices A and D are now irreducible nonsingular M-matrices. Thus, $I \otimes A + D^T \otimes I$ is an irreducible nonsingular M-matrix (the irreducibility is shown by a graph argument). Therefore, by Theorem 6.2.7 of [5], $(I \otimes A + D^T \otimes I)^{-1} > 0$. Thus, (3.4) is true and $S > 0$. □

More can be said about the matrix $D - CS$ when the minimal nonnegative solution S of (1.2) is positive.

THEOREM 3.3. *If* (3.1) *is an irreducible M-matrix and the minimal nonnegative solution S of* (1.2) *is positive, then* $D - CS$ *is an irreducible M-matrix* (*we use the convention that a* 1×1 *zero matrix is irreducible*).

Proof. We only need to prove that $D - CS$ is irreducible for $n \geq 2$. Write $D = (d_{ij})$. Let $V_1 = \{i \mid 1 \leq i \leq n,$ the ith row of C is zero$\}$ and $V_2 = \{i \mid 1 \leq i \leq n,$ the ith row of C is not zero$\}$. Since (3.1) is irreducible, its graph is strongly connected (see Theorem 2.2.7 of [5]). Therefore, for any $i \in V_1$ we can find $i_1, \ldots, i_{k-1} \in V_1$ (void if $k = 1$) and $i_k \in V_2$ such that $d_{ii_1}, d_{i_1i_2}, \ldots, d_{i_{k-1}i_k}$ are nonzero. Now, take any $i : 1 \leq i \leq n$. If $i \in V_2$, then the off-diagonal elements of $D - CS$ in the ith row are negative since S is positive; the diagonal element of $D - CS$ in the ith row must be positive since it is shown in the proof of Theorem 3.1 that $(D - CS)v_1 \geq 0$ for some $v_1 > 0$. If $i \in V_1$, then the ith row of $D - CS$ is the same as the ith row of D. It follows readily that the graph of $D - CS$ is strongly connected. So, $D - CS$ is irreducible. □

The equation (1.2) with no prescribed sign structure for the matrices $A, B, C,$ and D has been considered in [8] and [30]. It is shown that the solution of (1.2) with minimal Frobenius norm can be found by FP3 and Newton's method starting with $X_0 = 0$, if $\kappa < 1/4$ for FP3 and $\kappa < 1/12$ for Newton's method, where $\kappa = \|B\|_F\|C\|_F/s^2$ and s is the smallest singular value of $I \otimes A + D^T \otimes I$. If the matrix (3.1) is a singular M-matrix with no zero elements, for example, then the minimal positive solution can be found by FP3 and Newton's method with $X_0 = 0$. It is interesting to see how often the condition $\kappa < 1/4$ is satisfied when (3.1) is a singular M-matrix with no zero elements. We use MATLAB to obtain a 4×4 positive matrix R using rand(4,4), so $W = \text{diag}(Re) - R$ is a singular M-matrix with no zero elements. We let the matrix W be in the form (3.1), so the 2×2 matrices A, B, C, D are determined. We find that $\kappa < 1/4$ is satisfied 198 times for 10000 random matrices R ($\kappa < 1/5$ is satisfied 35 times). When we use rand(6,6) to get 3×3 matrices A, B, C, D in the same way, we find that $\kappa < 1/4$ is satisfied 2 times for 10000 random matrices.

It is interesting to note that the equation (1.2) from transport theory also satisfies the conditions in Theorem 3.1.

PROPOSITION 3.4. *Let the matrices* A, B, C, D *be defined by* (1.5). *Then the matrix* K *given by* (3.1) *is irreducible. The matrix* K *is a nonsingular M-matrix for* $0 < \beta < 1$ *and is a singular M-matrix for* $\beta = 1$.

Proof. By definition,

$$K = \begin{pmatrix} \frac{1}{\beta(1-\alpha)}W^{-1} - qe^T & -qq^T \\ -ee^T & \frac{1}{\beta(1+\alpha)}W^{-1} - eq^T \end{pmatrix}.$$

It is clear that K is irreducible. Since K is a singular (nonsingular) M-matrix if and only if

$$\begin{pmatrix} (1-\alpha)W & 0 \\ 0 & (1+\alpha)W \end{pmatrix} K = \begin{pmatrix} \frac{1}{\beta}I - (1-\alpha)Wqe^T & -(1-\alpha)Wqq^T \\ -(1+\alpha)Wee^T & \frac{1}{\beta}I - (1+\alpha)Weq^T \end{pmatrix}$$

is a singular (nonsingular) M-matrix, we only need to find a positive vector v such that $Qv = v$ for the positive matrix

$$Q = \begin{pmatrix} (1-\alpha)Wqe^T & (1-\alpha)Wqq^T \\ (1+\alpha)Wee^T & (1+\alpha)Weq^T \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(1-\alpha)ce^T & \frac{1}{4}(1-\alpha)cc^TW^{-1} \\ (1+\alpha)Wee^T & \frac{1}{2}(1+\alpha)Wec^TW^{-1} \end{pmatrix},$$

where we have used $q = \frac{1}{2}W^{-1}c$. Now, since $e^Tc = c^Te = 1$, direct computation shows that $Qv = v$ for

$$v = \begin{pmatrix} (1-\alpha)c \\ 2(1+\alpha)We \end{pmatrix} > 0.$$

This completes the proof. □

Therefore, with Remark 2.1 in mind, the existence of positive solutions of (1.2) with A, B, C, D given by (1.5) is established as a special case of Theorem 3.1. The existence in this special case was proved in [16] using the degree theory and was proved in [17] using the secular equation and other tools.

