Annals of Mathematical Sciences and Applications
Volume 7, Number 2, 169–191, 2022

On nonlinear matrix equations from the first standard form

Changli Liu, Jungong Xue*, and Ren-Cang Li†

In numerically solving nonlinear matrix equations, including algebraic Riccati equations, that are associated with the eigenspaces of certain regular matrix pencils by the doubling algorithms, the matrix pencils must first be brought into one of the two standard forms. Conversely, each standard form leads to a kind of nonlinear matrix equations, which are of interest in their own right. In this paper, we are concerned with the nonlinear matrix equations associated with the first standard form (SF1). Under the nonnegativeness assumption, we investigate solution existence and the convergence of the doubling algorithm. We obtain several results that resemble the ones for SF1 derived from an M-matrix algebraic Riccati equation.

AMS 2000 subject classifications: Primary 15A24, 65F30; secondary 65H10.
Keywords and phrases: First standard form, SF1, nonnegative matrix, M-matrix, minimal nonnegative solution, doubling algorithm.

1. Introduction

Huang, Li, and Lin [16] summarized a general framework to apply doubling algorithms for the numerical solutions of several types of nonlinear matrix equations in association with the eigenspaces of certain regular matrix pencils. Within the framework, there are two standard forms called the first standard form (SF1) and the second standard form (SF2) that the regular matrix pencils must be transformed into, if they are not already in, before a doubling algorithm can be applied. SF1 was inspired by the developments previously in [6, 7, 8, 15, 25], while SF2 differs subtly from what has been used in the past [6, 10, 11, 12, 13, 14].

*Supported in part by the NNSFC, China 12171101 and Laboratory of Mathematics for Nonlinear Science, Fudan University, China.
†Supported in part by NSF grants NSF DMS-1719620 and DMS-2009689.

Conversely, SF1 corresponds to a nonlinear matrix equation in the form

$$X = X_0 + F_0 X (I - Y_0 X)^{-1} E_0 \tag{1.1}$$

of unknown $X \in \mathbb{C}^{n\times m}$, where $X_0 \in \mathbb{C}^{n\times m}$, $Y_0 \in \mathbb{C}^{m\times n}$, $E_0 \in \mathbb{C}^{m\times m}$ and $F_0 \in \mathbb{C}^{n\times n}$ are known constant matrices. Several special nonlinear matrix equations must be transformed into (1.1) before the doubling algorithm (DA) can be applied for their efficient numerical solution [16]. They include the continuous-time algebraic Riccati equation (CARE) and the discrete-time algebraic Riccati equation (DARE) arising from the optimal control theory [27], and the M-matrix algebraic Riccati equation (MARE) from applied probability and transportation theory, Markov-modulated fluid queue theory [3, 5, 17, 18, 22, 23], among others. In general, a nonlinear matrix equation has more than one solution, but usually one of the them is of interest for the underlying application. In fact, for (1.1) derived from CARE, DARE, MARE, and the quasi-birth-and-death (QBD) equation, which particular solution to look for is dictated by the applications, and the subsequent success of DA is critically dependent on the properties of the source equations of (1.1). All of these make that nonlinear matrix equation (1.1) is so much more general than CARE, DARE, MARE, and the QBD equation, and begs further studies on its own. This is the purpose of this article.

The rest of this paper is organized as follows. In section 2, we briefly outline SF1 for which the doubling algorithm that is built for. Really, SF1 is defined upon four constant matrices E_0, F_0, X_0 and Y_0 of apt sizes. In section 3, we investigate solution existence and numerical solution of nonlinear matrix equation (1.1), assuming the nonnegativeness of the four matrices. We briefly comment on nonlinear matrix equation (1.1) in the symmetric case straightly from DARE. Finally, we draw our concluding remarks in section 5. We also list a few basic facts about nonnegative and M-matrices in appendix A needed for our analysis.

Notation. Throughout this paper, $\mathbb{C}^{n\times m}$ is the set of all $n\times m$ complex matrices, $\mathbb{C}^n = \mathbb{C}^{n\times 1}$, and $\mathbb{C} = \mathbb{C}^1$. Similarly define $\mathbb{R}^{n\times m}$, $\mathbb{R}^n$, and $\mathbb{R}$ except replacing the word *complex* by *real*. I_n (or simply I if its dimension is clear from the context) is the $n\times n$ identity matrix, and e_j is its jth column. The superscript "$\cdot^{\mathrm{T}}$" takes transpose of a matrix or vector. For $X \in \mathbb{R}^{m\times m}$, $X_{(i,j)}$ refers to its (i,j)th entry. Inequality $X \le Y$ means $X_{(i,j)} \le Y_{(i,j)}$ for all (i,j), and similarly for $X < Y$, $X \ge Y$, and $X > Y$. In particular, $X \ge 0$ means that X is entrywise nonnegative. For a square matrix X, denote by $\rho(X)$ its spectral radius.

2. The first standard form (SF1)

In this section, we will outline the framework for (SF1) only and its associated doubling algorithm. The interested reader is referred to [16] for detail on (SF2) and its associated doubling algorithm.

The first standard form (SF1) [16] takes the form

$$\text{(SF1)} \qquad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_0 \\ 0 & F_0 \end{bmatrix} \end{array}.$$

It generally comes from a nonlinear matrix equation that is equivalent to an eigenvalue problem of a regular matrix pencil $\mathscr{A} - \lambda \mathscr{B}$:

$$\mathscr{A} \begin{bmatrix} I_m \\ X \end{bmatrix} = \mathscr{B} \begin{bmatrix} I_m \\ X \end{bmatrix} M, \tag{2.1a}$$

where $X \in \mathbb{C}^{n \times m}$, $M \in \mathbb{C}^{m \times m}$, and

$$\mathscr{A} = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \end{array}, \quad \mathscr{B} = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \end{array}. \tag{2.1b}$$

Note that X and M are not independent in the sense if one is known, so is the other. To go from the matrix pencil $\mathscr{A} - \lambda \mathscr{B}$ to (SF1), while preserving the eigenspace, i.e., the column space of $\begin{bmatrix} I_m \\ X \end{bmatrix}$, we basically multiply $\mathscr{A}$ and $\mathscr{B}$ from the left by the same nonsingular matrices to yield a new matrix pencil $\mathscr{A}_0 - \lambda \mathscr{B}_0$, taking the standard form (SF1), while ensuring

$$\mathscr{A}_0 \begin{bmatrix} I_m \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I_m \\ X \end{bmatrix} \mathscr{M}. \tag{2.2}$$

Usually, the spectral radius $\rho(\mathscr{M}) < 1$ (sometimes $\rho(\mathscr{M}) = 1$ in the so-called critical case). Next, the doubling algorithm is applied to iteratively produce a sequence of matrix pencils $\mathscr{A}_i - \lambda \mathscr{B}_i$, taking the same form as $\mathscr{A}_0 - \lambda \mathscr{B}_0$:

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{array} \quad \text{for } i = 0, 1, \ldots, \tag{2.3}$$

as in Algorithm 2.1, such that

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i} \quad \text{for } i = 0, 1, \ldots. \tag{2.4}$$

Algorithm 2.1 Doubling Algorithm for (SF1)

Input: $X_0 \in \mathbb{C}^{n\times m}$, $Y_0 \in \mathbb{C}^{m\times n}$, $E_0 \in \mathbb{C}^{m\times m}$, $F_0 \in \mathbb{C}^{n\times n}$.
Output: X_∞ as the limit of X_i if it converges.

for $i = 0, 1, \dots,$ until convergence **do**
  compute E_{i+1}, F_{i+1}, X_{i+1}, Y_{i+1} according to

$$E_{i+1} = E_i(I_m - Y_iX_i)^{-1}E_i, \tag{2.5a}$$
$$F_{i+1} = F_i(I_n - X_iY_i)^{-1}F_i, \tag{2.5b}$$
$$X_{i+1} = X_i + F_i(I_n - X_iY_i)^{-1}X_iE_i \tag{2.5c}$$
$$= X_i + F_iX_i(I_m - Y_iX_i)^{-1}E_i, \tag{2.5d}$$
$$Y_{i+1} = Y_i + E_i(I_m - Y_iX_i)^{-1}Y_iF_i \tag{2.5e}$$
$$= Y_i + E_iY_i(I_n - X_iY_i)^{-1}F_i. \tag{2.5f}$$

end for
return X_i at convergence as the computed solution.

