SIAM J. MATRIX ANAL. APPL.
Vol. 33, No. 1, pp. 170–194

ALTERNATING-DIRECTIONAL DOUBLING ALGORITHM FOR M-MATRIX ALGEBRAIC RICCATI EQUATIONS*

WEI-GUO WANG†, WEI-CHAO WANG‡, AND REN-CANG LI‡

Abstract. A new doubling algorithm—the alternating-directional doubling algorithm (ADDA)—is developed for computing the unique minimal nonnegative solution of an M-matrix algebraic Riccati equation (MARE). It is argued by both theoretical analysis and numerical experiments that ADDA is always faster than two existing doubling algorithms: SDA of Guo, Lin, and Xu (*Numer. Math.*, 103 (2006), pp. 393–412) and SDA-ss of Bini, Meini, and Poloni (*Numer. Math.*, 116 (2010), pp. 553–578) for the same purpose. Also demonstrated is that all three methods are capable of delivering minimal nonnegative solutions with entrywise relative accuracies as warranted by the defining coefficient matrices of a MARE. The three doubling algorithms, differing only in their initial setups, correspond to three special cases of the general *bilinear* (also called *Möbius*) transformation. It is explained that ADDA is the best among all possible doubling algorithms resulted from all bilinear transformations.

Key words. matrix Riccati equation, M-matrix, minimal nonnegative solution, doubling algorithm, bilinear transformation

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

DOI. 10.1137/110835463

1. Introduction.

An *M-matrix algebraic Riccati equation*[1] (MARE) is the matrix equation

$$XDX - AX - XB + C = 0, \tag{1.1}$$

for which A, B, C, and D are matrices whose sizes are determined by the partitioning

$$W = \begin{array}{c} \\ m \\ n \end{array} \begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{pmatrix} B & -D \\ -C & A \end{pmatrix} \end{array} \tag{1.2}$$

and W is a nonsingular or an irreducible singular M-matrix. This kind of Riccati equation arises in applied probability and transportation theory and has been attracting a lot of attention lately. See [16, 19, 21, 22, 23, 24, 29] and the references therein. It is shown in [16, 21] that (1.1) has a unique minimal nonnegative solution Φ, i.e., entrywise

$$\Phi \le X \quad \text{for any other nonnegative solution } X \text{ of (1.1).}$$

*Received by the editors May 26, 2011; accepted for publication (in revised form) December 7, 2011; published electronically March 8, 2012.
http://www.siam.org/journals/simax/33-1/83546.html

†School of Mathematical Sciences, Ocean University of China, Qingdao, 266100, People's Republic of China (wgwang@ouc.edu.cn). This author's work was supported in part by National Natural Science Foundation of China grants 10971204 and 11071228, the China Scholarship Council, Shandong Province Natural Science Foundation grant Y2008A07, and Fundamental Research Funds for the Central Universities grant 201013048. This work was done while this author was a visiting scholar at the Department of Mathematics, University of Texas at Arlington, from September 2010 to August 2011.

‡Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019 (weichao.wang@mavs.uta.edu, rcli@uta.edu). The research of the second author was supported in part by National Science Foundation grant DMS-0810506. The research of the third author was supported in part by National Science Foundation grants DMS-0810506 and DMS-1115834.

[1]Previously it was called a nonsymmetric algebraic Riccati equation, a name that seems to be too broad to be descriptive. MARE was recently coined in [36] to better reflect its characteristics.

In [22], a structure-preserving doubling algorithm (SDA) was proposed by Guo, Lin, and Xu and analyzed for a MARE with W being a nonsingular M-matrix. SDA is very fast and efficient for small to medium MAREs as it is globally and quadratically convergent. Later in [20], it was argued that SDA still works for the case in which W is an irreducible singular M-matrix. The algorithm has to select a parameter that is no smaller than the largest diagonal entries in both A and B. Such a choice of the parameter ensures the following:

1. an elegant theory of global and quadratic convergence [20, 22], except for the *null recurrent* or *critical case* [20, p. 1085] (see also Theorem 3.1(d) below) for which only linear convergence is ensured [11];
2. computed Φ that has an entrywise relative accuracy as the input data deserves, as argued recently in [36].

Consequently, SDA has since emerged as one of the most efficient algorithms.

But as we shall argue in this paper, SDA has room to improve. One situation is when A and B differ in magnitude. But since SDA is blind to any magnitude difference between A and B, it still picks *one* parameter. Conceivably, if A and B could be treated differently with regard to their own characteristics, better algorithms would be possible. This is the motivational thought that drives our study in this paper. Specifically, we will propose a new doubling algorithm—the *alternating-directional doubling algorithm* (ADDA)—that also imports the idea from the alternating-directional-implicit (ADI) iteration for Sylvester equations [6, 33]. Our new doubling algorithm ADDA includes two parameters that can be tailored to reflect each individual characteristic of A and B, and consequently ADDA converges at least as fast as SDA and can be much faster when A and B have very different magnitudes.

We are not the first to notice that SDA often takes many iterations for a MARE with A and B having very different magnitudes. Guo [18] knew it. Bini, Meini, and Poloni [9] recently developed a doubling algorithm called *SDA-ss* using a shrink-and-shift approach of Ramaswami [28]. SDA-ss has shown dramatic improvements over SDA in some of the numerical tests in [9]. But it can happen that sometimes SDA-ss runs slower than SDA, although not by much. Later we will show our ADDA is always the fastest among all possible doubling algorithms derivable from all bilinear transformations, including these three methods.

Throughout this article, A, B, C, and D, unless explicitly stated differently, are reserved for the coefficient matrices of MARE (1.1) for which

(1.3) W defined by (1.2) is a nonsingular M-matrix or an irreducible singular M-matrix.

The rest of this paper is organized as follows. Section 2 presents several known results about M-matrices, as well as a new result on optimizing the product of two spectral radii of the generalized Cayley transforms of two M-matrices. This new result, which may be of independent interest of its own, will be used to develop our optimal ADDA. Section 3 is devoted to the development of ADDA, whose application to the M-matrix Sylvester equation leads to an improvement of the Smith method [30] in section 4. A detailed comparison on rates of convergence among ADDA, SDA, and SDA-ss is given in section 5. Section 6 enumerates all possible doubling algorithms derivable from the general bilinear transformation and concludes that ADDA is the best among all. Numerical results to demonstrate the efficiency of the three doubling methods are presented in section 7. Finally, we give our concluding remarks in section 8.

Notation. $\mathbb{R}^{n\times m}$ is the set of all $n\times m$ real matrices, $\mathbb{R}^n=\mathbb{R}^{n\times 1}$, and $\mathbb{R}=\mathbb{R}^1$. 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. $\mathbf{1}_{n,m}\in\mathbb{R}^{n\times m}$ is the matrix of all ones, and $\mathbf{1}_n=\mathbf{1}_{n,1}$. The superscript T takes the transpose of a matrix or a vector. For $X\in\mathbb{R}^{n\times m}$,

1. $X_{(i,j)}$ refers to its (i,j)th entry;
2. when $m=n$, $\mathrm{diag}(X)$ is the diagonal matrix with the same diagonal entries as X's, $\rho(X)$ is the spectral radius of X, and

$$\varrho(X)=\rho([\mathrm{diag}(X)]^{-1}[\mathrm{diag}(X)-X]).$$

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 $X>Y$. $\|X\|$ denotes some (general) matrix norm of X.

2. Preliminary results on *M*-matrices. $A\in\mathbb{R}^{n\times n}$ is called a *Z-matrix* if $A_{(i,j)}\le 0$ for all $i\ne j$ [7, p. 284]. Any Z-matrix A can be written as $sI-N$ with $N\ge 0$, and it is called an *M-matrix* if $s\ge\rho(N)$, a *singular M-matrix* if $s=\rho(N)$, and a *nonsingular M-matrix* if $s>\rho(N)$.

In this section, we first collect a few well-known results about M-matrices in Lemmas 2.1 and 2.2 that are needed later in this paper. They can be found in, e.g., [7, 14, 32]. Then we establish a new result on optimizing the product of two spectral radii of the generalized Cayley transforms of two M-matrices.

Lemma 2.1 gives four equivalent statements about when a Z-matrix is an M-matrix.

LEMMA 2.1. *For a Z-matrix A, the following are equivalent:*

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

Lemma 2.2 collects a few properties of M-matrices, important to our later analysis, where item (e) can be found in [27].

LEMMA 2.2. *Let* $A,B\in\mathbb{R}^{n\times n}$, *and suppose A is an M-matrix and B is a Z-matrix.*

(a) *If* $B\ge A$, *then B is an M-matrix. In particular,* $\theta I+A$ *is an M-matrix for* $\theta\ge 0$ *and a nonsingular M-matrix for* $\theta>0$.
(b) *If* $B\ge A$ *and A is nonsingular, then B is a nonsingular M-matrix, and* $A^{-1}\ge B^{-1}$.
(c) *If A is nonsingular and irreducible, then* $A^{-1}>0$.
(d) *The one with the smallest absolute value among all eigenvalues of A, denoted by* $\lambda_{\min}(A)$, *is nonnegative, and* $\lambda_{\min}(A)\le\max_i A_{(i,i)}$.
(e) *If A is a nonsingular M-matrix or an irreducible singular M-matrix and is partitioned as*

$$A=\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix},$$

where A_{11} *and* A_{22} *are square matrices, then* A_{11} *and* A_{22} *are nonsingular M-matrices, and their Schur complements*

$$A_{22}-A_{21}A_{11}^{-1}A_{12},\quad A_{11}-A_{12}A_{22}^{-1}A_{21}$$

are nonsingular M-matrices if A is a nonsingular M-matrix or an irreducible singular M-matrices if A is an irreducible singular M-matrix.

Theorem 2.3 below, which may have independent interest of its own, lays the foundation of our optimal ADDA in terms of its rate of convergence subject to certain nonnegativity condition. To the best of our knowledge, it is new.

Define the *generalized Cayley transformation*

$$\mathscr{C}(A;\alpha,\beta) \stackrel{\text{def}}{=} (A-\alpha I)(A+\beta I)^{-1} \tag{2.1}$$

of a square matrix A, where α,β are scalars such that $A+\beta I$ is nonsingular. Given square matrices A and B, define

$$f(\alpha,\beta) \stackrel{\text{def}}{=} \rho(\mathscr{C}(A;\alpha,\beta))\cdot\rho(\mathscr{C}(B;\beta,\alpha)), \tag{2.2}$$

$$g(\beta) \stackrel{\text{def}}{=} \rho\big((A+\beta I)^{-1}\big)\cdot\rho(B-\beta I), \tag{2.3}$$

provided all involved inverses exist. It can be seen that $g(\beta)\equiv f(0,\beta)$.

THEOREM 2.3. *For two M-matrices $A\in\mathbb{R}^{n\times n}$ and $B\in\mathbb{R}^{m\times m}$, define f and g by* (2.2) *and* (2.3)*, and set*

$$\alpha_{\text{opt}} \stackrel{\text{def}}{=} \max_i A_{(i,i)}, \quad \beta_{\text{opt}} \stackrel{\text{def}}{=} \max_i B_{(i,i)}. \tag{2.4}$$

(a) *If both A and B are singular, then $f(\alpha,\beta)\equiv 1$ for $\alpha>\alpha_{\text{opt}}$ and $\beta>\beta_{\text{opt}}$, and $g(\beta)\equiv 1$ for $\beta>\beta_{\text{opt}}$.*

(b) *If one of A and B is nonsingular, then $f(\alpha,\beta)$ for $\alpha>\alpha_{\text{opt}}$ and $\beta>\beta_{\text{opt}}$ is strictly increasing in α and β and $f(\alpha,\beta)<1$, and $g(\beta)$ for $\beta>\beta_{\text{opt}}$ is strictly increasing in β and $g(\beta)<1$.*

Consequently, f can be defined by continuity for all $\alpha\ge\alpha_{\text{opt}}$ and $\beta\ge\beta_{\text{opt}}$ and g can be defined by continuity for all $\beta\ge\beta_{\text{opt}}$. Moreover, we have

$$\min_{\alpha\ge\alpha_{\text{opt}},\beta\ge\beta_{\text{opt}}} f(\alpha,\beta) = f(\alpha_{\text{opt}},\beta_{\text{opt}}), \quad \min_{\beta\ge\beta_{\text{opt}}} g(\beta) = g(\beta_{\text{opt}}). \tag{2.5}$$

Proof. Both $A+\beta I$ and $B+\alpha I$ are nonsingular M-matrices for $\alpha>0$ and $\beta>0$; thus f and g are well-defined for $\alpha>\alpha_{\text{opt}}$ and $\beta>\beta_{\text{opt}}$ since $\alpha_{\text{opt}}\ge 0$ and $\beta_{\text{opt}}\ge 0$. In what follows, we will prove the claims for f only. Similar arguments work for g and thus are omitted.