4. Wiener-Hopf factorization for M-matrices. The Wiener-Hopf factorization for Q-matrices associated with finite Markov chains has been studied in [1, 23, 28, 29, 35]. The factorization was obtained by using probabilistic results and interpretations. In this section, we will establish Wiener-Hopf factorization for M-matrices. Our results include Wiener-Hopf factorization for Q-matrices as a special case. The proof will be purely algebraic.

THEOREM 4.1. *If the matrix* (3.1) *is a nonsingular M-matrix or an irreducible singular M-matrix, then there exist nonnegative matrices S_1 and S_2 such that*

$$\begin{pmatrix} D & -C \\ B & -A \end{pmatrix} \begin{pmatrix} I & S_2 \\ S_1 & I \end{pmatrix} = \begin{pmatrix} I & S_2 \\ S_1 & I \end{pmatrix} \begin{pmatrix} G_1 & 0 \\ 0 & -G_2 \end{pmatrix}, \tag{4.1}$$

where G_1 and G_2 are M-matrices.

Proof. By Theorem 3.1, the equation (1.2) has a nonnegative solution S_1 such that $D - CS_1$ is an M-matrix. By taking $G_1 = D - CS_1$, we get

$$\begin{pmatrix} D & -C \\ B & -A \end{pmatrix} \begin{pmatrix} I \\ S_1 \end{pmatrix} = \begin{pmatrix} I \\ S_1 \end{pmatrix} G_1. \tag{4.2}$$

Since

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

is also a nonsingular M-matrix or an irreducible singular M-matrix, Theorem 3.1 implies that the equation

$$XBX - XA - DX + C = 0 \tag{4.3}$$

has a nonnegative solution S_2 such that $A - BS_2$ is an M-matrix. Letting $G_2 = A - BS_2$, we have

$$\begin{pmatrix} D & -C \\ B & -A \end{pmatrix} \begin{pmatrix} S_2 \\ I \end{pmatrix} = \begin{pmatrix} S_2 \\ I \end{pmatrix} (-G_2). \tag{4.4}$$

The factorization (4.1) is obtained by combining (4.2) and (4.4). □

We can make stronger statements when the matrix (3.1) is a nonsingular M-matrix.

THEOREM 4.2. *If the matrix* (3.1) *is a nonsingular M-matrix, then the only matrices S_1 and S_2 satisfying* (4.1) *with G_1 and G_2 being M-matrices are the minimal*

nonnegative solution of (1.2) *and the minimal nonnegative solution of* (4.3)*, respectively. In this case,* G_1 *and* G_2 *are nonsingular* M*-matrices and the matrix*

$$\begin{pmatrix} I & S_2 \\ S_1 & I \end{pmatrix} \tag{4.5}$$

is nonsingular.

Proof. Let S_1 and S_2 be the minimal nonnegative solutions of (1.2) and (4.3), respectively. Let $G_1 = D - CS_1$ and $G_2 = A - BS_2$. Then (4.1) holds and G_1, G_2 are nonsingular M-matrices (see Theorem 3.1). Let $v_1, v_2 > 0$ be as in the proof of Theorem 3.1. Then, $S_1 v_1 < v_2$ and $S_2 v_2 < v_1$. Since

$$\begin{pmatrix} I & -S_2 \\ -S_1 & I \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} > 0,$$

the matrix (4.5) is a generalized strictly diagonally dominant matrix and hence nonsingular. Thus, (4.1) gives a similarity transformation and, as a result, the matrix (1.3) has n eigenvalues in $\mathbb{C}_>$ and m eigenvalues in $\mathbb{C}_<$. Now, if $\tilde{S}_1 \in \mathbb{R}^{m\times n}$ and $\tilde{S}_2 \in \mathbb{R}^{n\times m}$ satisfy (4.1) with $\tilde{G}_1$ and $\tilde{G}_2$ being M-matrices, then

$$\begin{pmatrix} D & -C \\ B & -A \end{pmatrix} \begin{pmatrix} I \\ \tilde{S}_1 \end{pmatrix} = \begin{pmatrix} I \\ \tilde{S}_1 \end{pmatrix} \tilde{G}_1.$$

Therefore, the eigenvalues of $\tilde{G}_1$ are precisely the n eigenvalues of (1.3) in $\mathbb{C}_>$. Since the column spaces of $(I \ \tilde{S}_1^T)^T$ and $(I \ S_1^T)^T$ are the same invariant subspace associated with these eigenvalues, we conclude that $\tilde{S}_1 = S_1$. Similarly, $\tilde{S}_2 = S_2$. □

From the above theorem and its proof, it is already clear that we can find the minimal solution using an appropriate invariant subspace (details will be provided in the next section).

The rest of this section is devoted to the case where (3.1) is an irreducible singular M-matrix. The minimal nonnegative solutions of (1.2) and (4.3) will be denoted by S_1 and S_2, respectively.