In [16], it was shown that (2.2), after eliminating the matrix $\mathscr{M}$, yields a nonlinear matrix equation

$$X = X_0 + F_0X(I - Y_0X)^{-1}E_0 \tag{2.6a}$$

that was called the *primal equation*. Conversely, by letting $\mathscr{M} = (I - Y_0X)^{-1}E_0$, we find that (2.6a) results in (2.2). Associated with this primal equation, there is the so-called *dual equation*

$$Y = Y_0 + E_0Y(I - X_0Y)^{-1}F_0. \tag{2.6b}$$

Letting $\mathscr{N} = (I - X_0Y)^{-1}F_0$, we find (2.6b) yields

$$\mathscr{A}_0^{(\mathrm{d})}\begin{bmatrix} I \\ Y \end{bmatrix} = \mathscr{B}_0^{(\mathrm{d})}\begin{bmatrix} I \\ Y \end{bmatrix}\mathscr{N}, \tag{2.7}$$

where $\mathscr{A}_0^{(\mathrm{d})} := \Pi_{m,n}^{\mathrm{T}}\mathscr{B}_0\Pi_{m,n}$ and $\mathscr{B}_0^{(\mathrm{d})} := \Pi_{m,n}^{\mathrm{T}}\mathscr{A}_0\Pi_{m,n}$ with $\Pi_{m,n} = \begin{bmatrix} 0 & I_m \\ I_n & 0 \end{bmatrix}$, yielding

$$\mathscr{A}_0^{(\mathrm{d})} = \begin{array}{c} \\ n \\ m \end{array}\!\!\begin{array}{c} \begin{array}{cc} n & m \end{array} \\ \begin{bmatrix} F_0 & 0 \\ -Y_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0^{(\mathrm{d})} = \begin{array}{c} \\ n \\ m \end{array}\!\!\begin{array}{c} \begin{array}{cc} n & m \end{array} \\ \begin{bmatrix} I & -X_0 \\ 0 & E_0 \end{bmatrix} \end{array}. \tag{2.8}$$

All of this is part of the primal-dual view developed in [16]. It can be verified that (2.7) is equivalent to

$$\mathscr{A}_0 \begin{bmatrix} Y \\ I_n \end{bmatrix} \mathscr{N} = \mathscr{B}_0 \begin{bmatrix} Y \\ I_n \end{bmatrix}. \tag{2.9}$$

Returning to Algorithm 2.1, if $\rho(\mathscr{M}) < 1$, then (2.4) implies, as $i \to \infty$,

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} \to 0 \quad \Rightarrow \quad X_i \to X,$$

a solution of the nonlinear equation of interest. But how to ensure $\rho(\mathscr{M}) < 1$? Note that $\mathscr{M}$ does not show up in Algorithm 2.1. In fact, the algorithm takes input $\{X_0, Y_0, E_0, Y_0\}$ and produces output $\{X_i, Y_i, E_i, Y_i\}_{i=0}^{\infty}$, provided all the inverses exist. It turns out that the existing analysis of the doubling algorithm on CARE, DARE, MARE, and the QBD equation [5, 16] traces $\mathscr{M}$ back to the matrix M in (2.1) to gain the information on $\rho(\mathscr{M})$. Departing from this approach, in this paper, we are going to look at Algorithm 2.1 from a different perspective, namely to build a convergence theory only around (2.6) by asking what properties the initial input matrices X_0, Y_0, E_0, Y_0 have in order to ensure the convergence of X_i to a solution of (2.6a). Conceivably such a theory may expand the domain where Algorithm 2.1 may be applied beyond the types of nonlinear matrix equations investigated in the literature.

The reader is referred to [16] for what suitable stopping criteria should be used in Algorithm 2.1 in different circumstances.

3. Nonnegative case

The nonnegative case of (SF1) refers to one with

$$\begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_0 & Y_0 \\ X_0 & F_0 \end{bmatrix} \end{array} \geq 0.$$

We will also assume that there is a positive vector $\boldsymbol{u} \equiv \begin{array}{c} m \\ n \end{array} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} > 0$ such that either

$$\begin{bmatrix} E_0 & Y_0 \\ X_0 & F_0 \end{bmatrix} \boldsymbol{u} < \boldsymbol{u}, \tag{3.1a}$$

or a combination of the following inequalities

$$\begin{bmatrix} E_0 & Y_0 \\ X_0 & F_0 \end{bmatrix} \boldsymbol{u} \le \boldsymbol{u}, \tag{3.1b}$$

$$\begin{bmatrix} E_0 & Y_0 \\ X_0 & F_0 \end{bmatrix} \boldsymbol{u} = \boldsymbol{u}, \tag{3.1c}$$

$$Y_0 u_2 < u_1, \tag{3.1d}$$

$$X_0 u_1 < u_2, \tag{3.1e}$$

$$\text{either } E_0 u_1 > 0 \text{ or } F_0 u_2 > 0. \tag{3.1f}$$

hold. Specifically, our investigation covers two cases:

$$\boxed{\begin{array}{rl} \text{Case I:} & (3.1a); \\ \text{Case II:} & (3.1b) \text{ with one of } (3.1d), (3.1e), \text{ and } (3.1f). \end{array}} \tag{3.2}$$

Remark 3.1. There are a few comments in order.

1. Condition (3.1c) is a special case of (3.1b). It is just that under (3.1c) sometimes we may have stronger conclusions in terms of equalities.
2. Assuming (3.1b), we find that the first condition in (3.1f), i.e., $E_0 u_1 > 0$, implies (3.1d) because $Y_0 u_2 \le u_1 - E_0 u_1 < u_1$, and similarly, the second condition in (3.1f), i.e., $F_0 u_2 > 0$, implies (3.1e).
3. To understand what $E_0 u_1 > 0$ means, we note that $E_0 \ge 0$ and $u_1 > 0$ and thus $E_0 u_1 > 0$ is equivalent to that no row of E_0 is zero. For the same reason, $F_0 u_2 > 0$ is same as that no row of F_0 is zero.

Later we will further divide (3.1a) into two conditions that may or may not be satisfied simultaneously:

$$[E_0, Y_0]\boldsymbol{u} < u_1, \tag{3.1a-1}$$

$$[X_0, F_0]\boldsymbol{u} < u_2. \tag{3.1a-2}$$

One source of such an (SF1) is the numerical solution of an MARE:

$$XDX - AX - XB + C = 0 \tag{3.3a}$$

by the doubling algorithm, where A, B, C, D are matrices whose sizes are determined by the partitioning

$$W = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \left[\begin{array}{cc} B & -D \\ -C & A \end{array} \right] \end{array} \tag{3.3b}$$

and W is a nonsingular M-matrix or an irreducible singular M-matrix. MARE (3.3) corresponds to an eigenvalue problem (2.1) with

$$\mathscr{A} = \begin{bmatrix} I_m & \\ & -I_n \end{bmatrix} \begin{bmatrix} B & -D \\ C & -A \end{bmatrix}, \quad \mathscr{B} = I_{m+n},$$

which eventually gives rise to (SF1) after proper transformations [16]. But we emphasize that not all such (SF1) come from some MARE (3.3).

3.1. Solution existence

Our main goal of this subsection is to establish a collection of results as to when one or both equations in (2.6) have minimal nonnegative solutions, denoted by Φ and Ψ, respectively, under subsets of conditions listed in (3.1), specifically either Case I or Case II in (3.2). Whenever they are proved to exist in what follows, it will also be shown at the same time that

$$\rho(Y_0\Phi) < 1, \quad \rho(X_0\Psi) < 1. \tag{3.4}$$

The minimality of Φ and Ψ is conditional in the following sense:

- for (2.6a), Φ is minimal among all nonnegative solutions X of (2.6a) subject to $\rho(Y_0X) < 1$;
- for (2.6b), Ψ is minimal among all nonnegative solutions Y of (2.6b) subject to $\rho(X_0Y) < 1$.