Assume for the moment that both A and B are irreducible M-matrices. Write $A=sI-N$, where $s\ge 0$ and $N\ge 0$ and N is irreducible. By the Perron–Frobenius theorem [7, p. 27], there is a positive vector u such that $Nu=\rho(N)u$. It can be seen that $\lambda_{\min}(A)=s-\rho(N)\ge 0$, where $\lambda_{\min}(A)$ is as defined in Lemma 2.2(d). We have

$$-\mathscr{C}(A;\alpha,\beta)u = (\alpha I-A)(A+\beta I)^{-1}u = [\alpha-\lambda_{\min}(A)][\lambda_{\min}(A)+\beta]^{-1}u.$$

Since $-\mathscr{C}(A;\alpha,\beta)\ge 0$ and irreducible for $\alpha>\alpha_{\text{opt}}$ and $\beta>0$, it follows from the Perron–Frobenius theorem that

$$\rho(\mathscr{C}(A;\alpha,\beta)) = \rho(-\mathscr{C}(A;\alpha,\beta)) = [\alpha-\lambda_{\min}(A)][\lambda_{\min}(A)+\beta]^{-1}.$$

Similarly, we have for $\alpha>0$ and $\beta>\beta_{\text{opt}}$,

$$\rho(\mathscr{C}(B;\beta,\alpha)) = [\beta-\lambda_{\min}(B)][\lambda_{\min}(B)+\alpha]^{-1}.$$

Finally for $\alpha>\alpha_{\text{opt}}$ and $\beta>\beta_{\text{opt}}$,

$$\begin{aligned} f(\alpha,\beta) &= \rho(\mathscr{C}(A;\alpha,\beta))\cdot\rho(\mathscr{C}(B;\beta,\alpha)) \\ &= \frac{\alpha-\lambda_{\min}(A)}{\lambda_{\min}(A)+\beta}\cdot\frac{\beta-\lambda_{\min}(B)}{\lambda_{\min}(B)+\alpha} \\ &= h_1(\alpha)h_2(\beta), \end{aligned}$$

where

$$h_1(\alpha) = \frac{\alpha - \lambda_{\min}(A)}{\lambda_{\min}(B) + \alpha}, \quad h_2(\beta) = \frac{\beta - \lambda_{\min}(B)}{\lambda_{\min}(A) + \beta}.$$

Now if both A and B are singular, then $\lambda_{\min}(A) = \lambda_{\min}(B) = 0$ and thus $f(\alpha, \beta) \equiv 1$, which proves item (a). If one of A and B is nonsingular, then $\lambda_{\min}(A) + \lambda_{\min}(B) > 0$ and thus

$$h_1'(\alpha) = \frac{\lambda_{\min}(A) + \lambda_{\min}(B)}{(\lambda_{\min}(B) + \alpha)^2} > 0, \quad h_2'(\beta) = \frac{\lambda_{\min}(A) + \lambda_{\min}(B)}{(\lambda_{\min}(A) + \beta)^2} > 0.$$

So $f(\alpha, \beta)$ is strictly increasing in α and β for $\alpha > \alpha_{\text{opt}}$ and $\beta > \beta_{\text{opt}}$ and

$$f(\alpha, \beta) < \lim_{\substack{\alpha \to \infty \\ \beta \to \infty}} f(\alpha, \beta) = 1.$$

This is item (b).

Suppose now that A and B are possibly reducible. Let $\Pi_1 \in \mathbb{R}^{n \times n}$ and $\Pi_2 \in \mathbb{R}^{m \times m}$ be two permutation matrices such that

$$\Pi_1^{\mathrm{T}} A \Pi_1 = \begin{pmatrix} A_{11} & -A_{12} & \dots & -A_{1q} \\ & A_{22} & \dots & -A_{2q} \\ & & \ddots & \vdots \\ & & & A_{qq} \end{pmatrix}, \quad \Pi_2^{\mathrm{T}} B \Pi_2 = \begin{pmatrix} B_{11} & -B_{12} & \dots & -B_{1p} \\ & B_{22} & \dots & -B_{2p} \\ & & \ddots & \vdots \\ & & & B_{pp} \end{pmatrix},$$

where $A_{ij} \in \mathbb{R}^{n_i \times n_j}$, $B_{ij} \in \mathbb{R}^{m_i \times m_j}$, all A_{ii} and B_{jj} are irreducible M-matrices, and all $A_{ij} \geq 0$ and $B_{ij} \geq 0$ for $i \neq j$. It can be seen that

$$f(\alpha, \beta) = \max_{i,j} \rho(\mathscr{C}(A_{ii}; \alpha, \beta)) \cdot \rho(\mathscr{C}(B_{jj}; \beta, \alpha)).$$

If one of A and B is nonsingular, then one of A_{ii} and B_{jj} is nonsingular for each pair (A_{ii}, B_{jj}) and thus all $\rho(\mathscr{C}(A_{ii}; \alpha, \beta)) \cdot \rho(\mathscr{C}(B_{jj}; \beta, \alpha))$ are strictly increasing in α and β for $\alpha > \alpha_{\text{opt}}$ and $\beta > \beta_{\text{opt}}$; so is $f(\alpha, \beta)$. Now if both A and B are singular, then there is at least one pair (A_{ii}, B_{jj}) for which both A_{ii} and B_{jj} are singular and irreducible. By item (a) we just proved for the irreducible and singular case, for that pair $\rho(\mathscr{C}(A_{ii}; \alpha, \beta)) \cdot \rho(\mathscr{C}(B_{jj}; \beta, \alpha)) \equiv 1$ for $\alpha \geq \alpha_{\text{opt}}$ and $\beta \geq \beta_{\text{opt}}$. Since for all other pairs (A_{ii}, B_{jj}),

$$\rho(\mathscr{C}(A_{ii}; \alpha, \beta)) \cdot \rho(\mathscr{C}(B_{jj}; \beta, \alpha)) \leq 1$$

by item (a). Thus $f(\alpha, \beta) \equiv 1$. ∎

3. ADDA. The basic idea of the doubling algorithm for an iterative scheme is to compute only the 2^kth approximations, instead of every approximation in the process. It traces back to the 1970s (see [2] and references therein). Recent resurgence of interest in the idea has led to efficient doubling algorithms for various nonlinear matrix equations. The interested reader is referred to [11] for a more general presentation. The use of an SDA to solve a MARE was first proposed and analyzed by Guo, Lin, and Xu [22]. For MARE (1.1), SDA simultaneously computes the minimal nonnegative solutions of (1.1) and its *complementary M-matrix algebraic Riccati equation* (cMARE)

$$YCY - YA - BY + D = 0. \tag{3.1}$$

In what follows, we shall present our ADDA for MARE in this way: framework, analysis, and then optimal ADDA. We name it ADDA after taking into consideration that it is a doubling algorithm and relates to the ADI iteration for Sylvester equations (see section 4).

3.1. Framework. The framework in this subsection works for any algebraic Riccati equation, provided all involved inverses exist. In general, we are not able to establish a convergence theory similar to the one to be given in the next subsection for a MARE.

For any solution X of MARE (1.1) and Y of cMARE (3.1), it can be verified that

$$H\begin{pmatrix} I \\ X \end{pmatrix} = \begin{pmatrix} I \\ X \end{pmatrix} R, \quad H\begin{pmatrix} Y \\ I \end{pmatrix} = \begin{pmatrix} Y \\ I \end{pmatrix} (-S), \tag{3.2}$$

where

$$H = \begin{pmatrix} B & -D \\ C & -A \end{pmatrix}, \quad R = B - DX, \quad S = A - CY. \tag{3.3}$$

Given any scalars α and β, we have

$$\begin{aligned} (H - \beta I)\begin{pmatrix} I \\ X \end{pmatrix}(\quad R + \alpha I) &= (H + \alpha I)\begin{pmatrix} I \\ X \end{pmatrix}(\quad R - \beta I), \\ (H - \beta I)\begin{pmatrix} Y \\ I \end{pmatrix}(-S + \alpha I) &= (H + \alpha I)\begin{pmatrix} Y \\ I \end{pmatrix}(-S - \beta I). \end{aligned}$$

If $R + \alpha I$ and $S + \beta I$ are nonsingular, then

$$(H - \beta I)\begin{pmatrix} I \\ X \end{pmatrix} = (H + \alpha I)\begin{pmatrix} I \\ X \end{pmatrix}\mathscr{C}(R; \beta, \alpha), \tag{3.4a}$$

$$(H - \beta I)\begin{pmatrix} Y \\ I \end{pmatrix}\mathscr{C}(S; \alpha, \beta) = (H + \alpha I)\begin{pmatrix} Y \\ I \end{pmatrix}. \tag{3.4b}$$

Suppose for the moment that $A + \beta I$ and $B + \alpha I$ are nonsingular and set

$$A_\beta = A + \beta I, \qquad B_\alpha = B + \alpha I, \tag{3.5}$$

$$U_{\alpha\beta} = A_\beta - CB_\alpha^{-1}D, \quad V_{\alpha\beta} = B_\alpha - DA_\beta^{-1}C, \tag{3.6}$$

and

$$Z_1 = \begin{pmatrix} B_\alpha^{-1} & 0 \\ -CB_\alpha^{-1} & I \end{pmatrix}, \; Z_2 = \begin{pmatrix} I & 0 \\ 0 & -U_{\alpha\beta}^{-1} \end{pmatrix}, \; Z_3 = \begin{pmatrix} I & B_\alpha^{-1}D \\ 0 & I \end{pmatrix}.$$

It can be verified that

$$M_0 \overset{\text{def}}{=} Z_3Z_2Z_1(H - \beta I) = \begin{pmatrix} E_0 & 0 \\ -X_0 & I \end{pmatrix}, \tag{3.7a}$$

$$L_0 \overset{\text{def}}{=} Z_3Z_2Z_1(H + \alpha I) = \begin{pmatrix} I & -Y_0 \\ 0 & F_0 \end{pmatrix}, \tag{3.7b}$$

where

$$E_0 = I - (\beta + \alpha)V_{\alpha\beta}^{-1}, \quad Y_0 = (\beta + \alpha)B_\alpha^{-1}DU_{\alpha\beta}^{-1}, \tag{3.8a}$$

$$F_0 = I - (\beta + \alpha)U_{\alpha\beta}^{-1}, \quad X_0 = (\beta + \alpha)U_{\alpha\beta}^{-1}CB_\alpha^{-1}. \tag{3.8b}$$

Premultiply the equations in (3.4) by $Z_3Z_2Z_1$ to get

$$M_0\begin{pmatrix} I \\ X \end{pmatrix} = L_0\begin{pmatrix} I \\ X \end{pmatrix}\mathscr{C}(R;\beta,\alpha), \quad M_0\begin{pmatrix} Y \\ I \end{pmatrix}\mathscr{C}(S;\alpha,\beta) = L_0\begin{pmatrix} Y \\ I \end{pmatrix}. \tag{3.9}$$

Our development up to this point differs from SDA of [22] only in our inclusion of two parameters α and β. The significance of doing so will be demonstrated in our later comparisons on convergence rates in section 5 and numerical examples in section 7. From this point forward, ours is the same as in [22]. The idea is to construct a sequence of pairs $\{M_k, L_k\}$, $k = 0, 1, 2, \ldots$, such that

$$M_k\begin{pmatrix} I \\ X \end{pmatrix} = L_k\begin{pmatrix} I \\ X \end{pmatrix}[\mathscr{C}(R;\beta,\alpha)]^{2^k}, \quad M_k\begin{pmatrix} Y \\ I \end{pmatrix}[\mathscr{C}(S;\alpha,\beta)]^{2^k} = L_k\begin{pmatrix} Y \\ I \end{pmatrix}, \tag{3.10}$$

and at the same time M_k and L_k have the same forms as M_0 and L_0, respectively, i.e.,

$$M_k = \begin{pmatrix} E_k & 0 \\ -X_k & I \end{pmatrix}, \quad L_k = \begin{pmatrix} I & -Y_k \\ 0 & F_k \end{pmatrix}. \tag{3.11}$$