Let v_1, v_2, u_1, u_2 be positive vectors such that

$$\begin{pmatrix} D & -C \\ -B & A \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0, \quad (u_1^T \ u_2^T) \begin{pmatrix} D & -C \\ -B & A \end{pmatrix} = 0. \tag{4.6}$$

Multiplying (4.1) by $\begin{pmatrix} u_1^T & -u_2^T \end{pmatrix}$ from the left gives

$$(u_1^T - u_2^T S_1)G_1 = 0, \quad (u_1^T S_2 - u_2^T)G_2 = 0.$$

If G_1 is nonsingular, then $u_1^T - u_2^T S_1 = 0$. Moreover, we see from the proof of Theorem 3.1 that $S_1 v_1 \le v_2$ and $S_1 v_1 \ne v_2$ (if $S_1 v_1 = v_2$, we would have $G_1 v_1 = (D - CS_1)v_1 = Dv_1 - Cv_2 = 0$, which is contradictory to the nonsingularity of G_1). So $u_1^T v_1 < u_2^T v_2$. Similarly, $u_1^T v_1 > u_2^T v_2$ when G_2 is nonsingular. Therefore, the following result is true.

LEMMA 4.3. *When* (3.1) *is an irreducible singular* M*-matrix,* G_1 *is singular if* $u_1^T v_1 > u_2^T v_2$*;* G_2 *is singular if* $u_1^T v_1 < u_2^T v_2$*; both* G_1 *and* G_2 *are singular if* $u_1^T v_1 = u_2^T v_2$.

Further discussions will be dependent on the positivity of S_1 and S_2.

LEMMA 4.4. *Assume that* (3.1) *is an irreducible singular* M*-matrix and* $S_1, S_2 > 0$*. Then the matrix* (1.3) *has* $n-1$ *eigenvalues in* $\mathbb{C}_>$*,* $m-1$ *eigenvalues in* $\mathbb{C}_<$*,*

one zero eigenvalue, and one more eigenvalue which is either zero or has nonzero real part.

Proof. By Theorem 3.3, G_1 and G_2 are irreducible M-matrices. Therefore, $G_1(G_2)$ has $n(m)$ eigenvalues in $\mathbb{C}_>$ when it is nonsingular; $G_1(G_2)$ has a zero eigenvalue and $n-1(m-1)$ eigenvalues in $\mathbb{C}_>$ when it is singular. By (4.1), the eigenvalues of G_1(resp., $-G_2$) are precisely the eigenvalues of the matrix (1.3) restricted to the column space of $(I \;\; S_1^T)^T$ (resp., $(S_2^T \;\; I)^T$). The result follows immediately. □

LEMMA 4.5. *Under the assumptions of Lemma* 4.4*, zero is a double eigenvalue of* (1.3) *if and only if* $u_1^T v_1 = u_2^T v_2$.

Proof. If zero is a double eigenvalue of (1.3), then the Jordan canonical form for (1.3) is

$$P^{-1} \begin{pmatrix} D & -C \\ B & -A \end{pmatrix} P = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix},$$

where $J_1 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and J_2 consists of Jordan blocks associated with nonzero eigenvalues (note that the null space of (1.3) is one-dimensional since (3.1) is an irreducible M-matrix). By (4.6), we get

$$(u_1^T \;\; -u_2^T)P = k_1 e_2^T, \quad P^{-1} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = k_2 e_1, \tag{4.7}$$

where e_1, e_2 are the first two standard unit vectors and k_1, k_2 are nonzero constants. Multiplying the two equations in (4.7) gives $u_1^T v_1 = u_2^T v_2$. If zero is a simple eigenvalue of (1.3), then we have $J_1 = (0)$ instead and we have (4.7) with e_2 replaced by e_1. Thus, $u_1^T v_1 \neq u_2^T v_2$. □

We will also need the following general result, which can be found in [26], for example.

LEMMA 4.6. *If* X *is any solution of* (1.2)*, then*

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

Thus, the eigenvalues of $D - CX$ *are eigenvalues of* (1.3) *and the eigenvalues of* $A - XC$ *are the negative of the remaining eigenvalues of* (1.3).

The next result determines the signs of the real parts for all eigenvalues of the matrix (1.3) and it also paves the way for finding S_1 and S_2 using subspace methods.

THEOREM 4.7. *Assume that the matrix* (3.1) *is an irreducible singular* M*-matrix and* $S_1, S_2 > 0$*. Let the vectors* u_1, u_2, v_1, v_2 *be as in* (4.6)*. Then*

1. *If* $u_1^T v_1 = u_2^T v_2$*, then* (1.3) *has* $n-1$ *eigenvalues in* $\mathbb{C}_>$*,* $m-1$ *eigenvalues in* $\mathbb{C}_<$*, and two zero eigenvalues. Moreover,* G_1 *and* G_2 *are singular* M*-matrices.*
2. *If* $u_1^T v_1 > u_2^T v_2$*, then* (1.3) *has* $n-1$ *eigenvalues in* $\mathbb{C}_>$*,* m *eigenvalues in* $\mathbb{C}_<$*, and one zero eigenvalue. Moreover,* G_1 *is a singular* M*-matrix and* G_2 *is a nonsingular* M*-matrix.*
3. *If* $u_1^T v_1 < u_2^T v_2$*, then* (1.3) *has* n *eigenvalues in* $\mathbb{C}_>$*,* $m-1$ *eigenvalues in* $\mathbb{C}_<$*, and one zero eigenvalue. Moreover,* G_1 *is a nonsingular* M*-matrix and* G_2 *is a singular* M*-matrix.*