Under (3.4), $I - Y_0\Phi$ and $I - X_0\Psi$ are nonsingular M-matrices, and we can define

$$\mathscr{M} := (I - Y_0\Phi)^{-1}E_0, \quad \mathscr{N} := (I - X_0\Psi)^{-1}F_0. \tag{3.5}$$

Both are nonnegative. We summarize our main results spread among 4 lemmas into Theorem 3.1, but we point out that it does not include all detailed conclusions of the lemmas. Further, Table 3.1 provides a table-view of the results, where a question mark means nonexistence or an answer yet to be found.

Theorem 3.1. *1. If (3.1a) holds, then both equations in (2.6) have minimal nonnegative solutions, Φ and Ψ, respectively. Moreover*

$$\Phi u_1 < u_2, \quad \mathscr{M} u_1 < u_1, \quad \rho(\mathscr{M}) < 1, \tag{3.6a}$$

$$\Psi u_2 < u_1, \quad \mathscr{N} u_2 < u_2, \quad \rho(\mathscr{N}) < 1. \tag{3.6b}$$

Table 3.1: Summary of solution existence

condition	(3.1a)	(3.1b)			
		(3.1f)	(3.1d)	(3.1e)	(3.1d) & (3.1e)
Φ to (2.6a)	Y	Y	Y	?	Y
Ψ to (2.6b)	Y	Y	?	Y	Y
$\rho(\mathscr{M})$	< 1	≤ 1	≤ 1; < 1 if (3.1a-2)	?	≤ 1; < 1 if (3.1a-2)
$\rho(\mathscr{N})$	< 1	≤ 1	?	≤ 1; < 1 if (3.1a-1)	≤ 1; < 1 if (3.1a-1)
$\rho(\mathscr{M}) \cdot \rho(\mathscr{N})$	< 1	≤ 1	?	?	≤ 1

2. *If (3.1b) and (3.1f) hold, then both equations in (2.6) have minimal nonnegative solutions, Φ and Ψ, respectively. Moreover,*

$$\Phi u_1 \le u_2, \quad \mathscr{M} u_1 \le u_1, \quad \rho(\mathscr{M}) \le 1, \tag{3.7a}$$

$$\Psi u_2 \le u_1, \quad \mathscr{N} u_2 \le u_2, \quad \rho(\mathscr{N}) \le 1. \tag{3.7b}$$

3. *If (3.1b) and (3.1d) hold, then (2.6a) has minimal nonnegative solution Φ. Moreover, we have (3.7a). Additionally, if also (3.1a-2) holds, i.e., $[X_0, F_0]\boldsymbol{u} < u_2$, then we have (3.6a).*
4. *If (3.1b) and (3.1e) hold, then (2.6b) has minimal nonnegative solution Ψ. Moreover, we have (3.7b). Additionally, if also (3.1a-1) holds, i.e., $[E_0, Y_0]\boldsymbol{u} < u_1$, then we have (3.6b).*

Later we will see that a sufficient condition to guarantee for convergence of the doubling algorithm is $\rho(\mathscr{M})\cdot\rho(\mathscr{N}) < 1$. Extracting from this theorem, we get the following corollary.

Corollary 3.1. *If (3.1a) holds, or if (3.1b), (3.1d), and (3.1e), together with one of (3.1a-1) and (3.1a-2), hold, then both equations in (2.6) have minimal nonnegative solutions Φ and Ψ, respectively, for which $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) < 1$.*

For the purpose of establishing this theorem, we design the following fixed-point iterations

$$\widehat{X}_0 = 0, \quad \widehat{X}_{i+1} = X_0 + F_0\widehat{X}_i(I - Y_0\widehat{X}_i)^{-1}E_0 \quad \text{for } i \ge 0, \tag{3.8a}$$

$$\widehat{Y}_0 = 0, \quad \widehat{Y}_{i+1} = Y_0 + E_0\widehat{Y}_i(I - X_0\widehat{Y}_i)^{-1}F_0 \quad \text{for } i \ge 0 \tag{3.8b}$$

to solve the equations in (2.6), respectively. The scheme (3.8a) will generate a sequence $\{\widehat{X}_i\}_{i=0}^{\infty}$ intended to approximate a solution of (2.6a), provided that all $I - Y_0\widehat{X}_i$ are invertible. Similarly, (3.8b) will generate a sequence $\{\widehat{Y}_i\}_{i=0}^{\infty}$ intended to approximate a solution of (2.6b), provided that all $I - X_0\widehat{Y}_i$ are invertible. We will need auxiliary iterations

$$\text{(3.9a)}\qquad \widehat{F}_0 = F_0, \quad \widehat{F}_{i+1} = F_0(I - \widehat{X}_i Y_0)^{-1}\widehat{F}_i \quad \text{for } i \ge 0,$$

$$\text{(3.9b)}\qquad \widehat{E}_0 = E_0, \quad \widehat{E}_{i+1} = E_0(I - \widehat{Y}_i X_0)^{-1}\widehat{E}_i \quad \text{for } i \ge 0$$

in our later proofs. It is clear that $\widehat{X}_1$, $\widehat{Y}_1$, $\widehat{E}_1$, and $\widehat{F}_1$ are well defined and nonnegative because $\widehat{X}_0 = 0$ and $\widehat{Y}_0 = 0$. In fact,

$$\text{(3.10)}\qquad \widehat{X}_1 = X_0,\ \widehat{Y}_1 = Y_0,\ \widehat{F}_1 = F_0^2,\ \widehat{E}_1 = E_0^2.$$

Lemma 3.1. *Suppose that (3.1b) and (3.1d) hold, then for all $i \ge 1$*

(i) $I - Y_0\widehat{X}_{i-1}$ *and* $I - \widehat{X}_{i-1}Y_0$ *are nonsingular M-matrices;*
(ii) $\widehat{X}_i \ge 0$ *and* $\widehat{F}_i \ge 0$*;*
(iii) $[\widehat{X}_i, \widehat{F}_{i-1}]\boldsymbol{u} \le u_2$*;*
(iv) $[\widehat{X}_i, \widehat{F}_{i-1}]\boldsymbol{u} = u_2$ *under (3.1c), a stronger condition than (3.1b).*

As a consequence, $0 = \widehat{X}_0 \le \widehat{X}_1 \le \widehat{X}_2 \le \cdots$, $\widehat{X}_i$ *is bounded, and the sequence* $\{\widehat{X}_i\}_{i=0}^{\infty}$ *converges to, say* $\Phi \ge X_0$*. Furthermore, we have the following statements.*

(a) $I - Y_0\Phi$ *and* $I - \Phi Y_0$ *are nonsingular M-matrices;*
(b) Φ *is the minimal nonnegative solution of (2.6a) among all nonnegative solutions X subject to* $\rho(Y_0X) < 1$*;*
(c) $\Phi u_1 \le [X_0, F_0]\boldsymbol{u} \le u_2$ *and* $\mathscr{M}u_1 \le u_1$*. In particular,* $\rho(\mathscr{M}) \le 1$*;*
(d) *If* $[X_0, F_0]\boldsymbol{u} < u_2$*, then* $\Phi u_1 < u_2$ *and* $\mathscr{M}u_1 < u_1$ *which implies* $\rho(\mathscr{M}) < 1$*.*

Proof. We use the mathematical induction to prove items (i) – (iv). We start by proving items (i) – (iv) for $i = 1$ for which both matrices in item (i) are the identity matrices of apt sizes. Item (ii) holds due to (3.10), item (iii) due to (3.1b), and item (iv) due to (3.1c).