The technique for constructing $\{M_{k+1}, L_{k+1}\}$ from $\{M_k, L_k\}$ is not entirely new and can be traced back to the 1980s in [10, 15, 26] and more recently in [3, 5, 31]. The idea is to seek suitable $\check{M}, \check{L} \in \mathbb{R}^{(m+n)\times(m+n)}$ such that

$$\operatorname{rank}\left((\check{M}, \check{L})\right) = m + n, \quad (\check{M}, \check{L})\begin{pmatrix} L_k \\ -M_k \end{pmatrix} = 0 \tag{3.12}$$

and set $M_{k+1} = \check{M}M_k$ and $L_{k+1} = \check{L}L_k$. It is not hard to verify that if the equations in (3.10) hold, then they hold for k replaced by $k+1$, i.e., for the newly constructed M_{k+1} and L_{k+1}. The only problem is that not every pair $\{\check{M}, \check{L}\}$ satisfying (3.12) leads to $\{M_{k+1}, L_{k+1}\}$ having the forms of (3.11). For this, we turn to the constructions of $\{\check{M}, \check{L}\}$ in [12, 13, 22, 25]:

$$\check{M} = \begin{pmatrix} E_k(I_m - Y_kX_k)^{-1} & 0 \\ -F_k(I_n - X_kY_k)^{-1}X_k & I_n \end{pmatrix}, \quad \check{L} = \begin{pmatrix} I_m & -E_k(I_m - Y_kX_k)^{-1}Y_k \\ 0 & -F_k(I_n - X_kY_k)^{-1} \end{pmatrix},$$

with which $M_{k+1} = \check{M}M_k$ and $L_{k+1} = \check{L}L_k$ have the forms of (3.11) with

$$E_{k+1} = E_k(I_m - Y_kX_k)^{-1}E_k, \tag{3.13a}$$

$$F_{k+1} = F_k(I_n - X_kY_k)^{-1}F_k, \tag{3.13b}$$

$$X_{k+1} = X_k + F_k(I_n - X_kY_k)^{-1}X_kE_k, \tag{3.13c}$$

$$Y_{k+1} = Y_k + E_k(I_m - Y_kX_k)^{-1}Y_kF_k. \tag{3.13d}$$

By now we have presented the framework of ADDA:

1. Pick suitable α and β for (best) convergence rate.
2. Compute M_0 and L_0 of (3.7) by (3.5), (3.6), and (3.8).
3. Iteratively compute M_k and L_k by (3.13) until convergence.

Associated with this general framework arise a few questions:

1. Are the iterative formulas in (3.13) well-defined, i.e., do all the inverses exist?
2. How do we choose best parameters α and β for fast convergence?
3. What do X_k and Y_k converge to if they are convergent?
4. How much better is ADDA than the doubling algorithms: SDA of Guo, Lin, and Xu [22] and SDA-ss of Bini, Meini, and Poloni [9]?

The first three questions will be addressed in the next subsection, while the last question will be answered in section 5.

3.2. Analysis. Recall that W defined by (1.2) is a nonsingular or an irreducible singular M-matrix. MARE (1.1) has a unique minimal nonnegative solution Φ [19] and cMARE (3.1) has a unique minimal nonnegative solution Ψ. Some properties of Φ and Ψ are summarized in Theorem 3.1. They are needed in order to answer the questions we posed at the end of the previous subsection.

THEOREM 3.1 (see [16, 17, 19]). *Assume* (1.3).

(a) *MARE* (1.1) *has a unique minimal nonnegative solution* Φ*, and its cMARE* (3.1) *has a unique minimal nonnegative solution* Ψ*.*

(b) *If* W *is* irreducible*, then* $\Phi > 0$ *and* $A - \Phi D$ *and* $B - D\Phi$ *are* irreducible M*-matrices.*

(c) *If* W *is* nonsingular*, then* $A - \Phi D$ *and* $B - D\Phi$ *are* nonsingular M*-matrices.*

(d) *Suppose* W *is* irreducible *and* singular*. Let* $u_1, v_1 \in \mathbb{R}^m$ *and* $u_2, v_2 \in \mathbb{R}^n$ *be positive vectors such that*

$$W \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0, \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}^{\mathrm{T}} W = 0. \tag{3.14}$$

1. *If* $u_1^{\mathrm{T}} v_1 > u_2^{\mathrm{T}} v_2$*, then* $B - D\Phi$ *is a singular* M*-matrix with*[2] $(B - D\Phi)v_1 = 0$ *and* $A - \Phi D$ *is a nonsingular* M*-matrix, and* $\Phi v_1 = v_2$ *and* $\Psi v_2 < v_1$*.*
2. *If* $u_1^{\mathrm{T}} v_1 = u_2^{\mathrm{T}} v_2$ *(the so-called critical case), then both* $B - D\Phi$ *and* $A - \Phi D$ *are singular* M*-matrices, and* $\Phi v_1 = v_2$ *and* $\Psi v_2 = v_1$*.*
3. *If* $u_1^{\mathrm{T}} v_1 < u_2^{\mathrm{T}} v_2$*, then* $B - D\Phi$ *is a nonsingular* M*-matrix and* $A - \Phi D$ *is a singular* M*-matrix, and* $\Phi v_1 < v_2$ *and* $\Psi v_2 = v_1$*.*

(e) $I - \Phi\Psi$ *and* $I - \Psi\Phi$ *are* M*-matrices and they are nonsingular, except for the critical case in which both are singular.*

Recall that our goal is to compute Φ as efficiently and accurately as possible and, as a by-product, Ψ, too. In view of this goal, we identify $X = \Phi$ and $Y = \Psi$ in all appearances of X and Y in subsection 3.1. In particular,

$$S = A - C\Psi, \quad R = B - D\Phi, \tag{3.3$'$}$$

and (3.10) and (3.11) yield immediately

$$E_k = (I - Y_k\Phi)\ [\mathscr{C}(R;\beta,\alpha)]^{2^k}, \tag{3.15a}$$

$$\Phi - X_k = \quad F_k\Phi \quad [\mathscr{C}(R;\beta,\alpha)]^{2^k}, \tag{3.15b}$$

$$\Psi - Y_k = \quad E_k\Psi \quad [\mathscr{C}(S;\alpha,\beta)]^{2^k}, \tag{3.15c}$$

$$F_k = (I - X_k\Psi)\,[\mathscr{C}(S;\alpha,\beta)]^{2^k}. \tag{3.15d}$$

Examining (3.15), we see that ADDA will converge if X_k and Y_k are uniformly bounded with respect to k, and if

$$\rho(\mathscr{C}(R;\beta,\alpha)) < 1, \quad \rho(\mathscr{C}(S;\alpha,\beta)) < 1, \tag{3.16a}$$

because then E_k and F_k are uniformly bounded with respect to k, and

$$[\mathscr{C}(R;\beta,\alpha)]^{2^k} \to 0, \quad [\mathscr{C}(S;\alpha,\beta)]^{2^k} \to 0 \tag{3.16b}$$

[2] [16, Theorem 4.8] says in this case $D\Phi v_1 = Dv_2$, which leads to $(B - D\Phi)v_1 = Bv_1 - Dv_2 = 0$.

as $k \to \infty$, implying that $\Phi - X_k \to 0$ and $\Psi - Y_k \to 0$ as $k \to \infty$. This is one of the guiding principles in [22] which enforces

$$\alpha = \beta \geq \max_{i,j}\{A_{(i,i)}, B_{(j,j)}\}, \tag{3.17}$$

which in turn ensures that X_k and Y_k are uniformly bounded and also ensures (3.16a) and thus (3.16b) because, by Theorem 3.1(c), both[3] S and R are nonsingular M-matrices if[4] W is a nonsingular M-matrix. Later, Guo, Iannazzo, and Meini [20] observed that the SDA of [22] still converges even if W is a singular irreducible M-matrix. This observation was formally proved in [11]. Guo, Iannazzo, and Meini [20, Theorem 4.4] also proved that taking

$$\alpha = \beta = \max_{i,j}\{A_{(i,i)}, B_{(j,j)}\} \tag{3.18}$$

makes the resulting SDA converge the fastest subject to (3.17). Another critical implication of (3.17) is that it makes $-E_0$ and $-F_0$, E_k and F_k for $k \geq 1$, and X_k and Y_k for $k \geq 0$ all nonnegative [22], a property that enables the SDA of [22] (with some minor but crucial implementation changes [36]) to compute Φ with deserved entrywise relative accuracy as argued in [36].

We would like our ADDA to have such a capability as well, i.e., computing Φ with deserved entrywise relative accuracy. To this end, we require

$$\alpha \geq \alpha_{\rm opt} \stackrel{\rm def}{=} \max_i A_{(i,i)}, \quad \beta \geq \beta_{\rm opt} \stackrel{\rm def}{=} \max_j B_{(j,j)} \tag{3.19}$$

but allow α and β to be different, and we seek to minimize the product of the spectral radii

$$\rho(\mathscr{C}(R;\beta,\alpha)) \cdot \rho(\mathscr{C}(S;\alpha,\beta)),$$

rather than each individual spectral radius. Later in Theorem 3.4, we will see that it is this product, not each individual spectral radius, that ultimately reflects the true rate of convergence. In particular, convergence is guaranteed if the product is less than 1, even if one of the spectral radii is bigger than 1. Moreover, the smaller the product, the faster the convergence.

That the rate of convergence of a doubling algorithm on a matrix Riccati type equation is dependent on the product of some two spectral radii is not new. In fact, the convergence analyses in [20, 22, 25] all suggested that.

The assumption (1.3) implies that A and B are nonsingular M-matrices by Lemma 2.2(e). Therefore both $\alpha_{\rm opt} > 0$ and $\beta_{\rm opt} > 0$.

LEMMA 3.2. *Assume* (1.3). *If* $\alpha > 0$ *and* $\beta > 0$, *then* A_β, B_α, $U_{\alpha\beta}$, *and* $V_{\alpha\beta}$ *defined in* (3.5) *and* (3.6) *are nonsingular* M*-matrices. Furthermore, both* $U_{\alpha\beta}$ *and* $V_{\alpha\beta}$ *are irreducible if* W *is irreducible.*

Proof. If $\alpha > 0$ and $\beta > 0$,

$$\widehat{W} = W + \begin{pmatrix} \alpha I & 0 \\ 0 & \beta I \end{pmatrix} = \begin{pmatrix} B + \alpha I & -D \\ -C & A + \beta I \end{pmatrix} \geq \min\{\alpha, \beta\} \cdot I + W$$

[3]That R is a nonsingular M-matrix is stated explicitly in Theorem 3.1(c). For S, we apply Theorem 3.1(c) to cMARE (3.1) identified as a MARE in the form of (1.1) with its coefficient matrix as $\left(\begin{smallmatrix} A & -C \\ -D & B \end{smallmatrix}\right)$.

[4]This is the case studied in [22].

is a nonsingular M-matrix. As the diagonal blocks of $\widehat{W}$, A_β, and B_α are nonsingular M-matrices, so are their corresponding Schur complements $V_{\alpha\beta}$ and $U_{\alpha\beta}$ in $\widehat{W}$ by Lemma 2.2(e). If also W is irreducible, then $\widehat{W}$ is a nonsingular irreducible M-matrix, and thus both $U_{\alpha\beta}$ and $V_{\alpha\beta}$ are nonsingular irreducible M-matrices again by Lemma 2.2(e). □

THEOREM 3.3. *Assume* (1.3) *and* (3.19).

(a) *We have*

$$E_0 \le 0,\ F_0 \le 0,\ \mathscr{C}(R;\beta,\alpha) \le 0,\ \mathscr{C}(S;\alpha,\beta) \le 0, \tag{3.20}$$

$$0 \le X_0 \le \Phi,\ 0 \le Y_0 \le \Psi. \tag{3.21}$$

If W is also irreducible, then

$$E_0 < 0,\ F_0 < 0,\ \mathscr{C}(R;\beta,\alpha) < 0,\ \mathscr{C}(S;\alpha,\beta) < 0, \tag{3.20$'$}$$

$$0 \le X_0 < \Phi,\ 0 \le Y_0 < \Psi. \tag{3.21$'$}$$

(b) *Both $I - Y_kX_k$ and $I - X_kY_k$ are nonsingular M-matrices for all $k \ge 0$.*

(c) *We have*

(3.22)

$$E_k \ge 0,\ F_k \ge 0,\ 0 \le X_{k-1} \le X_k \le \Phi,\ 0 \le Y_{k-1} \le Y_k \le \Psi \quad \textit{for } k \ge 1.$$

If W is also irreducible, then

(3.22$'$)

$$E_k > 0,\ F_k > 0,\ 0 \le X_{k-1} < X_k < \Phi,\ 0 \le Y_{k-1} < Y_k < \Psi \quad \textit{for } k \ge 1.$$

Proof. Our proof is largely the same as the proofs in [20, p. 1088].