Proof. Assertion (1) follows from Lemmas 4.5, 4.4, and 4.3. We will prove assertion (2) only since the proof of assertion (3) is very similar. When $u_1^T v_1 > u_2^T v_2$, $G_1 = D - CS_1$ is singular by Lemma 4.3. The last-mentioned eigenvalue in Lemma 4.4 cannot be zero by Lemma 4.5 and we need to show that it is in $\mathbb{C}_<$. If this

eigenvalue were in $\mathbb{C}_>$, the matrix $A - S_1C$ would have $m-1$ eigenvalues in $\mathbb{C}_>$ and one eigenvalue in $\mathbb{C}_<$, in view of Lemma 4.6. Since the eigenvalues of $G = I \otimes (A - S_1C) + (D - CS_1)^T \otimes I$ are the sums of eigenvalues of $A - S_1C$ and eigenvalues of $D - CS_1$, the matrix G would then have an eigenvalue in $\mathbb{C}_<$. This is a contradiction since G is an M-matrix by Corollary 2.6. A similar argument then shows that the eigenvalues of $G_2 = A - BS_2$ must be the negative of the m eigenvalues of (1.3) in $\mathbb{C}_<$. Therefore, G_2 is a nonsingular M-matrix by Theorem 1.1. □

Remark 4.1. Case (1) of Theorem 4.7 poses a great challenge to basic fixed-point iterations. Since the matrix (2.4) is a singular M-matrix in this case, the convergence of the fixed-point iterations for (1.2) is sublinear (see Theorems 2.5 and 2.7). The convergence of Newton's method for (1.2) is typically linear with rate 1/2 in this case if condition (3.4) is also satisfied, but the performance of Newton's method can be improved by using a double Newton step (see discussions in [14]).

When the matrix (3.1) is an irreducible singular M-matrix, we know from the proof of Theorem 3.1 that $S_1v_1 \le v_2$ and $S_2v_2 \le v_1$. With the additional assumption that $S_1, S_2 > 0$, we can say something more about S_1 and S_2.

THEOREM 4.8. *Under the conditions of Theorem* 4.7*, we have*

(1) *If $u_1^Tv_1 = u_2^Tv_2$, then $CS_1v_1 = Cv_2$ and $BS_2v_2 = Bv_1$. Consequently, $S_1v_1 = v_2$ and $S_2v_2 = v_1$ if C and B have no zero columns.*
(2) *If $u_1^Tv_1 > u_2^Tv_2$, then $S_2v_2 \ne v_1$ and $CS_1v_1 = Cv_2$. Consequently, $S_1v_1 = v_2$ if C has no zero columns.*
(3) *If $u_1^Tv_1 < u_2^Tv_2$, then $S_1v_1 \ne v_2$ and $BS_2v_2 = Bv_1$. Consequently, $S_2v_2 = v_1$ if B has no zero columns.*

Moreover, the matrix (4.5) *is singular if and only if $S_1v_1 = v_2$ and $S_2v_2 = v_1$.*

Proof. We will prove (3). The proof of (1) and (2) is similar. If $u_1^Tv_1 < u_2^Tv_2$, then G_1 is a nonsingular M-matrix and G_2 is a singular M-matrix by Theorem 4.7. That $S_1v_1 \ne v_2$ has been proved in the discussions leading to Lemma 4.3. Note that G_2 is irreducible and $G_2v_2 = (A - BS_2)v_2 \ge Av_2 - Bv_1 = 0$. If $BS_2v_2 \ne Bv_1$, then G_2v_2 would be nonnegative and nonzero. Therefore, G_2 would be a nonsingular M-matrix by Theorem 6.2.7 of [5]. The contradiction shows that $BS_2v_2 = Bv_1$. Thus, $B(v_1 - S_2v_2) = 0$. It follows that $v_1 - S_2v_2 = 0$ if B has no zero columns. The proof of (3) is completed.

If $S_1v_1 = v_2$ and $S_2v_2 = v_1$, then

$$\begin{pmatrix} I & S_2 \\ S_1 & I \end{pmatrix} \begin{pmatrix} v_1 \\ -v_2 \end{pmatrix} = 0.$$

Thus, the matrix (4.5) is singular. If $S_1v_1 = v_2$ and $S_2v_2 = v_1$ are not both true, then

$$\begin{pmatrix} I & -S_2 \\ -S_1 & I \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{4.8}$$

is nonnegative and nonzero. Since S_1 and S_2 are positive, the matrix on the left side of (4.8) is irreducible. Therefore, it is a nonsingular M-matrix by Theorem 6.2.7 of [5] and hence the matrix (4.5) is nonsingular. □

For the equation (1.2) from transport theory, the matrix (3.1) is an irreducible singular M-matrix if $\beta = 1$ (see Proposition 3.4). The next result shows that, for this special equation, only cases (1) and (3) are possible in Theorems 4.7 and 4.8.