Suppose items (i) – (iv) hold for $i = j$. We will have to prove them for $i = j + 1$. Since $Y_0u_2 < u_1$ by (3.1d), we have

$$Y_0\widehat{X}_j u_1 \le Y_0 u_2 < u_1,$$

which implies $\rho(Y_0\widehat{X}_j) < 1$ by Lemma A.1 in the appendix. Hence $\rho(Y_0\widehat{X}_j) = \rho(\widehat{X}_jY_0) < 1$, too. Therefore both $I - Y_0\widehat{X}_j$ and $I - \widehat{X}_jY_0$ are nonsingular M-matrices. As a consequence, $(I - Y_0\widehat{X}_j)^{-1} \ge 0$ and $(I - \widehat{X}_jY_0)^{-1} \ge 0$, and $\widehat{X}_{j+1}$ and $\widehat{F}_{j+1}$ are well-defined and nonnegative as well. Next, we note the following

$$E_0u_1 \le u_1 - Y_0u_2, \tag{3.11}$$

$$\widehat{X}_j(I - Y_0\widehat{X}_j)^{-1} = (I - \widehat{X}_jY_0)^{-1}\widehat{X}_j, \tag{3.12}$$

$$\begin{aligned}\widehat{X}_j(I - Y_0\widehat{X}_j)^{-1}Y_0 &= (I - \widehat{X}_jY_0)^{-1}\widehat{X}_jY_0\\ &= -I + (I - \widehat{X}_jY_0)^{-1},\end{aligned} \tag{3.13}$$

$$\widehat{X}_ju_1 - u_2 \le -\widehat{F}_{j-1}u_2, \tag{3.14}$$

$$X_0u_1 + F_0u_2 \le u_2. \tag{3.15}$$

It follows from (3.8a) that

$$\begin{aligned}\widehat{X}_{j+1}u_1 &= X_0u_1 + F_0\widehat{X}_j(I - Y_0\widehat{X}_j)^{-1}E_0u_1\\ &\le X_0u_1 + F_0\widehat{X}_j(I - Y_0\widehat{X}_j)^{-1}(u_1 - Y_0u_2) \qquad \text{(by (3.11))}\\ &= X_0u_1 + F_0\widehat{X}_j(I - Y_0\widehat{X}_j)^{-1}u_1 - F_0\widehat{X}_j(I - Y_0\widehat{X}_j)^{-1}Y_0u_2.\end{aligned}$$

Upon using (3.12) and (3.13), we get

$$\begin{aligned}\widehat{X}_{j+1}u_1 &= X_0u_1 + F_0(I - \widehat{X}_jY_0)^{-1}\widehat{X}_ju_1 + F_0u_2 - F_0(I - \widehat{X}_jY_0)^{-1}u_2\\ &= X_0u_1 + F_0u_2 + F_0(I - \widehat{X}_jY_0)^{-1}(\widehat{X}_ju_1 - u_2)\\ &\le u_2 - F_0(I - \widehat{X}_jY_0)^{-1}\widehat{F}_{j-1}u_2 \qquad \text{(by (3.14) and (3.15))}\\ &= u_2 - \widehat{F}_ju_2, \end{aligned} \tag{3.16}$$

i.e., $[\widehat{X}_{j+1}, \widehat{F}_j]\boldsymbol{u} \le u_2$. Furthermore, in the case of (3.1c), all "less than or equal to" signs from (3.11) to (3.16) can be replaced by the "equal to" sign, which implies $[\widehat{X}_{j+1}, \widehat{F}_j]\boldsymbol{u} = u_2$. This completes the proof of items (i) – (iv).

$\widehat{X}_i$ is monotonically increasing and bounded because $\widehat{X}_iu_1 \le u_2$ for all i. Thus it has a limit $\lim_{i\to\infty}\widehat{X}_i =: \Phi$ and $\Phi u_1 \le u_2$. We have

$$Y_0\Phi u_1 \le Y_2u_2 < u_1,$$

implying $\rho(Y_0\Phi) < 1$ by Lemma A.1. Hence $\rho(Y_0\Phi) = \rho(\Phi Y_0) < 1$, too. Therefore both $I - Y_0\Phi$ and $I - \Phi Y_0$ are nonsingular M-matrices, proving item (a).

For item (b), by letting $i \to \infty$ in (3.8a), we see that Φ is a nonnegative solution of (2.6a). Consider any nonnegative solution $\widetilde{\Phi}$ of (2.6a) such that $\rho(Y_0\widetilde{\Phi}) < 1$. we claim that $\widehat{X}_i \le \widetilde{\Phi}$ for $i \ge 0$, which leads to $\Phi \le \widetilde{\Phi}$ upon letting $i \to \infty$. It is true that $\widehat{X}_0 = 0 \le \widetilde{\Phi}$. Suppose $\widehat{X}_i \le \widetilde{\Phi}$ for $i = j$. Then $I - Y_0\widehat{X}_i \ge I - Y_0\widetilde{\Phi}$ which is a nonsingular M-matrix because $\rho(Y_0\widetilde{\Phi}) < 1$. By Lemma A.3, $0 \le (I - Y_0\widehat{X}_i)^{-1} \le (I - Y_0\widetilde{\Phi})^{-1}$. Therefore using (3.8a), we conclude that

$$\widehat{X}_{j+1} \le X_0 + F_0\widetilde{\Phi}(I - Y_0\widetilde{\Phi})^{-1}E_0 = \widetilde{\Phi},$$

completing the induction step. Thus item (b) is proved.

For item (c), we note $\widehat{X}_i u_1 \le u_2$ by items (iii) and (iv). Letting $i \to \infty$, we find $\Phi u_1 \le u_2$. Hence $Y_0 u_2 \ge Y_0 \Phi u_1$ and

$$E_0 u_1 \le u_1 - Y_0 u_2 \le u_1 - Y_0\Phi u_1 = (I - Y_0\Phi)u_1 \tag{3.17}$$

to give $\mathscr{M} u_1 = (I - Y_0\Phi)^{-1}E_0 u_1 \le u_1$. Since Φ is a solution of (2.6a), we get $\Phi = X_0 + F_0\Phi\mathscr{M}$ and thus

$$\Phi u_1 = X_0 u_1 + F_0\Phi\mathscr{M} u_1 \le X_0 u_1 + F_0\Phi u_1 \le X_0 u_1 + F_0 u_2 = [X_0, F_0]\boldsymbol{u} \le u_1. \tag{3.18}$$

Lastly for item (d), suppose $[X_0, F_0]\boldsymbol{u} < u_2$. Then the first inequality in (3.17) becomes strict to give $\mathscr{M} u_1 < u_1$ which, by Lemma A.1, yields $\rho(\mathscr{M}) < 1$, and the last inequality in (3.18) becomes strict to give $\Phi u_1 < u_2$. □

Lemma 3.2. *Suppose that* (3.1b) *and* (3.1e) *hold, then for all* $i \ge 1$

(i) $I - X_0\widehat{Y}_{i-1}$ *and* $I - \widehat{Y}_{i-1}X_0$ *are nonsingular* M*-matrices;*
(ii) $\widehat{Y}_i \ge 0$ *and* $\widehat{E}_i \ge 0$*;*
(iii) $[\widehat{E}_i, \widehat{Y}_{i-1}]\boldsymbol{u} \le u_1$*;*
(iv) $[\widehat{E}_i, \widehat{Y}_{i-1}]\boldsymbol{u} = u_1$ *under* (3.1c), *a stronger condition than* (3.1b).

As a consequence, $0 = \widehat{Y}_0 \le \widehat{Y}_1 \le \widehat{Y}_2 \le \cdots$, $\widehat{Y}_i$ *is bounded, and the sequence* $\{\widehat{Y}_i\}_{i=0}^{\infty}$ *converges to, say* $\Psi \ge Y_0$. *Furthermore, we have the following statements.*

(a) $I - X_0\Psi$ *and* $I - \Psi X_0$ *are nonsingular* M*-matrices;*
(b) Ψ *is the minimal nonnegative solution of* (2.6b) *among all nonnegative solutions* Y *subject to* $\rho(X_0 Y) < 1$*;*
(c) $\Psi u_2 \le [E_0, Y_0]\boldsymbol{u} \le u_1$ *and* $\mathscr{N} u_2 \le u_2$. *In particular,* $\rho(\mathscr{N}) \le 1$*;*
(d) *If, also,* $[E_0, Y_0]\boldsymbol{u} < u_1$, *then* $\Psi u_2 < u_1$ *and* $\mathscr{N} u_2 < u_2$ *which implies* $\rho(\mathscr{N}) < 1$.