(a) That $\mathscr{C}(R;\beta,\alpha) \le 0$ and $\mathscr{C}(S;\alpha,\beta) \le 0$ is fairly straightforward because R and S are M-matrices and α and β are restricted by (3.19). For E_0 and F_0, we note

$$E_0 = V_{\alpha\beta}^{-1}[V_{\alpha\beta} - (\beta+\alpha)I] \tag{3.23a}$$

$$= V_{\alpha\beta}^{-1}(B - \beta I - DA_\beta^{-1}C), \tag{3.23b}$$

$$F_0 = U_{\alpha\beta}^{-1}[U_{\alpha\beta} - (\beta+\alpha)I] \tag{3.23c}$$

$$= U_{\alpha\beta}^{-1}(A - \alpha I - CB_\alpha^{-1}D). \tag{3.23d}$$

Since A_β, B_α, $V_{\alpha\beta}$, and $U_{\alpha\beta}$ are nonsingular M-matrices by Lemma 3.2, we have

$$A_\beta^{-1} \ge 0, \quad B_\alpha^{-1} \ge 0, \quad V_{\alpha\beta}^{-1} \ge 0, \quad U_{\alpha\beta}^{-1} \ge 0.$$

Therefore $E_0 \le 0$, $F_0 \le 0$, $X_0 \ge 0$, and $Y_0 \ge 0$. Equations (3.15b) and (3.15c) for $k = 0$ yield $\Phi - X_0 \ge 0$ and $\Psi - Y_0 \ge 0$, respectively.

Now suppose W is irreducible. By Lemma 3.2, both $U_{\alpha\beta}$ and $V_{\alpha\beta}$ are irreducible. So $U_{\alpha\beta}^{-1} > 0$, $V_{\alpha\beta}^{-1} > 0$, and no columns of $V_{\alpha\beta} - (\beta+\alpha)I$ and $U_{\alpha\beta} - (\beta+\alpha)I$ both of which are nonpositive, are zeros. Therefore $E_0 < 0$ and $F_0 < 0$ by (3.23a) and (3.23c). Theorem 3.1(b) implies that $(S+\beta I)^{-1} > 0$, $(R+\alpha I)^{-1} > 0$, and no columns of $S - \alpha I$ and $R - \beta I$, both of which are nonpositive are zeros, and thus

$$\mathscr{C}(S;\alpha,\beta) = (S+\beta I)^{-1}(S-\alpha I) < 0,\ \mathscr{C}(R;\beta,\alpha) = (R+\alpha I)^{-1}(R-\beta I) < 0.$$

Finally

$$\Phi - X_0 = F_0\Phi\mathscr{C}(R;\beta,\alpha) > 0, \quad \Psi - Y_0 = E_0\Psi\mathscr{C}(S;\alpha,\beta) > 0$$

because $\Phi > 0$ and $\Psi > 0$ by Theorem 3.1(b) and (3.20$'$).

(b)–(c) We have $I - X_0Y_0 \geq I - \Phi\Psi$ and $I - Y_0X_0 \geq I - \Psi\Phi$. Suppose for the moment that W is nonsingular. Then both $I - \Phi\Psi$ and $I - \Psi\Phi$ are nonsingular M-matrices by Theorem 3.1(e), and thus $I - X_0Y_0$ and $I - Y_0X_0$ are nonsingular M-matrices, too, by Lemma 2.2(b).

Now suppose W is an irreducible singular matrix. By Theorem 3.1(d), we have $\Psi\Phi v_1 \leq v_1$, where $v_1 > 0$ is defined in Theorem 3.1(d). So $\rho(\Psi\Phi) \leq 1$ by [7, Theorem 1.11, p. 28]. By part (a) of this theorem, $0 \leq X_0 < \Phi$ and $0 \leq Y_0 < \Psi$. Therefore $0 \leq X_0Y_0 < \Phi\Psi$. Since $\Phi\Psi$ is irreducible, we conclude by [7, Corollary 1.5, p. 27]

$$\rho(Y_0X_0) = \rho(X_0Y_0) < \rho(\Psi\Phi) = \rho(\Phi\Psi) \leq 1,$$

and thus $I - Y_0X_0$ and $I - X_0Y_0$ are nonsingular M-matrices. This proves part (b) for $k = 0$.

Since $E_0 \leq 0$ and $F_0 \leq 0$, and $I - Y_0X_0$ and $I - X_0Y_0$ are nonsingular M-matrices, we deduce from (3.13) that

$$E_1 \geq 0, \quad F_1 \geq 0, \quad X_1 \geq X_0, \quad Y_1 \geq Y_0.$$

By (3.15b) and (3.15c),

$$\Phi - X_1 = F_1\Phi\left[\mathscr{C}(R;\beta,\alpha)\right]^2, \quad \Psi - Y_1 = E_1\Psi\left[\mathscr{C}(S;\alpha,\beta)\right]^2, \tag{3.24}$$

yielding $\Phi - X_1 \geq 0$ and $\Psi - Y_1 \geq 0$, respectively. Consider now W is also irreducible. We have, by (3.20′) and (3.21′) and (3.13),

$$E_1 > 0, \quad F_1 > 0, \quad X_1 > X_0 \geq 0, \quad Y_1 > Y_0 \geq 0,$$

and then $X_1 < \Phi$ and $Y_1 < \Psi$ by (3.24). This proves part (c) for $k = 1$.

Part (b) for $k \geq 1$ and part (c) for $k \geq 2$ can be proved together through the induction argument. Details are omitted. □

One important implication of Theorem 3.3 is that all formulas in subsection 3.1 for ADDA are well-defined under the assumptions (1.3) and (3.19).

Next we look into choosing α and β subject to (3.19) to optimize the convergence speed. We have (3.15), which yields

$$0 \leq \Phi - X_k = (I - X_k\Psi)\left[\mathscr{C}(S;\alpha,\beta)\right]^{2^k}\Phi\left[\mathscr{C}(R;\beta,\alpha)\right]^{2^k} \tag{3.25a}$$

$$\leq \left[\mathscr{C}(S;\alpha,\beta)\right]^{2^k}\Phi\left[\mathscr{C}(R;\beta,\alpha)\right]^{2^k}, \tag{3.25b}$$

$$0 \leq \Psi - Y_k = (I - Y_k\Phi)\left[\mathscr{C}(R;\beta,\alpha)\right]^{2^k}\Psi\left[\mathscr{C}(S;\alpha,\beta)\right]^{2^k} \tag{3.25c}$$

$$\leq \left[\mathscr{C}(R;\beta,\alpha)\right]^{2^k}\Psi\left[\mathscr{C}(S;\alpha,\beta)\right]^{2^k}. \tag{3.25d}$$

Now if W is a nonsingular M-matrix, then both R and S are nonsingular M-matrices, too, by Theorem 3.1(c). Therefore

$$\rho(\mathscr{C}(R;\beta,\alpha)) < 1, \ \rho(\mathscr{C}(S;\alpha,\beta)) < 1 \text{ under (3.17)}, \tag{3.26}$$

implying $X_k \to \Phi$ and $Y_k \to \Psi$ as $k \to \infty$. This is what was proved in [22]. But for an irreducible singular M-matrix W with $u_1^{\mathrm{T}}v_1 \neq u_2^{\mathrm{T}}v_2$, it is proved in [20] that one of the spectral radii in (3.26) is less than 1, while the other is equal to 1, still implying $X_k \to \Phi$ and $Y_k \to \Psi$ as $k \to \infty$. Furthermore, [20, Theorem 4.4] implies that the best choice is given by (3.18) in the sense that both spectral radii in $\rho(\mathscr{C}(R;\alpha,\alpha))$ and $\rho(\mathscr{C}(S;\alpha,\alpha))$ are minimized subject to $\alpha \geq \max_{i,j}\{A_{(i,i)}, B_{(j,j)}\}$.

We can do better by allowing α and β to be different with the help of Theorem 2.3. The main result is summarized in the following theorem.

THEOREM 3.4. *Assume* (1.3) *and* (3.19). *We have*

$$\limsup_{k\to\infty} \|\Phi - X_k\|^{1/2^k} \le \rho(\mathscr{C}(S;\alpha,\beta)) \cdot \rho(\mathscr{C}(R;\beta,\alpha)), \tag{3.27a}$$

$$\limsup_{k\to\infty} \|\Psi - Y_k\|^{1/2^k} \le \rho(\mathscr{C}(R;\beta,\alpha)) \cdot \rho(\mathscr{C}(S;\alpha,\beta)). \tag{3.27b}$$

The optimal α and β that minimize the right-hand sides of (3.27) *are $\alpha = \alpha_{\rm opt}$ and $\beta = \beta_{\rm opt}$.*

Proof. Since all matrix norms are equivalent, we may assume that $\|\cdot\|$ is consistent. By (3.25b), we have

$$\|\Phi - X_k\|^{1/2^k} \le \left\|[\mathscr{C}(S;\alpha,\beta)]^{2^k}\right\|^{1/2^k} \cdot \|\Phi\|^{1/2^k} \cdot \left\|[\mathscr{C}(R;\beta,\alpha)]^{2^k}\right\|^{1/2^k},$$

which goes to $\rho(\mathscr{C}(S;\alpha,\beta))\cdot\rho(\mathscr{C}(R;\beta,\alpha))$ as $k\to\infty$, unless $\Phi = 0$, in which case both sides are 0 for all k. Thus (3.27a) holds. Similarly we have (3.27b). Since $R = B - D\Phi$ and $S = A - C\Psi$ are M-matrices and $D\Phi \ge 0$ and $C\Psi \ge 0$,

$$\alpha \ge \max_i A_{(i,i)} \ge \max_i S_{(i,i)}, \quad \beta \ge \max_j B_{(j,j)} \ge \max_j R_{(j,j)}.$$

By Theorem 2.3, $\rho(\mathscr{C}(R;\beta,\alpha)) \cdot \rho(\mathscr{C}(S;\alpha,\beta))$ is either strictly increasing if at least one of R and S is nonsingular or identically 1, subject to (3.19). In any case, $\alpha = \alpha_{\rm opt}$ and $\beta = \beta_{\rm opt}$ minimize the product $\rho(\mathscr{C}(S;\alpha,\beta)) \cdot \rho(\mathscr{C}(R;\beta,\alpha))$. ∎

3.3. Optimal ADDA. We are now ready to present our ADDA, based on the framework in subsection 3.1 and analysis in subsection 3.2.

ALGORITHM 3.1. ADDA for MARE $XDX - AX - XB + C = 0$ and, as a by-product, for cMARE $YCY - YA - BY + D = 0$.

1 Pick $\alpha \ge \alpha_{\rm opt}$ and $\beta \ge \beta_{\rm opt}$;
2 $A_\beta \stackrel{\rm def}{=} A + \beta I$, $B_\alpha \stackrel{\rm def}{=} B + \alpha I$;
3 Compute A_β^{-1} and B_α^{-1};
4 Compute $V_{\alpha\beta}$ and $U_{\alpha\beta}$ as in (3.6) and then their inverses;
5 Compute E_0 by (3.23b), F_0 by (3.23d), X_0 and Y_0 by (3.8);
6 Compute $(I - X_0Y_0)^{-1}$ and $(I - Y_0X_0)^{-1}$;
7 Compute X_1 and Y_1 by (3.13c) and (3.13d);
8 For $k = 1, 2, \ldots$, until convergence
9 Compute E_k and F_k by (3.13a) and (3.13b)
(after substituting $k+1$ for k);
10 Compute $(I - X_kY_k)^{-1}$ and $(I - Y_kX_k)^{-1}$;
11 Compute X_{k+1} and Y_{k+1} by (3.13c) and (3.13d);
12 Enddo

Remark 3.1. ADDA differs from SDA of [22] only in its initial setup—lines 1–5 that build two parameters α and β into the algorithm. In [36], we explained in detail how to make critical implementation changes to ensure computed Φ and Ψ by SDA to have entrywise relative accuracy as much as the input data deserves. The key is to use the GTH-like algorithm [1, 35] to invert all nonsingular M-matrices. Every comment in [36, Remark 4.1], except the selection of its sole parameter for SDA, applies here. We shall not repeat most of those comments to save space.