PROPOSITION 4.9. *For the equation* (1.2) *with A, B, C, D given by* (1.5) *with $\beta = 1$, we have $u_1^Tv_1 = u_2^Tv_2$ for $\alpha = 0$ and $u_1^Tv_1 < u_2^Tv_2$ for $0 < \alpha < 1$.*

Proof. By the proof of Proposition 3.4, we can take $v_1 = (1-\alpha)c$ and $v_2 = 2(1+\alpha)We$. Similarly, we can take $u_1 = 2(1-\alpha)We$ and $u_2 = (1+\alpha)c$. The conclusions follow immediately. □

5. The Schur method. In this section, we will explain how to use the Schur method to find the minimal nonnegative solution of the equation (1.2).

THEOREM 5.1. *Assume that* (3.1) *is a nonsingular M-matrix or an irreducible singular M-matrix such that the minimal nonnegative solutions of* (1.2) *and* (4.3) *are positive. Let H be the matrix given by* (1.3). *Let U be an orthogonal matrix such that*

$$U^T H U = F$$

is a real Schur form of H, where the 1×1 *or* 2×2 *diagonal blocks of F are arranged in the order for which the real parts of the corresponding eigenvalues are nonincreasing. If U is partitioned as*

$$\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

where $U_{11} \in \mathbb{R}^{n\times n}$, *then* U_{11} *is nonsingular and* $U_{21}U_{11}^{-1}$ *is the minimal nonnegative solution of* (1.2).

Proof. From the discussions in Sections 3 and 4, we know that the minimal nonnegative solution S exists and the n-dimensional column space of $(I \;\; S^T)^T$ is either the n-dimensional invariant subspace of H corresponding to the eigenvalues in $\mathbb{C}_>$ or, when the largest invariant subspace $\mathcal{V}$ of H corresponding to the eigenvalues in $\mathbb{C}_>$ is only $(n-1)$-dimensional, the direct sum of $\mathcal{V}$ and the one-dimensional eigenspace of H corresponding to the zero eigenvalue. From $HU = UF$ and the specified ordering of the diagonal blocks of F, we can see that the column space of $(U_{11}^T \;\; U_{21}^T)^T$ is the same n-dimensional invariant subspace (no difficulties will arise when H has a double zero eigenvalue, since there is only one eigenvector (up to a factor) associated with the zero eigenvalue). So,

$$\begin{pmatrix} I \\ S \end{pmatrix} = \begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix} W$$

for some nonsingular $W \in \mathbb{R}^{n\times n}$. Thus, U_{11} is nonsingular and $S = U_{21}U_{11}^{-1}$. □

Remark 5.1. A Wiener-Hopf factorization for (3.1) can be obtained by solving also the dual equation (4.3).

We now consider (1.2) with conditions (1.7) and (1.8). In this case, any nonnegative solution of (1.2) must be positive (see Remark 2.1).

THEOREM 5.2. *Consider* (1.2) *with conditions* (1.7) *and* (1.8). *Let H be the matrix given by* (1.3). *Let U be an orthogonal matrix such that*

$$U^T H U = F$$

is a real Schur form of F, where the 1×1 *or* 2×2 *diagonal blocks of F are arranged in the order for which the real parts of the corresponding* $n+m$ *eigenvalues are nonincreasing.*

(1) *Assume that* λ_n *and* λ_{n+1} *are a conjugate pair corresponding to a* 2×2 *diagonal block,* $\mathrm{Re}(\lambda_{n-1}) > \mathrm{Re}(\lambda_n)$ (*if* $n > 1$), *and* $\mathrm{Re}(\lambda_{n+1}) > \mathrm{Re}(\lambda_{n+2})$ (*if* $m > 1$). *Then* (1.2) *has no positive solutions.*

(2) *Assume that* $\mathrm{Re}(\lambda_n) > \mathrm{Re}(\lambda_{n+1})$ *and* U *is partitioned as*

$$\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

where $U_{11} \in \mathbb{R}^{n\times n}$. *If* U_{11} *is nonsingular and* $S = U_{21}U_{11}^{-1}$ *is positive, then* S *is the minimal positive solution of* (1.2). *Otherwise,* (1.2) *has no positive solutions.*

(3) *Assume that* $\lambda_n = \lambda_{n+1}$ *are real,* $\mathrm{Re}(\lambda_{n-1}) > \lambda_n$ (*if* $n > 1$), *and* $\lambda_{n+1} > \mathrm{Re}(\lambda_{n+2})$ (*if* $m > 1$). *Assume further that there is only one eigenvector* (*up to a factor*) *associated with* $\lambda_n = \lambda_{n+1}$ *and let* U *be partitioned as in part* (2). *If* U_{11} *is nonsingular and* $S = U_{21}U_{11}^{-1}$ *is positive, then* S *is the minimal positive solution of* (1.2). *Otherwise,* (1.2) *has no positive solutions.*