Proof. The proof is similar to that of Lemma 3.1. □

Lemma 3.3. *Suppose*[1] (3.1a) *holds. Then all claims of both Lemmas 3.1 and 3.2, except both items (iv) there, are valid.*

Proof. The inequality (3.1a) implies (3.1b), (3.1d), and (3.1e), but not (3.1c). Hence this lemma is a corollary of Lemmas 3.1 and 3.2. □

Lemma 3.4. *Suppose that* (3.1b) *and* (3.1f) *hold, then for all* $i \geq 1$

(i) $I - Y_0\widehat{X}_{i-1}$ *and* $I - \widehat{X}_{i-1}Y_0$ *are nonsingular* M*-matrices;*
(ii) $\widehat{X}_i \geq 0$ *and* $\widehat{F}_i \geq 0$*;*
(iii) $\widehat{F}_i u_2 > 0$ *if* $F_0 u_2 > 0$*;*
(iv) $[\widehat{X}_i, \widehat{F}_{i-1}]\boldsymbol{u} \leq u_2$*;*
(v) $[\widehat{X}_i, \widehat{F}_{i-1}]\boldsymbol{u} = u_2$ *under* (3.1c)*, a stronger condition than* (3.1b)*;*
(vi) $I - X_0\widehat{Y}_{i-1}$ *and* $I - \widehat{Y}_{i-1}X_0$ *are nonsingular* M*-matrices;*
(vii) $\widehat{Y}_i \geq 0$ *and* $\widehat{E}_i \geq 0$*;*
(viii) $\widehat{E}_i u_2 > 0$ *if* $E_0 u_1 > 0$*;*
(ix) $[\widehat{E}_{i-1}, \widehat{Y}_i]\boldsymbol{u} \leq u_1$*;*
(x) $[\widehat{E}_{i-1}, \widehat{Y}_i]\boldsymbol{u} = u_2$ *under* (3.1c)*, a stronger condition than* (3.1b)*.*

As a consequence, $0 = \widehat{X}_0 \leq \widehat{X}_1 \leq \widehat{X}_2 \leq \cdots$ *and* $0 = \widehat{Y}_0 \leq \widehat{Y}_1 \leq \widehat{Y}_2 \leq \cdots$*,* $\widehat{X}_i$ *and* $\widehat{Y}_i$ *are bounded, and the sequence* $\{\widehat{X}_i\}_{i=0}^{\infty}$ *and* $\{\widehat{Y}_i\}_{i=0}^{\infty}$ *converge to, say* $\Phi \geq X_0$ *and* $\Psi \geq Y_0$*, respectively. Items (a) – (d) of both Lemmas 3.1 and 3.2 remain valid.*

Proof. We use the mathematical induction to prove items (i) – (v). Items (vi) – (x) can be proved in the same way.

For $i = 1$, $\widehat{X}_0 = 0$ and $\widehat{F}_0 = F_0$. Items (i) – (v) are evident by assumption (3.1b) or more strongly (3.1c), and (3.1f), and $\widehat{F}_1 = F_0^2$.

Suppose items (i) – (v) hold for $i = j$. We will have to prove them for $i = j + 1$. If $F_0 u_2 > 0$, then $\widehat{F}_j u_2 > 0$ and thus $\widehat{X}_j u_1 \leq u_2 - \widehat{F}_j u_2 < u_2$ to give

$$\widehat{X}_j Y_0 u_2 \leq \widehat{X}_j u_1 < u_2$$

which implies $\rho(\widehat{X}_j Y_0) < 1$ by Lemma A.1. Hence we have $\rho(Y_0\widehat{X}_j) = \rho(\widehat{X}_j Y_0) < 1$. Therefore both $I - Y_0\widehat{X}_j$ and $I - \widehat{X}_j Y_0$ are nonsingular M-matrices. As a consequence, $(I - Y_0\widehat{X}_j)^{-1} \geq 0$ and $(I - \widehat{X}_j Y_0)^{-1} \geq 0$, and $\widehat{X}_{j+1}$ and $\widehat{F}_{j+1}$ are well-defined and nonnegative. Also $\widehat{F}_{j+1}u_2 > 0$ in the case $F_0 u_2 > 0$. Next, we note that the reasonings between (3.11) and (3.16) remain valid, and thus $[\widehat{X}_{j+1}, \widehat{F}_j]\boldsymbol{u} \leq u_2$. Furthermore, in the case of (3.1c), all "less than or equal to" signs as the results of directly using (3.11) can

[1]This implies, in particular, that the conditions in both items (d) of Lemmas 3.1 and 3.2 are satisfied.

be replaced by the "equal to" sign, which implies $[\widehat{X}_{j+1}, \widehat{F}_j]\boldsymbol{u} = u_2$. This completes the proof of items (i) – (v). □

3.2. Eigenvalue problem for $\mathscr{A}_0 - \lambda\mathscr{B}_0$

Any solution X to (2.6a) decouples the eigenvalue problem for $\mathscr{A}_0 - \lambda\mathscr{B}_0$. In fact, we will have (2.2) with $\mathscr{M} = (I - Y_0X)^{-1}E_0$ and

$$\mathscr{A}_0 \begin{bmatrix} I_m & 0 \\ X & I_n \end{bmatrix} \begin{bmatrix} I_m & \\ & F_0 \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I_m & 0 \\ X & I_n \end{bmatrix} \begin{bmatrix} \mathscr{M} & -Y_0 \\ & I - XY_0 \end{bmatrix}. \tag{3.19}$$

Consequently, the eigenvalues of $\mathscr{A}_0 - \lambda\mathscr{B}_0$ is the multiset union of those of

$$\mathscr{M} = (I - Y_0X)^{-1}E_0 \tag{3.20}$$

and the reciprocals of those of $F_0(I - XY_0)^{-1}$. In particular, the infinite is an eigenvalue of $\mathscr{A}_0 - \lambda\mathscr{B}_0$ if F_0 is singular.

When we have a solution pair (X, Y) to both equations in (2.6) such that $I - XY$ is nonsingular, it can be verified that

$$\mathscr{A}_0 \begin{bmatrix} I_m & Y \\ X & I_n \end{bmatrix} \begin{bmatrix} I_m & \\ & \mathscr{N} \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I_m & Y \\ X & I_n \end{bmatrix} \begin{bmatrix} \mathscr{M} & \\ & I_n \end{bmatrix}, \tag{3.21}$$

where $\mathscr{M}$ is the same as (3.20) and

$$\mathscr{N} = (I - X_0Y)^{-1}F_0. \tag{3.22}$$

Immediately, we conclude that the eigenvalues of $\mathscr{A}_0 - \lambda\mathscr{B}_0$ is the multiset union of those of $\mathscr{M}$ and the reciprocals of those of $\mathscr{N}$.

3.3. Convergence of doubling algorithm

Now consider running Algorithm 2.1 with nonnegative matrices X_0, Y_0, E_0, F_0 as inputs. The next theorem says, among many others, the iteration will not breakdown, i.e., all inverse exists, under assumption of (3.1a), or (3.1b) with (3.1f). We point out again that (3.1c) is considered a special case of (3.1b).