About selecting the parameters α and β, Theorem 3.4 suggests $\alpha = \alpha_{\rm opt}$ and $\beta = \beta_{\rm opt}$ for the best convergence rate. But when the diagonal entries of A and B are not known exactly or are not exactly floating point numbers, the diagonal entries of $A - \alpha I$ and $B - \beta I$ needed for computing E_0 by (3.23b) and F_0 by (3.23d) may suffer catastrophic cancelations. One remedy to avoid this is to take $\alpha = \eta \cdot \alpha_{\rm opt}$ and $\beta = \eta \cdot \beta_{\rm opt}$ for some $\eta > 1$ but not too close to 1. This will slow the convergence, but the gain is to ensure computed Φ and Ψ by ADDA have deserved entrywise relative accuracy. Usually ADDA converges so fast that such a little degradation in the optimality of α and β does not increase the number of iteration steps needed for convergence.

Recall the convergence of ADDA does not depend on both spectral radii $\rho(\mathscr{C}(S;\alpha,\beta))$ and $\rho(\mathscr{C}(R;\beta,\alpha))$ being less than 1. In fact, often the larger is bigger than 1 while the smaller is less than 1 but the product is less than 1. It can happen that the larger one is so big that implemented as exactly given in Algorithm 3.1, ADDA can encounter overflown in E_k or F_k before X_k and Y_k converge with a desired accuracy. This happened in one of our test runs. To cure this, we notice that scaling E_k and F_k to ηE_k and $\eta^{-1}F_k$ for some $\eta > 0$ has no effect on X_{k+1} and Y_{k+1} and thereafter. In view of this, we devise the following strategy: at every iteration step after E_k and F_k are computed, we pick η such that $\|\eta E_k\| = \|\eta^{-1}F_k\|$, i.e., $\eta = \sqrt{\|F_k\|/\|E_k\|}$, and scale E_k and F_k to ηE_k and $\eta^{-1}F_k$. Which matrix norm $\|\cdot\|$ is not particularly important and in our tests—we used the ℓ_1-operator norm $\|\cdot\|_1$.

The *optimal ADDA* is the one with $\alpha = \alpha_{\rm opt}$ and $\beta = \beta_{\rm opt}$. Since there is little reason not to use the optimal ADDA, except for the situation we mentioned in Remark 3.1, for ease of presentation in what follows we always mean the optimal ADDA whenever we refer to an ADDA, unless explicitly stated differently.

4. Application to *M*-matrix Sylvester equation. When $D = 0$, MARE (1.1) degenerates to a Sylvester equation:

$$AX + XB = C. \tag{4.1}$$

The assumption (1.3) on its associated $\left(\begin{smallmatrix} B & 0 \\ -C & A \end{smallmatrix}\right)$ implies that A and B are nonsingular M-matrices and $C \geq 0$. Thus (4.1) is an *M-matrix sylvester equation* (MSE) as defined in [35]: both A and B have positive diagonal entries and nonpositive off-diagonal entries and $P = I_m \otimes A + B^{\rm T} \otimes I_n$ is a nonsingular M-matrix, and $C \geq 0$.

MSE (4.1) has the unique solution $\Phi \geq 0$ and its cMARE has the solution $\Psi = 0$. Apply ADDA to (4.1) to get

$$E_0 = \mathscr{C}(B;\beta,\alpha) \equiv (B+\alpha I)^{-1}(B-\beta I), \tag{4.2a}$$

$$F_0 = \mathscr{C}(A;\alpha,\beta) \equiv (A+\beta I)^{-1}(A-\alpha I), \tag{4.2b}$$

$$X_0 = (\beta+\alpha)(A+\beta I)^{-1}C(B+\alpha I)^{-1}, \tag{4.2c}$$

and for $k \geq 0$

$$E_{k+1} = E_k^2, \qquad F_{k+1} = F_k^2, \tag{4.2d}$$

$$X_{k+1} = X_k + F_k X_k E_k. \tag{4.2e}$$

The associated error equation is

$$0 \leq \Phi - X_k = [\mathscr{C}(A;\alpha,\beta)]^{2^k}\, \Phi\, [\mathscr{C}(B;\beta,\alpha)]^{2^k}. \tag{4.3}$$

Smith's method [30, 35] is obtained after setting $\alpha = \beta$ in (4.2) always.

Alternatively, we can derive (4.2) through a combination of an ADI iteration and Smith's idea in [30]. Given an approximation $\boldsymbol{X} \approx \Phi$, we compute the next approximation $\boldsymbol{Z}$ by one step of ADI:

1. Solve $(A+\beta I)\boldsymbol{Y} = C - \boldsymbol{X}(B-\beta I)$ for $\boldsymbol{Y}$.
2. Solve $\boldsymbol{Z}(B+\alpha I) = C - (A-\alpha I)\boldsymbol{Y}$ for $\boldsymbol{Z}$.

Eliminate $\boldsymbol{Y}$ to get

$$\boldsymbol{Z} = X_0 + F_0 \boldsymbol{X} E_0, \tag{4.4}$$

where E_0, F_0, and X_0 are the same as in (4.2a)–(4.2c). With $\boldsymbol{X} = 0$, keep iterating (4.4) to get

$$\boldsymbol{Z}_k = \sum_{i=0}^{k} F_0^i X_0 E_0^i.$$

If it converges, it converges to the solution $\Phi = \boldsymbol{Z}_\infty = \sum_{i=0}^{\infty} F_0^i X_0 E_0^i$. It can be verified that $\{\boldsymbol{Z}_i\}$ relates to $\{X_i\}$ by $X_k = \boldsymbol{Z}_{2^k}$. Namely, instead of computing every member in the sequence $\{\boldsymbol{Z}_i\}$, (4.2) computes only the 2^kth members. In view of its connection to ADI and Smith's method [30], we call (4.2) an *alternating-directional Smith method* (ADSM) for MSE (4.1). This connection to ADI is also the reason we name our Algorithm 3.1 an *alternating-directional doubling algorithm*.

Equation (4.3) gives

$$\limsup_{k\to\infty} \|\Phi - X_k\|^{1/2^k} \le \rho(\mathscr{C}(A;\alpha,\beta)) \cdot \rho(\mathscr{C}(B;\beta,\alpha)), \tag{4.5}$$

suggesting we pick α and β to minimize the right-hand side of (4.5) for fastest convergence. Subject to again

$$\alpha \ge \alpha_{\text{opt}} \stackrel{\text{def}}{=} \max_i A_{(i,i)}, \quad \beta \ge \beta_{\text{opt}} \stackrel{\text{def}}{=} \max_j B_{(j,j)} \tag{3.19}$$

in order to ensure $F_0 \le 0$, $E_0 \le 0$ and all $F_k \ge 0$ and $E_k \ge 0$ for $k \ge 1$, we conclude by Theorem 2.3 that $\alpha = \alpha_{\text{opt}}$ and $\beta = \beta_{\text{opt}}$ minimize the right-hand side of (4.5).

5. Comparisons with existing doubling algorithms. In this section, we will compare the rates of convergence among our ADDA, the SDA of [22], and SDA combined with the SDA-ss of [9].

The right-hand sides in (3.27) provide an upper bound on convergence rate of ADDA. It is possible that the bound may overestimate the rate, but we expect in general it is tight. To facilitate our comparisons in what follows, we shall simply regard the upper bound as the *true* rate and without loss of generality assume

$$\alpha_{\text{opt}} \stackrel{\text{def}}{=} \max_i A_{(i,i)} \ge \beta_{\text{opt}} \stackrel{\text{def}}{=} \max_i B_{(i,i)}. \tag{5.1}$$

Let $\lambda_{\min}(S)$ be the eigenvalue of S in (3.3′) with the smallest real part among all its eigenvalues. We know $\lambda_{\min}(S) \ge 0$ and let $\lambda_{\min}(R)$ be the same for R as in (3.3′).

We have the convergence rate for the optimal ADDA

$$r_{\text{adda}} = \frac{\alpha_{\text{opt}} - \lambda_{\min}(S)}{\beta_{\text{opt}} + \lambda_{\min}(S)} \cdot \frac{\beta_{\text{opt}} - \lambda_{\min}(R)}{\alpha_{\text{opt}} + \lambda_{\min}(R)}. \tag{5.2}$$

Estimates in (3.27) with $\alpha = \beta$ hold for SDA. Apply [20, Theorem 4.4] to conclude that the convergence rate for the optimal SDA is

$$r_{\rm sda} = \frac{\alpha_{\rm opt} - \lambda_{\min}(S)}{\alpha_{\rm opt} + \lambda_{\min}(S)} \cdot \frac{\alpha_{\rm opt} - \lambda_{\min}(R)}{\alpha_{\rm opt} + \lambda_{\min}(R)} \tag{5.3}$$

upon noticing (5.1).

In order to see the convergence rate of the optimal SDA-ss, we outline the algorithm below. For

$$\beta \ge \beta_{\rm opt} \stackrel{\rm def}{=} \max_j B_{(j,j)}, \tag{5.4}$$

set

$$\widehat{H} = I - \beta^{-1}H, \quad \widehat{A} = I + \beta^{-1}A, \quad \widehat{B} = I - \beta^{-1}B, \tag{5.5}$$

where H is defined as in (3.3). With S and R given by (3.3′), we have

$$\widehat{H}\begin{pmatrix} I \\ \Phi \end{pmatrix} = \begin{pmatrix} I \\ \Phi \end{pmatrix}\widehat{R}, \qquad \widehat{H}\begin{pmatrix} \Psi \\ I \end{pmatrix}\widehat{S} = \begin{pmatrix} \Psi \\ I \end{pmatrix}, \tag{5.6a}$$

$$\widehat{R} = I - \beta^{-1}R, \qquad \widehat{S} = (I + \beta^{-1}S)^{-1}. \tag{5.6b}$$

Note that $\widehat{A}$ is a nonsingular M-matrix, and let

$$\widehat{M}_0 = \begin{pmatrix} \widehat{E}_0 & 0 \\ -\widehat{X}_0 & I \end{pmatrix}, \quad \widehat{L}_0 = \begin{pmatrix} I & -\widehat{Y}_0 \\ 0 & \widehat{F}_0 \end{pmatrix}, \tag{5.7}$$

where

$$\widehat{E}_0 = \widehat{B} + \beta^{-2}D\widehat{A}^{-1}C, \quad \widehat{Y}_0 = \beta^{-1}D\widehat{A}^{-1}, \tag{5.8a}$$

$$\widehat{F}_0 = \widehat{A}^{-1}, \qquad \widehat{X}_0 = \beta^{-1}\widehat{A}^{-1}C. \tag{5.8b}$$

It can be verified that $\widehat{H} = \widehat{L}_0^{-1}\widehat{M}_0$, substituting which into the equations in (5.6) to get

$$\widehat{M}_0\begin{pmatrix} I \\ \Phi \end{pmatrix} = \widehat{L}_0\begin{pmatrix} I \\ \Phi \end{pmatrix}\widehat{R}, \quad \widehat{M}_0\begin{pmatrix} \Psi \\ I \end{pmatrix}\widehat{S} = \widehat{L}_0\begin{pmatrix} \Psi \\ I \end{pmatrix}.$$

The rest follows the idea in [22] (and also that in section 3). SDA-ss seeks to construct a sequence of pairs $\{\widehat{M}_k, \widehat{L}_k\}$, $k = 0, 1, 2, \ldots$, such that

$$\widehat{M}_k\begin{pmatrix} I \\ \Phi \end{pmatrix} = \widehat{L}_k\begin{pmatrix} I \\ \Phi \end{pmatrix}\widehat{R}^{2^k}, \quad \widehat{M}_k\begin{pmatrix} \Psi \\ I \end{pmatrix}\widehat{S}^{2^k} = \widehat{L}_k\begin{pmatrix} \Psi \\ I \end{pmatrix}, \tag{5.9}$$

and at the same time $\widehat{M}_k$ and $\widehat{L}_k$ have the same forms as $\widehat{M}_0$ and $\widehat{L}_0$, respectively, i.e.,

$$\widehat{M}_k = \begin{pmatrix} \widehat{E}_k & 0 \\ -\widehat{X}_k & I \end{pmatrix}, \quad \widehat{L}_k = \begin{pmatrix} I & -\widehat{Y}_k \\ 0 & \widehat{F}_k \end{pmatrix}. \tag{5.10}$$

The formulas (3.13) for advancing from the kth approximations to the $(k+1)$st ones remain valid here after placing a hat over every occurrence of E, F, X, and Y there.