Proof. If (1.2) has a positive solution, then it has a minimal positive solution S by Theorem 2.1. Since $I \otimes (A - SC) + (D - CS)^T \otimes I$ is an M-matrix by Theorem 2.1, the real part of each eigenvalue of $D - CS$ must be greater than or equal to the negative of the real part of each eigenvalue of $A - SC$. In other words, in view of Lemma 4.6, the real part of each eigenvalue of $D - CS$ must be greater than or equal to the real part of each of the remaining m eigenvalues of H. Under the assumptions of part (1), the eigenvalues of the real matrix $D - CS$ must be $\lambda_1, \ldots, \lambda_{n-1}$ and one of the eigenvalues λ_n and λ_{n+1}. This is impossible since λ_n and λ_{n+1} are a conjugate pair with nonzero imaginary parts. Part (1) is thus proved. Under the assumptions of part (2), if (1.2) has a positive solution, then the column space of $(I \;\; S^T)^T$ for the minimal positive solution S must be the n-dimensional invariant subspace of H corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$. The proof can thus be completed as in the proof of Theorem 5.1. Under the assumptions of part (3), if (1.2) has a positive solution, then the column space of $(I \;\; S^T)^T$ for the minimal positive solution S must be the direct sum of the $(n-1)$-dimensional invariant subspace of H corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_{n-1}$ and the one-dimensional eigenspace of H corresponding to $\lambda_n = \lambda_{n+1}$. The proof can again be completed as in the proof of Theorem 5.1. □

Remark 5.2. If the minimal positive solution found by the Schur method in Theorem 5.2 (2) is not accurate enough, we can use Newton's method as a correction method. Local quadratic convergence of Newton's method is guaranteed since the Fréchet derivative at the solution is nonsingular in this case.

Remark 5.3. In Theorem 5.2 (3), the additional assumption that there is only one eigenvector associated with $\lambda_n = \lambda_{n+1}$ is essential. Without this assumption, no definitive information can be obtained about positive solutions of (1.2) from the real Schur form. Newton's method can find the minimal positive solution of (1.2) if it has a positive solution, with or without the additional assumption. However, we cannot expect Newton's method to have quadratic convergence since the Fréchet derivative at the minimal solution is singular in this case.

As we can see from Theorems 5.1 and 5.2, the real Schur form with the prescribed ordering of the diagonal blocks is essential for finding the minimal nonnegative solution using the Schur method. This real Schur form can be obtained by using orthogonal transformations to reduce H to upper Hessenberg form and then using a slight modification of Stewart's algorithm HQR3 [31]. In Stewart's HQR3, the 1×1 or 2×2 diagonal blocks of the real Schur form are arranged in the order for which the moduli (not the real parts) of the corresponding eigenvalues are nonincreasing.

In Theorems 5.1 and 5.2, the minimal nonnegative solution S is found by solving

$SU_{11} = U_{21}$. The accuracy of S is thus dependent on $\kappa(U_{11})$, the condition number of the matrix U_{11}.

PROPOSITION 5.3. *Let $S = U_{21}U_{11}^{-1}$ be the minimal nonnegative solution of* (1.2)*, where U_{11} and U_{21} are as in Theorems* 5.1 *and* 5.2*. Then*

$$\kappa_2(U_{11}) \leq 1 + \|S\|_2^2.$$

Proof. The proof is omitted here since it is very similar to that of corresponding results in [15] and [18]. □

6. Comparison of solution methods. For the equation (1.2) with conditions (1.7) and (1.8), the minimal positive solution can be found by FP1, FP2, FP3, Newton's method, or the Schur method, whenever (1.2) has a positive solution. In this section, we will compare these methods on a few test examples.

For the Newton iteration (2.2), 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 [2] and [11]. If we use the Bartels-Stewart algorithm [2] to solve the Sylvester equation, the computational work for each Newton iteration is about $62n^3$ flops when $m = n$. 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. The Schur method needs roughly $200n^3$ flops to get an approximate solution.

Example 6.1. We generate (and save) a random 100×100 matrix R with no zero elements using rand(100,100) in MATLAB. Let $W = \text{diag}(Re) - R$. So W is a singular M-matrix with no zero elements. We introduce a real parameter α and let

$$\alpha I + W = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix},$$

where the matrices A, B, C, D are all 50×50. The existence of a positive solution of (1.2) is guaranteed for $\alpha \geq 0$. In tables 6.1–6.3, 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 $X_0 = 0$. The initial residual error is $\|\mathcal{R}(X_0)\|_\infty = \|B\|_\infty = 0.2978 \times 10^2$. As predicted by Theorem 2.7, FP2 has faster convergence than FP1 while FP3 has faster convergence than FP2. With the required computational work per iteration in mind, we find that, for this example, FP2 is the best among the three basic fixed-point iterations. When $\alpha = 10$, the fixed-point iterations are quite good. However, Newton's method is much better for $\alpha = 0$. As shown in [14], we can also use Newton's method after any number of fixed-point iterations and still have the monotone convergence. We now apply the Schur method (SM) to find the minimal solution. The method turns out to be very successful. The residual norm for the approximate solution obtained from the Schur method is listed in Table 6.4, along with the residual norm for the approximate solution obtained by Newton's method after $12, 6, 5$ iterations for $\alpha = 0, 1, 10$, respectively. The accuracy achieved by the Schur method is very impressive, although not as high as that achieved by Newton's method. The good performance of the Schur method is partly due to the small condition number of the matrix U_{11}. For $\alpha = 0$, for example, we find that $\kappa_2(U_{11}) = 1.4114$ and $\|S\|_2 = 0.9960$. A rough estimate can actually be obtained beforehand for any $\alpha \geq 0$. Since $Se \leq e$ by the proof of Theorem 3.1, we have $\|S\|_\infty \leq 1$. So, by Proposition 5.3, $\kappa_2(U_{11}) \leq 1 + (\sqrt{50}\|S\|_\infty)^2 \leq 51$. When $\alpha = 0$, the Schur method is much better