Theorem 3.2. *Suppose that in Algorithm 2.1, input matrices X_0, Y_0, E_0, F_0 are nonnegative matrices and satisfy (3.1a), or (3.1b) with (3.1f). Then Algorithm 2.1 does not breakdown, i.e., all inverses exist during the iterative process, and the following statements hold for $i \geq 0$.*

(a) *All* X_i, Y_i, E_i, F_i *are nonnegative;*
(b) *All* $I - X_iY_i$ *and* $I - Y_iX_i$ *are nonsingular* M*-matrices;*
(c) *If* $E_0u_1 > 0$, *then* $E_iu_1 > 0$;
(d) *If* $F_0u_2 > 0$, *then* $F_iu_2 > 0$;
(e) *We have*

$$\begin{bmatrix} E_i & Y_i \\ X_i & F_i \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \le \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{3.23}$$

where the inequality is strict under (3.1a) *or an identity under* (3.1c).

As a consequence of items (a) and (b) and the recursive formulas in (2.5), *we conclude that* $0 \le X_0 \le X_1 \le \cdots$, $0 \le Y_0 \le Y_1 \le \cdots$, *and* X_i *and* Y_i *are bounded.*

Proof. We will prove items (a) – (e) by induction. For $i = 0$, only item (b) needs a proof. It follows from (3.1a) that

$$Y_0u_2 < u_1 - E_0u_1 \le u_1, \quad X_0u_1 < u_2 - F_0u_2 \le u_2, \tag{3.24}$$

or from (3.1b) that

$$Y_0u_2 \le u_1 - E_0u_1 \le u_1, \quad X_0u_1 \le u_2 - F_0u_2 \le u_2. \tag{3.25}$$

If $E_0u_1 > 0$ then the first inequality in (3.25) can be improved to $Y_0u_2 < u_1$; if $F_0u_2 > 0$, then the second inequality in (3.25) can be improved to $X_0u_1 < u_2$. Hence under (3.1a) or under (3.1b) with (3.1f), we have either $Y_0u_2 < u_1$ or $X_0u_1 < u_2$ or both. Consider now $Y_0u_2 < u_1$. For any nonzero row $e_j^{\mathrm{T}}X_0$ of X_0, we have

$$e_j^{\mathrm{T}}X_0(u_1 - Y_0u_2) > 0 \quad \Rightarrow \quad e_j^{\mathrm{T}}X_0Y_0u_2 < e_j^{\mathrm{T}}X_0u_1 \le e_j^{\mathrm{T}}u_2.$$

Since for any zero row $e_j^{\mathrm{T}}X_0 = 0$ of X_0, if any, we clearly have $e_j^{\mathrm{T}}X_0Y_0u_2 = 0 < e_j^{\mathrm{T}}u_2$. Hence $X_0Y_0u_2 < u_2$ which implies $\rho(X_0Y_0) < 1$ by Lemma A.1 and $I - X_0Y_0$ is a nonsingular M-matrix. Since $\rho(Y_0X_0) = \rho(X_0Y_0) < 1$, $I - Y_0X_0$ is a nonsingular M-matrix, too. The case when $X_0u_1 < u_2$ can be dealt with in the same way.

Suppose now all the claims are true for $i \le j$. Consider $i = j+1$. Item (a) for $i = j+1$ is rather straightforward upon using items (a) and (b) for $i = j$ and the recursive formulas in (2.5). To prove items (c) and (d) for $i = j+1$, we note, by the inductive assumption, either $E_ju_1 > 0$, i.e., no row of E_j is zero, or $F_j u_2 > 0$, i.e., no row of F_j is zero. Since also $(I - X_j Y_j)^{-1} \geq 0$ and $(I - Y_j X_j)^{-1} \geq 0$ and both have no zero rows, we have

$$E_{j+1} u_1 = E_j (I_m - Y_j X_j)^{-1} E_j u_1 > 0 \quad \text{if } E_j u_1 > 0;$$
$$F_{j+1} u_2 = F_j (I_m - X_j Y_j)^{-1} F_j u_2 > 0 \quad \text{if } F_j u_2 > 0.$$

This proves items (c) and (d) for $i = j + 1$. Next, we prove item (e) for $i = j+1$. By the inductive assumption, $I - X_j Y_j$ and $I - Y_j X_j$ are nonsingular M-matrices, implying that $\begin{bmatrix} I_m & -Y_j \\ -X_j & I_n \end{bmatrix}$ is also a nonsingular M-matrix because

$$\begin{bmatrix} I_m & -Y_j \\ -X_j & I_n \end{bmatrix}^{-1} = \begin{bmatrix} (I - Y_j X_j)^{-1} & Y_j (I - X_j Y_j)^{-1} \\ (I - X_j Y_j)^{-1} X_j & (I - X_j Y_j)^{-1} \end{bmatrix} \geq 0. \tag{3.26}$$

By the induction hypothesis, we have

$$\begin{bmatrix} E_j & Y_j \\ X_j & F_j \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \leq \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} E_j & 0 \\ 0 & F_j \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \leq \left(I - \begin{bmatrix} 0 & Y_j \\ X_j & 0 \end{bmatrix} \right) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{3.27}$$

Pre-multiply both sides of the last inequality by the nonnegative matrix in (3.26) to get

$$\begin{bmatrix} I_m & -Y_j \\ -X_j & I_n \end{bmatrix}^{-1} \begin{bmatrix} E_j & 0 \\ 0 & F_j \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \leq \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{3.28}$$

Combining all the equations in (2.5) gives

$$\begin{bmatrix} E_{j+1} & Y_{j+1} \\ X_{j+1} & F_{j+1} \end{bmatrix} = \begin{bmatrix} 0 & Y_j \\ X_j & 0 \end{bmatrix} + \begin{bmatrix} E_j & 0 \\ 0 & F_j \end{bmatrix} \begin{bmatrix} I_m & -Y_j \\ -X_j & I_n \end{bmatrix}^{-1} \begin{bmatrix} E_j & 0 \\ 0 & F_j \end{bmatrix}. \tag{3.29}$$

Now using (3.28), we have

$$\begin{aligned} \begin{bmatrix} E_{j+1} & Y_{j+1} \\ X_{j+1} & F_{j+1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} &= \begin{bmatrix} 0 & Y_j \\ X_j & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &\quad + \begin{bmatrix} E_j & 0 \\ 0 & F_j \end{bmatrix} \begin{bmatrix} I_m & -Y_j \\ -X_j & I_n \end{bmatrix}^{-1} \begin{bmatrix} E_j & 0 \\ 0 & F_j \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &\leq \begin{bmatrix} 0 & Y_j \\ X_j & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} E_j & 0 \\ 0 & F_j \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} E_j & Y_j \\ X_j & F_j \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\
&\le \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \qquad (3.30)
\end{aligned}
$$

as expected. Further, it can be checked that all inequalities from (3.27) to (3.30) become equalities under (3.1c), and all inequalities in (3.27), (3.28), and (3.30) are strict under (3.1a).

Lastly for item (b) for $i = j+1$, in the same way as above for proving both $I - X_0Y_0$ and $I - Y_0X_0$ are nonsingular M-matrices, we can prove both $X_{j+1}Y_{j+1}u_2 < u_2$ and $Y_{j+1}X_{j+1}u_1 < u_1$ in the case of (3.1a), or, either $X_{j+1}Y_{j+1}u_2 < u_2$ or $Y_{j+1}X_{j+1}u_1 < u_1$, depending on either $E_{j+1}u_1 > 0$ or $F_{j+1}u_2 > 0$ in the case of (3.1b) with (3.1f). Consequently, $\rho(Y_{j+1}X_{j+1}) = \rho(X_{j+1}Y_{j+1}) < 1$ and $I - X_{j+1}Y_{j+1}$ and $I - Y_{j+1}X_{j+1}$ are nonsingular M-matrices.

The induction proof of items (a) – (e) is completed. □

Previously in Theorem 3.1 and Table 3.1, we summarized various cases for which both Φ and Ψ provably exist and at the same time $\rho(\mathscr{M})\cdot\rho(\mathscr{N}) \le 1$. These cases contain those covered by Theorem 3.2 that guarantees Algorithm 2.1 executes without any breakdown. Next we cite [16, Theorem 3.18] to our interest in this section regarding the convergence of the doubling algorithm in Algorithm 2.1.