At the end, we will have the following equations for errors in the approximations $\widehat{X}_k$ and $\widehat{Y}_k$:

$$\Phi - \widehat{X}_k = (I - \widehat{X}_k\Psi)\widehat{S}^{2^k}\Phi\widehat{R}^{2^k} \le \widehat{S}^{2^k}\Phi\widehat{R}^{2^k}, \tag{5.11}$$

$$\Psi - \widehat{Y}_k = (I - \widehat{Y}_k\Phi)\widehat{R}^{2^k}\Psi\widehat{S}^{2^k} \le \widehat{R}^{2^k}\Psi\widehat{S}^{2^k}. \tag{5.12}$$

Consequently

$$\limsup_{k\to\infty} \|\Phi - \widehat{X}_k\|^{1/2^k},\ \limsup_{k\to\infty} \|\Psi - \widehat{Y}_k\|^{1/2^k} \le \rho(\widehat{R})\cdot\rho(\widehat{S}). \tag{5.13}$$

In view of this inequality, (5.4), and Theorem 2.3, we conclude that the convergence rate of the optimal SDA-ss is

$$r_{\text{sda-ss}} = \frac{1-\beta_{\text{opt}}^{-1}\lambda_{\min}(R)}{1+\beta_{\text{opt}}^{-1}\lambda_{\min}(S)} = \frac{\beta_{\text{opt}}-\lambda_{\min}(R)}{\beta_{\text{opt}}+\lambda_{\min}(S)}. \tag{5.14}$$

Now we are ready to compare all three rates of convergence. To simplify notation, we drop the subscript opt to α and β and write $\lambda_S = \lambda_{\min}(S)$ and $\lambda_R = \lambda_{\min}(R)$. We have

$$\begin{aligned}\frac{r_{\text{adda}}}{r_{\text{sda}}} &= \frac{\beta-\lambda_R}{\alpha-\lambda_R}\cdot\frac{\alpha+\lambda_S}{\beta+\lambda_S}\\ &= 1 - \frac{(\lambda_R+\lambda_S)(\alpha-\beta)}{(\alpha-\lambda_R)(\beta+\lambda_S)},\end{aligned} \tag{5.15}$$

$$\begin{aligned}\frac{r_{\text{adda}}}{r_{\text{sda-ss}}} &= \frac{\alpha-\lambda_S}{\alpha+\lambda_R}\\ &= 1 - \frac{\lambda_R+\lambda_S}{\alpha+\lambda_R},\end{aligned} \tag{5.16}$$

$$\begin{aligned}\frac{r_{\text{sda-ss}}}{r_{\text{sda}}} &= \frac{\beta-\lambda_R}{\beta+\lambda_S}\cdot\frac{\alpha+\lambda_S}{\alpha-\lambda_S}\cdot\frac{\alpha+\lambda_R}{\alpha-\lambda_R}\\ &= 1 - \frac{(\lambda_R+\lambda_S)[\alpha(\alpha-\beta)-\lambda_S(\alpha-\lambda_R)-\alpha(\beta-\lambda_R)]}{(\beta+\lambda_S)(\alpha-\lambda_S)(\alpha-\lambda_R)}.\end{aligned} \tag{5.17}$$

If $\lambda_R+\lambda_S = 0$ (which happens in the critical case), then all three ratios are 1. In fact, for the critical case $r_{\text{adda}} = r_{\text{sda}} = r_{\text{sda-ss}} = 1$ and thus the three doubling algorithms converge linearly [11]. Suppose, in what follows, that $\lambda_R+\lambda_S > 0$, and recall (5.1). The first ratio

$$r_{\text{adda}}/r_{\text{sda}} \le 1 \quad \text{always},$$

with equality for $\alpha = \beta$, as expected. The ratio can be made much less than 1 if $\alpha/\beta \gg 1$. The second ratio

$$r_{\text{adda}}/r_{\text{sda-ss}} < 1 \quad \text{always}.$$

There is no definitive word on the third ratio because the sign of

$$\zeta \stackrel{\text{def}}{=} \alpha(\alpha-\beta) - \lambda_S(\alpha-\lambda_R) - \alpha(\beta-\lambda_R)$$

can change, dependent on different cases. If $\zeta > 0$, then SDA-ss is faster than SDA; otherwise it is slower.

It is worth pointing out that for SDA-ss it is very important how the shift-and-shrink (5.5) is done. For example, instead of (5.1), if

$$\max_i A_{(i,i)} < \max_i B_{(i,i)}. \tag{5.18}$$

Then we still have (5.14), but, instead of (5.3),

$$r_{\rm sda} = \frac{\beta-\lambda_S}{\beta+\lambda_S}\cdot\frac{\beta-\lambda_R}{\beta+\lambda_R}. \tag{5.19}$$

Then

$$\frac{r_{\rm sda}}{r_{\text{sda-ss}}} = \frac{\beta-\lambda_S}{\beta+\lambda_R} = 1-\frac{\lambda_R+\lambda_S}{\beta+\lambda_R} < 1$$

always, indicating SDA-ss is slower than SDA. To overcome this, when (5.18) holds, SDA-ss should be applied to cMARE (3.1) instead, and as a by-product, Φ is computed as the minimal nonnegative solution to the complementary MARE of cMARE (3.1).

6. Doubling algorithms by general bilinear transformations. The doubling algorithms SDA, SDA-ss, and ADDA are constructed, respectively, by

Cayley transformation:	$t \to \mathscr{C}(t;\alpha,\alpha) = (t-\alpha)/(t+\alpha)$	for SDA,
shrink-and-shift transformation:	$t \to t/\beta - 1$	for SDA-ss,
generalized Cayley transformation:	$t \to \mathscr{C}(t;\alpha,\beta) = (t-\alpha)/(t+\beta)$	for ADDA.

These transformations are three special cases of the following more general *bilinear* (also called *Möbius*) transformation:

$$t \to \mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) \stackrel{\rm def}{=} (\alpha_1 t-\alpha)/(\beta_1 t+\beta). \tag{6.1}$$

It is tempting to ask if some faster doubling algorithm than ADDA could be constructed with this bilinear transformation because of two additional parameters α_1 and β_1 to work with. In what follows we shall explain that optimal ADDA is still the best among all possible doubling algorithms coming out of (6.1).

The framework in subsection 3.1 can be modified to accommodate $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1)$ upon noticing that, similar to (3.4),

$$(\beta_1 H-\beta I)\begin{pmatrix} I\\ X\end{pmatrix} = (\alpha_1 H+\alpha I)\begin{pmatrix} I\\ X\end{pmatrix}\mathscr{B}(R;\beta,\beta_1,\alpha,\alpha_1), \tag{6.2a}$$

$$(\beta_1 H-\beta I)\begin{pmatrix} Y\\ I\end{pmatrix}\mathscr{B}(S;\alpha,\alpha_1,\beta,\beta_1) = (\alpha_1 H+\alpha I)\begin{pmatrix} Y\\ I\end{pmatrix}. \tag{6.2b}$$

Assuming no breakdown occurs, i.e., all involved inverses exist, in the end we will have error equations, similar to those in (3.15),

$$\Phi - X_k = (I-X_k\Psi)\left[\mathscr{B}(S;\alpha,\alpha_1,\beta,\beta_1)\right]^{2^k}\Phi\left[\mathscr{B}(R;\beta,\beta_1,\alpha,\alpha_1)\right]^{2^k}, \tag{6.3a}$$

$$\Psi - Y_k = (I-Y_k\Phi)\left[\mathscr{B}(R;\beta,\beta_1,\alpha,\alpha_1)\right]^{2^k}\Psi\left[\mathscr{B}(S;\alpha,\alpha_1,\beta,\beta_1)\right]^{2^k}. \tag{6.3b}$$

There are four cases to consider:

1. $\alpha_1 \neq 0$ and $\beta_1 \neq 0$. Since $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) = (\alpha_1/\beta_1)\cdot\mathscr{C}(t;\alpha/\alpha_1,\beta/\beta_1)$, both equations in (6.3) are the same as those for ADDA with the generalized Cayley transformation $\mathscr{C}(t;\alpha/\alpha_1,\beta/\beta_1)$. This implies that any resulting doubling algorithm is an ADDA.

2. $\alpha_1 \neq 0$, $\beta_1 = 0$ (and then $\beta \neq 0$ in order for $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1)$ to be well-defined):
   (a) If $\alpha = 0$, then $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) = (\alpha_1/\beta)t$ and thus the equations in (6.3) become

   $$\Phi - X_k = (I - X_k\Psi)S^{2^k}\Phi R^{-2^k}, \quad \Psi - Y_k = (I - Y_k\Phi)R^{-2^k}\Psi S^{2^k}. \tag{6.4}$$

   Convergence of X_k and Y_k to Φ and Ψ, respectively, is no longer guaranteed.
   (b) If $\alpha \neq 0$, then $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) = (\alpha/\beta)[t(\alpha/\alpha_1)^{-1} - 1]$ and thus the equations in (6.3) are the same as those for an SDA-ss. This implies that any resulting doubling algorithm is an SDA-ss.
3. $\alpha_1 = 0$ (and then $\alpha \neq 0$ in order for $\mathscr{B}(t;\beta,\beta_1,\alpha,\alpha_1)$ to be well-defined), $\beta_1 \neq 0$. This case is essentially the same as the previous one: $\alpha_1 \neq 0$, $\beta_1 = 0$.
4. $\alpha_1 = \beta_1 = 0$, i.e., $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1)$ is constant. This is the trivial case. Convergence of X_k and Y_k to Φ and Ψ, respectively, is not possible because no information on H is built into the algorithm.

In summary, possible doubling algorithms derivable from the general bilinear transformation are SDA, SDA-ss, ADDA, the trivial ones by $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) \equiv 1$ or $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) \equiv 0$, and the one by $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) = t$. Among these, optimal ADDA is the best.

In principle, possible doubling algorithms can also be constructed by noticing that, similar to (3.4) and (6.2),

$$h(H)\begin{pmatrix} I \\ X \end{pmatrix} = \begin{pmatrix} I \\ X \end{pmatrix} h(R), \quad h(H)\begin{pmatrix} Y \\ I \end{pmatrix}[h(S)]^{-1} = \begin{pmatrix} Y \\ I \end{pmatrix},$$

where $h(\cdot)$ is a rational function (or any other more complicated function). But without knowing a particular effective $h(\cdot)$, such a generality has no practical value.

7. Numerical examples. In this section, we shall present a few numerical examples to test numerical effectiveness of ADDA, in comparison with SDA and SDA-ss, as well as their ability to deliver entrywise relative accurate numerical solutions as argued in [36]. We will use two error measures to gauge accuracy in a computed solution $\widehat{\Phi}$: the normalized residual (NRes),

$$\text{NRes} = \frac{\|\widehat{\Phi}D\widehat{\Phi} - A\widehat{\Phi} - \widehat{\Phi}B + C\|}{\|\widehat{\Phi}\|(\|\widehat{\Phi}\|\|D\| + \|A\| + \|B\|) + \|C\|}, \tag{7.1}$$

a commonly used measure because it is readily available, and the entrywise relative error (ERErr),

$$\text{ERErr} = \max_{i,j} |(\widehat{\Phi} - \Phi)_{(i,j)}|/\Phi_{(i,j)}, \tag{7.2}$$

which is not available in actual computations but is made available here for testing purposes. In the case of ERErr, the indeterminant 0/0 is treated as 0. In using (7.1) hereafter, we use ℓ_1-operator norm $\|\cdot\|_1$ as an example. For all practical purposes, any matrix norm should work just fine.

Both errors defined by (7.1) and (7.2) are 0 if $\widehat{\Phi}$ is exact, but numerically they can only be made as small as $O(\boldsymbol{u})$ in general, where $\boldsymbol{u}$ is the unit machine roundoff. As we will see, to achieve $\widehat{\Phi}$ with deserved entrywise relative accuracy, tiny NRes, as tiny

as $O(\boldsymbol{u})$, is not sufficient. To get some idea about what deserved entrywise relative accuracy should be expected, we will first outline some of the main perturbation results in [36] and then present them along with our numerical results.

7.1. Deserved entrywise relative accuracy. Let[5] W be perturbed to $\widetilde{W}$ in such a way that

$$|\widetilde{A}-A|\le\epsilon|A|,\ |\widetilde{B}-B|\le\epsilon|B|,\ |\widetilde{C}-C|\le\epsilon C,\ |\widetilde{D}-D|\le\epsilon D, \tag{7.3}$$

where $0\le\epsilon<1$. It has been shown [36] that $\widetilde{\Phi}_{(i,j)}=0$ if and only if $\Phi_{(i,j)}=0$, under (7.3) and the assumption that both W and $\widetilde{W}$ are M-matrices. This fact paves the way to investigate how much each entry changes relatively.