than the basic fixed-point iterations. It is also considerably cheaper than Newton's method, although Newton's method produces a more accurate approximation. When $\alpha = -10^{-4}$, the equation also has a positive solution by Theorem 5.2 (2). The residual norm for the approximate solution obtained from the Schur method is 0.7463×10^{-12} while the residual norm for the approximate solution obtained by Newton's method after 13 iterations is 0.1955×10^{-13}. By Theorem 2.4, equation (1.2) has a positive solution for all $\alpha \geq -10^{-4}$. When $\alpha = -10^{-3}$, the equation does not have a positive solution. In this case, Newton's method exhibits no convergence and the Schur method produces a 2×2 block in the middle of the real Schur form (see Theorem 5.2 (1)). When $\alpha = 0$, the matrix (1.3) has 50 eigenvalues in $\mathbb{C}_>$, 49 eigenvalues in $\mathbb{C}_<$, and one zero eigenvalue. The eigenvalue with the smallest positive real part is the real eigenvalue $\lambda_{50} = 0.1790$. Thus, for all $\alpha \geq 0$, the convergence of Newton's method is quadratic and the convergence of basic fixed-point iterations is linear.

TABLE 6.1
Iteration counts for Example 6.1, $\alpha = 0$

ϵ	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
NM	6	9	10	11	12
FP1	196	1515	4003	6564	9125
FP2	154	1173	3050	4977	6904
FP3	96	758	2017	3313	4609

TABLE 6.2
Iteration counts for Example 6.1, $\alpha = 1$

ϵ	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
NM	4	5	5	6	6
FP1	40	71	101	131	161
FP2	30	52	75	97	119
FP3	19	33	47	62	76

TABLE 6.3
Iteration counts for Example 6.1, $\alpha = 10$

ϵ	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
NM	3	4	4	4	5
FP1	14	23	32	41	49
FP2	11	17	23	29	35
FP3	6	10	14	18	22

Example 6.1 is not particularly tough for the basic fixed-point iterations since $\lambda_{50} = 0.1790$ is not too close to zero. The next example *is*.

Example 6.2. Let

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$

be a doubly stochastic matrix (i.e., $R \geq 0, Re = e, R^T e = e$), where $R_{11}, R_{22} \in \mathbb{R}^{m \times m}$ are irreducible and $R_{21}, R_{12} \neq 0$. Let $W = a(I - R)$, where a is a given positive number. So W is a singular M-matrix satisfying the assumptions in Proposition 3.2 and the situation in Theorem 4.7 (1) happens. Let

$$W = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix},$$

TABLE 6.4
Residual errors for Example 6.1

α	0	1	10
NM	0.1999×10^{-13}	0.1570×10^{-13}	0.1149×10^{-13}
SM	0.6419×10^{-12}	0.5715×10^{-12}	0.6984×10^{-12}

where $A, B, C, D \in \mathbb{R}^{m\times m}$. We will find the minimal positive solution of the equation (1.2). As noted in Remark 4.1, the convergence of basic fixed-point iterations will be sublinear and the convergence of Newton's method will typically be linear with rate 1/2. However, the minimal positive solution can be found easily by the Schur method described in Theorem 5.1. Since $We = 0$, we have $Se \le e$ by the proof of Theorem 3.1. Since $W^T e = 0$, we can also get $S^T e \le e$ by taking transpose for (1.2) and applying the proof of Theorem 3.1 to the new equation. Therefore, $S^T Se \le S^T e \le e$. Thus, $\rho(S^T S) \le 1$. Now, by Proposition 5.3, $\kappa_2(U_{11}) \le 1 + \rho(S^T S) \le 2$. We apply the Schur method to a special example with $m = 100$, $B = C = I$, and

$$A = D = \begin{pmatrix} 2 & -1 & & \\ & 2 & \ddots & \\ & & \ddots & -1 \\ -1 & & & 2 \end{pmatrix}.$$

For this example, the minimal solution S must be doubly stochastic. In fact, $Se = e$ follows directly from Theorem 4.8 (1) and $S^T e = e$ is obtained by taking transpose for (1.2) and applying Theorem 4.8 (1) to the new equation. The approximate minimal solution is found by the Schur method with residual error 0.9896×10^{-13}. We also apply Newton's method to this equation. The residual error is 0.5683×10^{-13} after 22 iterations. The performance of Newton's method can be improved significantly by using the double Newton strategy as described in [14]. After 6 Newton iterations and one double Newton step, the residual error is 0.4649×10^{-14}. The basic fixed-point iterations are indeed extremely slow. We apply FP1 to the special example with $m = 5$ instead. It needs 399985 iterations to make the residual error less than 10^{-10}. For this example, if an approximate solution has more digits than needed, chopping is recommended. By using chopping instead of rounding, we will have a much better chance to secure $\sigma(D - CS) \subset \mathbb{C}_{\ge}$, by the theory of nonnegative matrices.