Theorem 3.3 ([16])**.** *Suppose that there are solutions $X = \Phi$ and $Y = \Psi$ to the equations in (2.6) such that $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) < 1$, and suppose that* Algorithm 2.1 *executes without any breakdown, i.e., all the inverses exist during the doubling iterations. Then X_i and Y_i converge to Φ and Ψ quadratically, and moreover,*

(3.31a)
$$\limsup_{i\to\infty} \|X_i - \Phi\|^{1/2^i} \le \rho(\mathscr{M}) \cdot \rho(\mathscr{N}), \quad \limsup_{i\to\infty} \|Y_i - \Psi\|^{1/2^i} \le \rho(\mathscr{M}) \cdot \rho(\mathscr{N}),$$

(3.31b)
$$\limsup_{i\to\infty} \|E_i\|^{1/2^i} \le \rho(\mathscr{M}), \qquad \limsup_{i\to\infty} \|F_i\|^{1/2^i} \le \rho(\mathscr{N}).$$

In this theorem, there are two major requirements: 1) the existence of solutions Φ and Ψ such that $\rho(\mathscr{M})\cdot\rho(\mathscr{N}) < 1$, and 2) Algorithm 2.1 executes without any breakdown. The first requirement is answered in Corollary 3.1, while the second requirement is answered by Theorem 3.2.

Convergence of the sequences generated by Algorithm 2.1 is also possible when $\rho(\mathscr{M})\cdot\rho(\mathscr{N})=1$, but additional conditions are needed. The case falls into the critical case. The reader is referred to [16, Theorem 3.26] for more detail.

3.4. Highly accurate implementation

The ideas in [20, 19, 21, 26] can be straightforwardly adopted to yield a highly accurate implementation of Algorithm 2.1 to solve the minimal nonnegative solutions Φ and Ψ, if exist, to the equations in (2.6) with high relative entrywise accuracy, provided

(3.32)

- input matrices X_0, Y_0, E_0, F_0 are nonnegative and have high relative entrywise accuracy;
- there is a positive vector $\boldsymbol{u} \equiv \begin{array}{c} m \\ n \end{array}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} > 0$, such that

  $\boldsymbol{v} := \boldsymbol{u} - \begin{bmatrix} E_0 & Y_0 \\ X_0 & F_0 \end{bmatrix}\boldsymbol{u} \geq 0$ evaluated to $\hat{\boldsymbol{v}} \equiv \begin{array}{c} m \\ n \end{array}\begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} \approx \boldsymbol{v}$ with high relative entrywise accuracy.

For completeness, we detail the implementation in Algorithm 3.1. There are a couple of explanations to make. In step 5, the concept of *triplet representation* is mentioned. It was introduced in [1, 2] as an alternative representation of an M-matrix, say A, as follows:

$$A = \{\text{offdiag}(A), \boldsymbol{p}, \boldsymbol{q}\}, \tag{3.33}$$

where $\text{offdiag}(A) \geq 0$ is the opposite of the off-diagonal part of A, $\boldsymbol{p} > 0$ and $\boldsymbol{q} = A\boldsymbol{p} \geq 0$. Numerically, Alfa, Xue, and Ye [1] presented the GTH-like algorithm that can compute the solution to $A\boldsymbol{x} = \boldsymbol{b} \geq 0$ with high relative entrywise accuracy if a highly relative entrywise accurate triplet representation (3.33) is provided. To save space, we will not give the detail of the GTH-like algorithm here, but refer the reader to [1, 16].

Algorithm 3.1 Highly Accurate ADDA (accADDA)

Input: nonnegative X_0, Y_0, E_0, F_0, and vectors $\boldsymbol{u}$ and $\hat{\boldsymbol{v}}$, as described in (3.32);
Output: the minimal nonnegative solutions Φ (and Ψ if needed).

1: $w_1^{(0)} = \hat{v}_1$, $w_2^{(0)} = \hat{v}_2$, $i = -1$;
2: **repeat**
3: $i = i + 1$;
4: compute

$$v_1^{(i)} = w_1^{(i)} + E_i u_1 + Y_i(F_i u_2 + w_2^{(i)}),$$
$$v_2^{(i)} = w_2^{(i)} + F_i u_2 + X_i(E_i u_1 + w_1^{(i)});$$

5: generate the triplet representations:

$$I_m - Y_i X_i = \{\text{offdiag}(I_m - Y_i X_i), u_1, v_1^{(i)}\},$$
$$I_n - X_i Y_i = \{\text{offdiag}(I_n - X_i Y_i), u_2, v_2^{(i)}\};$$

6: compute E_{i+1}, F_{i+1}, X_{i+1} and Y_{i+1} by (2.5) with the help of the GTH-like algorithm made possible by the triplet representations;
7: compute

$$w_1^{(i+1)} = w_1^{(i)} + E_i(I_m - Y_i X_i)^{-1}[w_1^{(i)} + Y_i w_2^{(i)}],$$
$$w_2^{(i+1)} = w_2^{(i)} + F_i(I_n - X_i Y_i)^{-1}[X_i w_1^{(i)} + w_2^{(i)}]$$

(reuse $E_i(I_m - Y_i X_i)^{-1}$ and $F_i(I_n - X_i Y_i)^{-1}$ that appear in implementing line 6 to reduce work);
8: **until** convergence;
9: **return** the last X_i and Y_i as approximations to Φ and Ψ, respectively.

4. Symmetric case

The symmetric case of (SF1) refers to one with $m = n$ and

$$F_0 = E_0^{\mathrm{H}}, \quad X_0 = X_0^{\mathrm{H}}, \quad Y_0 = Y_0^{\mathrm{H}},$$

where $(\,\cdot\,)^{\mathrm{H}}$ takes the complex conjugate transpose of a matrix. Now (2.6a) takes the form

$$X = X_0 + E_0^{\mathrm{H}} X (I - Y_0 X)^{-1} E_0.$$

This is exactly in one of the forms of DARE for which $X_0 \succeq 0$ (positive semidefinite), $Y_0 \preceq 0$ (negative semidefinite) [16, chapter 4]. In the DARE case, there is a classical theorem regarding its solution existence (see, e.g., [16, Theorem 4.3]), relying on certain stabilizable and detectable properties from the optimal control theory [27]. It is not clear what other conditions, besides those in DARE, to impose for solution existence in general.

5. Concluding remarks

In [16], the authors streamlined the applications of doubling algorithms to solve certain types of nonlinear matrix equations in association with the eigenspaces of regular matrix pencils. In doing so, they formally formulated two standard forms into which the regular matrix pencils must be transformed first, if they are not already in, before a doubling iterative procedure can be applied. This is especially true for nonlinear matrix equations currently solved through the first standard form (SF1), such as DARE, CARE, MARE, and the QBD equation [5]. In general, (SF1) corresponds to a nonlinear matrix equation that may not have any trace of its source equation, and it is of interest in its own right.

In this paper, we launched a study on the nonlinear matrix equation associated directly with (SF1) under the assumption that all involved constant matrices are nonnegative, among others. Various results analogously to the ones for MARE are proved to still hold. But still there are unanswered questions that warrant further investigation. For example, the minimality of the nonnegative solutions Φ and Ψ is conditional, e.g., Φ is minimal among all nonnegative solutions X of (2.6a) subject to $\rho(Y_0X) < 1$, as we pointed out at the beginning of subsection 3.1. Naturally, we would like to know if Φ is minimal among all nonnegative solutions X of (2.6a). Another question is about the critical case when $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) = 1$. We have practically nothing for the situation in the general case of SF1. Previously for CARE, DARE, MARE, and the QBD equation, we do have knowledge on the Jordan canonical form of the associated matrix pencil $\mathscr{A} - \lambda\mathscr{B}$ that allows us to tell the convergence behavior of the doubling algorithm [5, 16].

So far, we have not mentioned anything about the second standard form (SF2) at all. There is a reason for that. In fact, in [16], there is a wealth of studies on the nonlinear matrix equations solved by the doubling algorithm on SF2, and these equations can be directly translated from SF2, without going through any nontrivial transformation as we have to do for CARE, MARE, and the QBD equation.