Split A and B as

$$A=D_1-N_1,\quad D_1=\operatorname{diag}(A), \tag{7.4a}$$

$$B=D_2-N_2,\quad D_2=\operatorname{diag}(B). \tag{7.4b}$$

Correspondingly

$$A-\Phi D=D_1-N_1-\Phi D,\quad B-D\Phi=D_2-N_2-D\Phi,$$

and set

$$\lambda_1=\rho(D_1^{-1}(N_1+\Phi D)),\quad \lambda_2=\rho(D_2^{-1}(N_2+D\Phi)),\quad \lambda=\max\{\lambda_1,\lambda_2\}, \tag{7.5}$$

$$\tau_1=\frac{\min_i A_{(i,i)}}{\max_j B_{(j,j)}},\qquad \tau_2=\frac{\min_j B_{(j,j)}}{\max_i A_{(i,i)}}. \tag{7.6}$$

If W is nonsingular, then $A-\Phi D$ and $B-D\Phi$ are nonsingular M-matrices by Theorem 3.1; so $\lambda_1<1$ and $\lambda_2<1$ [32, Theorem 3.15, p. 90] and thus $0\le\lambda<1$. If W is an irreducible singular M-matrix, then by Theorem 3.1(d),

1. if $u_1^{\mathrm{T}}v_1>u_2^{\mathrm{T}}v_2$, then $\lambda_1<1$ and $\lambda_2=1$;
2. if $u_1^{\mathrm{T}}v_1<u_2^{\mathrm{T}}v_2$, then $\lambda_1=1$ and $\lambda_2<1$;
3. if $u_1^{\mathrm{T}}v_1=u_2^{\mathrm{T}}v_2$, then $\lambda_1=\lambda_2=1$.

The third case $u_1^{\mathrm{T}}v_1=u_2^{\mathrm{T}}v_2$, the so-called critical case, is rather extreme. It is argued in [19] that for the critical case for sufficiently small $\|\widetilde{W}-W\|$ there exists a constant θ such that

1. $\|\widetilde{\Phi}-\Phi\|\le\theta\|\widetilde{W}-W\|^{1/2}$;
2. $\|\widetilde{\Phi}-\Phi\|\le\theta\|\widetilde{W}-W\|$ if $\widetilde{W}$ is also singular.

This θ is known only by its existence.

The following results are taken from [36]. They are more informative but do not work for the critical case. Suppose that W is a nonsingular M-matrix or an irreducible singular M-matrix with $u_1^{\mathrm{T}}v_1\ne u_2^{\mathrm{T}}v_2$, ϵ in (7.3) is sufficiently small, and $\widetilde{W}$ is an M-matrix. We have

1.

$$|\Phi-\widetilde{\Phi}|\le\left[2\gamma\epsilon\,\mathbf{1}_{n,m}+O\left(\epsilon^2\right)\right]\Phi, \tag{7.7}$$

where γ is given by

$$(A-\Phi D)\Upsilon+\Upsilon(B-D\Phi)=D_1\Phi+\Phi D_2,\quad \gamma=\max_{i,j}\Upsilon_{(i,j)}/\Phi_{(i,j)}. \tag{7.8}$$

[5] We'll denote each perturbed counterpart by the same symbol but with a tilde.

2.

$$|\Phi - \widetilde{\Phi}| \le \left[2mn\,\kappa\chi\,\epsilon + O\left(\epsilon^2\right)\right]\Phi, \tag{7.9}$$

where κ is given by

$$(A - \Phi D)\Phi_1 + \Phi_1(B - D\Phi) = C, \quad \kappa = \max_{i,j} (\Phi_1)_{(i,j)}/\Phi_{(i,j)},$$

and dependent on different cases, χ is given by

(a) for nonsingular M-matrix W,

$$\chi = \max\left\{\frac{1+\lambda_1+(1+\lambda_2)\tau_1^{-1}}{1-\lambda_1+(1-\lambda_2)\tau_1^{-1}}, \frac{1+\lambda_2+(1+\lambda_1)\tau_2^{-1}}{1-\lambda_2+(1-\lambda_1)\tau_2^{-1}}\right\} \le \frac{1+\lambda}{1-\lambda}; \tag{7.10}$$

(b) for singular M-matrix W with $u_1^{\mathrm{T}} v_1 \ne u_2^{\mathrm{T}} v_2$,

$$\chi = 2 \times \begin{cases} \dfrac{1+\lambda_1+2\tau_1^{-1}}{1-\lambda_1} & \text{if } u_1^{\mathrm{T}} v_1 > u_2^{\mathrm{T}} v_2, \\ \dfrac{1+\lambda_2+2\tau_2^{-1}}{1-\lambda_2} & \text{if } u_1^{\mathrm{T}} v_1 < u_2^{\mathrm{T}} v_2. \end{cases} \tag{7.11}$$

It is proved that both γ and κ are finite [36]. Between (7.7) and (7.9), the linear term in the former is sharp while the one in the latter is not. But (7.9) is more informative in that it reveals the critical role played by the spectral radii λ_i in Φ's sensitivity.

In view of these perturbation results under (7.3) with $\epsilon = O(\boldsymbol{u})$, it is reasonable to define the *deserved entrywise relative accuracy* in any computed $\widehat{\Phi}$ to be that the associated ERErr is about $O(\gamma\boldsymbol{u})$ or $O(\kappa\chi\boldsymbol{u})$. In our examples in the next subsection, we shall compare ERErr against $(m+n)\gamma\boldsymbol{u}$ to verify if all our computed $\widehat{\Phi}$ at convergence have the deserved entrywise relative accuracy.

7.2. Examples. All computations are performed in MATLAB with $\boldsymbol{u} = 1.11 \times 10^{-16}$. Optimal parameters as specified in section 5 are used for ADDA, SDA, and SDA-ss. Kahan's stopping criteria [35],

$$\frac{(X_{k+1} - X_k)_{(i,j)}^2}{(X_k - X_{k-1})_{(i,j)} - (X_{k+1} - X_k)_{(i,j)}} \le \epsilon \cdot (X_{k+1})_{(i,j)} \quad \text{for all } i \text{ and } j, \tag{7.12}$$

are used to terminate iterations, where ϵ is a preselected tolerance. After numerous numerical experiments, we find that ϵ about 10^{-10} to 10^{-12} works the best for computed $\widehat{\Phi}$ to achieve its deserved accuracy without wasting the last iteration step.

Since ADDA is SDA if $\alpha_{\mathrm{opt}} = \beta_{\mathrm{opt}}$ for which there are numerous tests in the literature, our examples will mainly focus on the case

$$\alpha_{\mathrm{opt}} \stackrel{\mathrm{def}}{=} \max_i A_{(i,i)} \ne \beta_{\mathrm{opt}} \stackrel{\mathrm{def}}{=} \max_i B_{(i,i)}.$$

We will present three examples here. More examples can be found in [34]. Table 7.1 summarizes rates of convergence for ADDA, SDA-ss, and SDA for the examples, computed according to (5.2), (5.3), and (5.14). Also included in the table are quantities $\varrho(I - \Phi\Psi)$ and $\varrho(I - \Psi\Phi)$, which tell us how accurately all inverses of M-matrices $I - X_k Y_k$ and $I - Y_k X_k$ arising from the methods may be computed [35]. Table 7.2

TABLE 7.1
Rates of convergence of ADDA, SDA-ss, and SDA.

Example	r_{adda}	$r_{\mathrm{sda\text{-}ss}}$	r_{sda}	$\varrho(I-\Phi\Psi)$	$\varrho(I-\Psi\Phi)$
7.1 ($\xi = 1.5$)	0.58	0.75	0.64	0.5	0.5
7.1 ($\xi = 1+10^{-6}$)	$1-10^{-6}$	$1-7\cdot 10^{-7}$	$1-10^{-6}$	$1-2\cdot 10^{-6}$	$1-2\cdot 10^{-6}$
7.2	0.06	0.14	0.25	$6.3\cdot 10^{-2}$	$6.3\cdot 10^{-2}$
7.3	0.11	0.11	$1-2\cdot 10^{-4}$	$5.9\cdot 10^{-2}$	$1.1\cdot 10^{-1}$

TABLE 7.2
Parameters in the first order error bounds.

Example	λ_1	λ_2	2γ	κ	$\kappa\chi$
7.1 ($\xi = 1.5$)	0.78	1.0	15.0	3.0	84.0
7.1 ($\xi = 1+10^{-6}$)	$1-6.7\cdot 10^{-7}$	1.0	$6.0\cdot 10^{6}$	$1.0\cdot 10^{6}$	$1.2\cdot 10^{13}$
7.2	1	0.4	$3.2\cdot 10^{2}$	30.9	$1.6\cdot 10^{2}$
7.3	0.11	1	$2.1\cdot 10^{4}$	1.1	$4.8\cdot 10^{4}$

summarizes various stability parameters in the first order error bounds in subsection 7.1. They can and will be used to explain the entrywise relative accuracy in computed $\widehat{\Phi}$.

Example 7.1. In this example, $m = n = 2$ and

$$B = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}, \quad D = \mathbf{1}_{2,2}, \quad A = \xi \cdot B, \quad C = \xi \cdot D.$$

Making $\xi = 1$ and scaling W by 10^{-3} recovers a null recurrent case example in [4] (see also [20, Test 7.2]). It can be verified that $\Phi = \frac{1}{2}\mathbf{1}_{2,2}$ and $\Psi = \frac{1}{2\xi}\mathbf{1}_{2,2}$. Note also W is an irreducible singular M-matrix:

$$W\mathbf{1}_4 = 0, \quad \begin{pmatrix} \mathbf{1}_2 \\ \xi^{-1}\cdot\mathbf{1}_2 \end{pmatrix}^{\mathrm{T}} W = 0.$$

Figure 7.1 shows plots for $\xi = 1.5$ and $\xi = 1+10^{-6}$: the left ones for NRes and the right ones for ERErr. The horizontal dotted line in the right plots are $(m+n)\gamma\boldsymbol{u}$. If ERErr falls below the dotted line, we regard the computed $\widehat{\Phi}$ as having the deserved entrywise relative accuracy. We will follow this way of presenting iteration histories in the rest of the examples.

The case in which $\xi = 1$ is the critical case for which the doubling algorithms still converge but only linearly [11]. But for $0 < \xi \neq 1$ all three methods converge quadratically. In Figure 7.1 for $\xi = 1.5$, ADDA is the fastest, SDA comes in second, and SDA-ss is the slowest. There is little difference between SDA and ADDA for $\xi = 1+10^{-6}$ as expected and both are faster than SDA-ss, but not by much, and all three algorithms take about 24 iteration steps, about 3 times as many as that for $\xi = 1.5$.