Acknowledgments. The author thanks the referees for their very helpful comments.

REFERENCES

[1] M. T. BARLOW, L. C. G. ROGERS, AND D. WILLIAMS, *Wiener-Hopf factorization for matrices*, in Séminaire de Probabilités XIV, Lecture Notes in Math. 784, Springer, Berlin, 1980, pp. 324–331.

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

[3] P. BENNER AND R. BYERS, *An exact line search method for solving generalized continuous-time algebraic Riccati equations*, IEEE Trans. Automat. Control, 43 (1998), pp. 101–107.

[4] P. BENNER, V. MEHRMANN, AND H. XU, *A new method for computing the stable invariant subspace of a real Hamiltonian matrix*, J. Comput. Appl. Math., 86 (1997), pp. 17–43.

[5] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[6] R. Byers, *A Hamiltonian* QR *algorithm*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 212–229.
[7] W. A. Coppel, *Matrix Quadratic equations*, Bull. Austral. Math. Soc., 10 (1974), pp. 377–401.
[8] J. W. Demmel, *Three methods for refining estimates of invariant subspaces*, Computing, 38 (1987), pp. 43–57.
[9] M. Fiedler and V. Ptak, *On matrices with non-positive off-diagonal elements and positive principal minors*, Czechoslovak. Math. J., 12 (1962), pp. 382–400.
[10] I. Gohberg, P. Lancaster, and L. Rodman, *On Hermitian solutions of the symmetric algebraic Riccati equations*, SIAM J. Control Optim., 24 (1986), pp. 1323–1334.
[11] G. H. Golub, S. Nash, and C. Van Loan, *A Hessenberg-Schur method for the problem* $AX + XB = C$, IEEE Trans. Automat. Control, 24 (1979), pp. 909–913.
[12] C.-H. Guo and P. Lancaster, *Analysis and modification of Newton's method for algebraic Riccati equations*, Math. Comp., 67 (1998), pp. 1089–1105.
[13] C.-H. Guo and A. J. Laub, *On a Newton-like method for solving algebraic Riccati equations*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 694–698.
[14] 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.
[15] N. J. Higham and H.-M. Kim, *Numerical analysis of a quadratic matrix equation*, IMA J. Numer. Anal., 20 (2000), pp. 499–519.
[16] J. Juang, *Existence of algebraic matrix Riccati equations arising in transport theory*, Linear Algebra Appl., 230 (1995), pp. 89–100.
[17] J. Juang and W.-W. Lin, *Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 228–243.
[18] C. Kenney, A. J. Laub, and M. Wette, *A stability-enhancing scaling procedure for Schur-Riccati solvers*, Systems Control Lett., 12 (1989), pp. 241–250.
[19] D. L. Kleinman, *On an iterative technique for Riccati equation computations*, IEEE Trans. Automat. Control, 13 (1968), pp. 114–115.
[20] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995.
[21] P. Lancaster and M. Tismenetsky, *The Theory of Matrices, 2nd ed.*, Academic Press, Orlando, FL, 1985.
[22] A. J. Laub, *A Schur method for solving algebraic Riccati equations*, IEEE Trans. Automat. Control, 24 (1979), pp. 913–921.
[23] R. R. London, H. P. McKean, L. C. G. Rogers, and D. Williams, *A martingale approach to some Wiener-Hopf problems* II, in Séminaire de Probabilités XVI, Lecture Notes in Math. 920, Springer, Berlin, 1982, pp. 68–90.
[24] V. L. Mehrmann, *The Autonomous Linear Quadratic Control Problem*, Lecture Notes in Control and Inform. Sci. 163, Springer-Verlag, Berlin, 1991.
[25] 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.
[26] H.-B. Meyer, *The matrix equation* $AZ + B - ZCZ - ZD = 0$, SIAM J. Appl. Math., 30 (1976), pp. 136–142.
[27] C. Paige and C. Van Loan, *A Schur decomposition for Hamiltonian matrices*, Linear Algebra Appl., 41 (1981), pp. 11–32.
[28] L. C. G. Rogers, *Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains*, Ann. Appl. Probab., 4 (1994), pp. 390–413.
[29] L. C. G. Rogers and Z. Shi, *Computing the invariant law of a fluid model*, J. Appl. Probab., 31 (1994), pp. 885–896.
[30] G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Rev., 15 (1973), pp. 727–764.
[31] G. W. Stewart, *HQR3 and EXCHNG: Fortran subroutines for calculating and ordering the eigenvalues of a real upper Hessenberg matrix*, ACM Trans. Math. Software, 2 (1976), pp. 275–280.
[32] P. Van Dooren, *A generalized eigenvalue approach for solving Riccati equations*, SIAM J. Sci. Statist. Comput., 2 (1981), pp. 121–135.
[33] C. Van Loan, *A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix*, Linear Algebra Appl., 61 (1984), pp. 233–251.
[34] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[35] D. Williams, *A "potential-theoretic" note on the quadratic Wiener-Hopf equation for Q-matrices*, in Séminaire de Probabilités XVI, Lecture Notes in Math. 920, Springer, Berlin, 1982, pp. 91–94.