Appendix A. Preliminaries on M-matrices

In this appendix, we collect some well-known results on nonnegative matrices and M-matrices that are relevant to our arguments in this paper. They can be found in, e.g., [4, 9, 24].

Lemma A.1 ([4, Theorem 1.11])**.** *Let $A \in \mathbb{R}^{n\times n}$ be nonnegative and $\boldsymbol{u} \in \mathbb{R}^n$ be positive, i.e., $A \geq 0$ and $\boldsymbol{u} > 0$.*

(a) *If $A\boldsymbol{u} \leq \boldsymbol{u}$, then $\rho(A) \leq 1$.*
(b) *If $A\boldsymbol{u} < \boldsymbol{u}$, then $\rho(A) < 1$.*

$A \in \mathbb{R}^{n\times n}$ is called a *Z-matrix* if $A_{(i,j)} \leq 0$ for all $i \neq j$ [4, p.284]. Any Z-matrix A can be written as $sI - N$ with $N \geq 0$, and it is called an *M-matrix* if $s \geq \rho(N)$, a *singular M-matrix* if $s = \rho(N)$, and a *nonsingular M-matrix* if $s > \rho(N)$.

Lemma A.2. *For a Z-matrix A, the following are equivalent:*

(a) *A is a nonsingular M-matrix;*
(b) *$A^{-1} \geq 0$;*
(c) *$Au > 0$ for some vector $u > 0$;*
(d) *All eigenvalues of A have positive real parts.*

Lemma A.3. *Let $A \in \mathbb{R}^{n\times n}$ be a nonsingular M-matrix, $B \in \mathbb{R}^{n\times n}$ a Z-matrix. If $A \leq B$, then B is also a nonsingular M-matrix and, moreover, $0 \leq B^{-1} \leq A^{-1}$.*

References

[1] Alfa, A. S., Xue, J. and Ye, Q. (2002). Accurate computation of the smallest eigenvalue of a diagonally dominant M-matrix. *Math. Comp.* **71** 217–236. MR1862996

[2] Alfa, A. S., Xue, J. and Ye, Q. (2002). Entrywise perturbation theory for diagonally dominant M-matrices with applications. *Numer. Math.* **90** 401–414. MR1884223

[3] Bean, N. G., O′Reilly, M. M. and Taylor, P. G. (2005). Algorithms for return probabilities for stochastic fluid flows. *Stoch. Models* **21** 149–184. MR2124363

[4] Berman, A. and Plemmons, R. J. (1994). *Nonnegative Matrices in the Mathematical Sciences.* SIAM, Philadelphia. This SIAM edition is a corrected reproduction of the work first published in 1979 by Academic Press, San Diego, CA. MR0544666

[5] Chen, C., Li, R.-C. and Ma, C. (2019). Highly Accurate Doubling Algorithm for Quadratic Matrix Equation from Quasi-Birth-and-Death Process. *Linear Algebra Appl.* **583** 1–45. MR3997180

[6] Chiang, C.-Y., Chu, E. K.-W., Guo, C.-H., Huang, T.-M., Lin, W.-W. and Xu, S.-F. (2009). Convergence Analysis of the Doubling Algorithm for Several Nonlinear Matrix Equations in the Critical Case. *SIAM J. Matrix Anal. Appl.* **31** 227–247. MR2496417

[7] Chu, E. K. W., Fan, H. Y. and Lin, W. W. (2005). A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations. *Linear Algebra Appl.* **396** 55–80. MR2112199

[8] Chu, E. K. W., Fan, H. Y., Lin, W. W. and Wang, C. S. (2004). Structure-Preserving Algorithms for Periodic Discrete-Time Algebraic Riccati Equations. *Int. J. Control* **77** 767–788. MR2072208

[9] Fiedler, M. (2008). *Special Matrices and Their Applications in Numerical Mathematics*, 2nd ed. Dover Publications, Inc., Mineola, New York. MR2517585

[10] Guo, C. H., Kuo, Y. C. and Lin, W. W. (2011). Complex symmetric stabilizing solution of the matrix equation $X + A^\top X^{-1} A = Q$. *Linear Algebra Appl.* **435** 1187–1192. MR2807142

[11] Guo, C. H., Kuo, Y. C. and Lin, W. W. (2012). Numerical solution of nonlinear matrix equations arising from Green's function calculations in nano research. *J. Comput. Appl. Math.* **236** 4166–4180. MR2942418

[12] Guo, C. H., Kuo, Y. C. and Lin, W. W. (2012). On a Nonlinear Matrix Equation Arising in Nano Research. *SIAM J. Matrix Anal. Appl.* **33** 235–262. MR2902680

[13] Guo, C. H. and Lin, W. (2010). Solving a Structured Quadratic Eigenvalue Problem by a Structure-Preserving Doubling Algorithm. *SIAM J. Matrix Anal. Appl.* **31** 2784–2801. MR2740633

[14] Guo, C. H. and Lin, W. W. (2010). The matrix equation $X + A^\top X^{-1} A = Q$ and its application in nano research. *SIAM J. Sci. Comput.* **32** 3020–3038. MR2729449

[15] Guo, X., Lin, W. and Xu, S. (2006). A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation. *Numer. Math.* **103** 393–412. MR2221055

[16] Huang, T.-M., Li, R.-C. and Lin, W.-W. (2018). *Structure-Preserving Doubling Algorithms For Nonlinear Matrix Equations. Fundamentals of Algorithms* **14**. SIAM, Philadelphia. MR3909433

[17] Juang, J. (1995). Existence of algebraic matrix Riccati equations arising in transport theory. *Linear Algebra Appl.* **230** 89–100. MR1355689

[18] Latouche, G. and Taylor, P. G. (2009). A stochastic fluid model for an ad hoc mobile network. *Queueing Syst.* **63** 109–129. MR2576008

[19] Liu, C., Wang, W.-G., Xue, J. and Li, R.-C. (2021). Accurate numerical solution for structured M-matrix algebraic Riccati equations. *J. Comput. Appl. Math.* **396**. MR4254546

[20] Liu, C., Xue, J. and Li, R.-C. (2020). Accurate Numerical Solution For Shifted M-Matrix Algebraic Riccati Equations. *J. Sci. Comput.* **84**. MR4121186

[21] Nguyen, G. T. and Poloni, F. (2015). Componentwise accurate fluid queue computations using doubling algorithms. *Numer. Math.* **130** 763–792. MR3357656

[22] Ramaswami, V. (1999). Matrix analytic methods for stochastic fluid flows. *Proceedings of the 16th International Teletraffic Congress* 19–30. Elsevier Science, Edinburgh.

[23] Rogers, L. (1994). Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains. *Ann. Appl. Probab.* **4** 390–413. MR1272732

[24] Varga, R. S. (1962). *Matrix Iterative Analysis*. Prentice Hall, Englewood Cliffs, NJ. MR0158502

[25] Wang, W.-G., Wang, W.-C. and Li, R.-C. (2012). Alternating-Directional Doubling Algorithm for M-Matrix Algebraic Riccati Equations. *SIAM J. Matrix Anal. Appl.* **33** 170–194. MR2902677

[26] Xue, J. and Li, R.-C. (2017). Highly Accurate Doubling Algorithms for M-matrix Algebraic Riccati Equations. *Numer. Math.* **135** 733–767. MR3606461

[27] Zhou, K., Doyle, J. C. and Glover, K. (1995). *Robust and Optimal Control*. Prentice Hall, Upper Saddle River, New Jersey.

Changli Liu
College of Mathematics
Sichuan University
Chengdu 610065, PR China
E-mail address: chliliu@hotmail.com

Jungong Xue
School of Mathematical Science
Fudan University
Shanghai 200433, PR China
E-mail address: xuej@fudan.edu.cn

Ren-Cang Li
Department of Mathematics
University of Texas at Arlington
Arlington, TX 76019, USA,
and
Department of Mathematics
Hong Kong Baptist University
Kowloon, Hong Kong
E-mail address: rcli@uta.edu

Received February 22, 2022