Example 7.2.

$$A = \begin{pmatrix} 3 & -1 & & \\ & 3 & \ddots & \\ & & \ddots & -1 \\ -1 & & & 3 \end{pmatrix} \in \mathbb{R}^{n\times n},\ C = 2I_n,\ B = 10A,\ D = 10C.$$

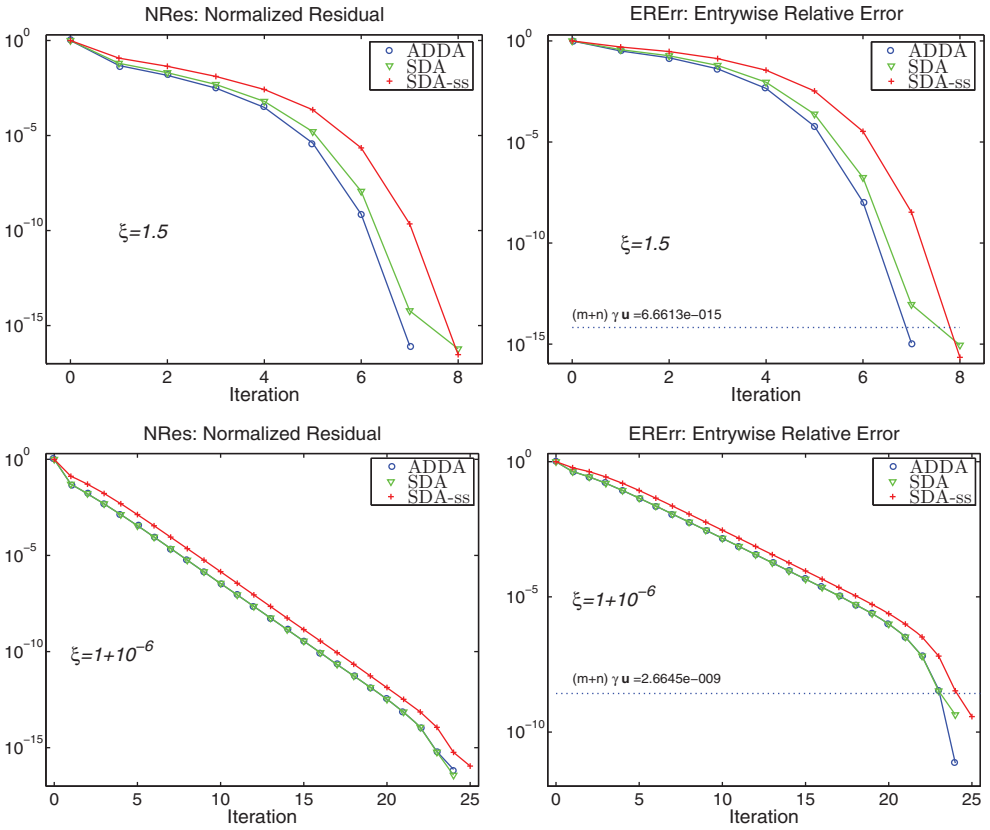

FIG. 7.1. *Example* 7.1 *for* $\xi = 1.5$ *and* $\xi = 1 + 10^{-6}$. *The case for* $\xi = 1 + 10^{-6}$ *is so close to the critical case, convergence by the three algorithms looks linear, except toward the very end. Note also much larger error bounds for the case* $\xi = 1 + 10^{-6}$ *than for the case* $\xi = 1.5$. *SDA-ss is actually slightly slower than SDA (and ADDA) for the two runs.*

W is an irreducible singular M-matrix: $W\mathbf{1}_{2n} = 0$, but $u_1^{\mathrm{T}}v_1 \neq u_2^{\mathrm{T}}v_2$. For testing purposes, we have computed for $n = 100$ an "exact" solution[6] Φ and Ψ by the computerized algebra system Maple with 100 decimal digits. This "exact" solution Φ's entries range from $5.7 \cdot 10^{-31}$ to $6.3 \cdot 10^{-2}$ and Ψ's entries range from $5.7 \cdot 10^{-30}$ to

[6]Thanks to an anonymous referee, these exact solutions can also be constructed explicitly. However, evaluating such explicitly constructed solutions does not guarantee the smallest entries in magnitude to be fully accurate due to harmful cancelations, unless the evaluation is done in a floating point arithmetic environment with precision about twice as much as the IEEE double precision floating point arithmetics. We outline the construction as follows. Since A is the sum of I_n and a special circulant matrix, we have [8, p. 356] $A = Q\Lambda Q^*$, where Q is unitary and Λ is diagonal and both are complex and known explicitly. Here Q^* is the complex conjugate transpose of Q. Let $\Phi_Q = Q^*\Phi Q$. MARE $\Phi D\Phi - A\Phi - \Phi B + C = 0$ can be transformed to $20\Phi_Q^2 - \Lambda\Phi_Q - 10\Phi_Q\Lambda + 2I = 0$ whose interested solution can be constructed from a basis matrix of the invariant subspace of $\left(\begin{smallmatrix} 10\Lambda & -20I \\ 2I & -\Lambda \end{smallmatrix}\right)$ associated with those eigenvalues of positive real parts. It can be seen that one such basis matrix takes the form $(X_1^{\mathrm{T}}, X_2^{\mathrm{T}})^{\mathrm{T}}$ with diagonal X_i, and consequently $\Phi_Q = X_2X_1^{-1}$ is diagonal. The n diagonal entries of Φ_Q can be computed by solving n scalar quadratic equations $20t^2 - 11\mu t + 2 = 0$ in t for each diagonal entry μ of Λ and picking the root t such that $\mu > t$ (because $B - D\Phi = Q(20\Lambda - 20\Phi_Q)Q^*$. Similarly $\Psi C\Psi - \Psi A - B\Psi + D = 0$ can be transformed to $2\Psi_Q^2 - \Psi_Q\Lambda - 10\Lambda\Psi_Q + 20I = 0$ whose interested solution is also diagonal for the same reason, where $\Psi_Q = Q^*\Psi Q$. As a by-product, one can argue that $\Psi_Q = 10\Phi_Q$ to conclude $\Psi = 10\Phi$.

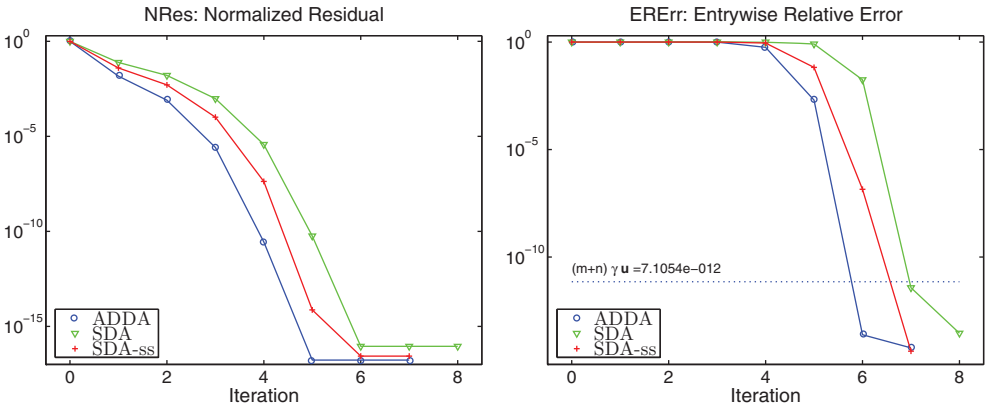

FIG. 7.2. *Example* 7.2. *Uneven convergence toward entries with widely different magnitudes. ERErr is still large even when NRes is already tiny before* $\widehat{\Phi}$ *is fully entrywise converged.*

$6.3 \cdot 10^{-1}$. Despite this wide range of magnitudes in their entries, all three methods are able to deliver computed $\widehat{\Phi}$ and $\widehat{\Psi}$ with entrywise relative errors at the level of $O(\boldsymbol{u})$. See Figure 7.2. Notice how few improvements are in ERErr for the first four iterations, even though NRes decreases substantially during the period. For example, at iteration 5,

	ADDA	SDA-ss	SDA
NRes	$1.6950 \cdot 10^{-17}$	$7.4124 \cdot 10^{-15}$	$5.7149 \cdot 10^{-11}$
ERErr	$2.0093 \cdot 10^{-3}$	$6.6470 \cdot 10^{-2}$	$8.1583 \cdot 10^{-1}$

This is because it takes a while for the tiny entries to gain some relative accuracy.

Example 7.3 (see [4, 20]). This is essentially the example of a positive recurrent Markov chain with nonsquare coefficients, originally from [4]. Here

$$A = 18 \cdot I_2, \quad B = 180002 \cdot I_{18} - 10^4 \cdot \mathbf{1}_{18,18}, \quad C = \mathbf{1}_{2,18}, \quad D = C^{\mathrm{T}}.$$

It is known $\Phi = \frac{1}{18} \cdot \mathbf{1}_{2,18} = \Psi^{\mathrm{T}}$. In this example, A and B differ a great deal in magnitude. Figure 7.3 shows the performance of the three methods. We see that ADDA and SDA-ss are about the same, and both are much faster than SDA.

Along with three examples above, we have conducted numerous other tests, including many random ones. We come up with the following two conclusions about speed and accuracy for the three doubling algorithms:

- ADDA is always the fastest among all three. SDA-ss can even run slower than SDA when $\max_i A_{(i,i)}$ and $\max_j B_{(j,j)}$ are about the same or differ within a factor of two. However, when $\max_i A_{(i,i)}$ and $\max_j B_{(j,j)}$ differ by a factor over, say, 10, ADDA and SDA-ss take about the same number of iterations to deliver fully converged $\widehat{\Phi}$ and both can be much faster than SDA.
- With the suggested optimal parameter selections in section 5, all three methods are capable of delivering computed $\widehat{\Phi}$ with the deserved entrywise relative accuracy as warranted by the input data.

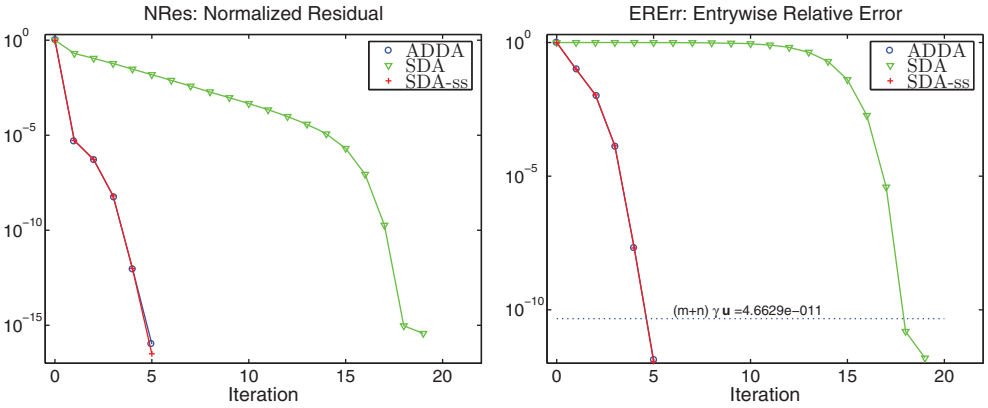

FIG. 7.3. *Example* 7.3. *ADDA and SDA-ss are barely distinguishable. Both are much faster than SDA.*

8. Concluding remarks. We have presented a new doubling algorithm for the unique minimal nonnegative solution Φ of MARE (1.1). It is the product of combining the alternating directional idea in ADI for the Sylvester equation (see [6, 33]) and the idea of SDA in [22]. For this reason, we name our new method the alternating-directional doubling algorithm. Compared with two existing double algorithms—SDA in [22] and SDA-ss in [9]—*our ADDA is always the fastest* as we argued first through theoretical convergence analysis and then numerical tests. Finally, all three methods are able to compute Φ as entrywise accurately as the perturbation analysis in [36] suggests.

All three doubling algorithms, SDA, SDA-ss, and ADDA, differing only in their initial setups, are constructed, respectively, by three special cases of the more general bilinear (or Möbius) transformation $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) = (\alpha_1 t - \alpha)/(\beta_1 t + \beta)$. Possible doubling algorithms derivable from $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1)$ are SDA, SDA-ss, ADDA, the trivial ones by $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) \equiv 1$ or $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) \equiv 0$, and the one by $\mathscr{B}(t;\alpha,\alpha_1,\beta,\beta_1) = t$. Among all, optimal ADDA is the best.

REFERENCES

[1] A. S. ALFA, J. XUE, AND Q. YE, *Accurate computation of the smallest eigenvalue of a diagonally dominant M-matrix*, Math. Comp., 71 (2002), pp. 217–236.
[2] B. D. O. ANDERSON, *Second-order convergent algorithms for the steady-state Riccati equation*, Internat. J. Control, 28 (1978), pp. 295–306.
[3] Z. BAI, J. DEMMEL, AND M. GU, *An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems*, Numer. Math., 76 (1997), pp. 279–308.
[4] N. G. BEAN, M. M. O'REILLY, AND P. G. TAYLOR, *Algorithms for return probabilities for stochastic fluid flows*, Stoch. Models, 21 (2005), pp. 149–184.
[5] P. BENNER, *Contributions to the Numerical Solution of Algebra Riccati Equations and Related Eigenvalue Problems*, Logos, Berlin, Germany, 1997.
[6] P. BENNER, R.-C. LI, AND N. TRUHAR, *On ADI method for Sylvester equations*, J. Comput. Appl. Math., 233 (2009), pp. 1035–1045.
[7] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, 1994.
[8] D. S. BERNSTEIN, *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed., Princeton University Press, Princeton, NJ, 2009.
[9] D. A. BINI, B. MEINI, AND F. POLONI, *Transforming algebraic Riccati equations into unilateral quadratic matrix equations*, Numer. Math., 116 (2010), pp. 553–578.
[10] A. Y. BULGAKOV AND S. K. GODUNOV, *Circular dichotomy of the spectrum of a matrix*, Siberian Math. J., 29 (1988), pp. 734–744.

[11] C.-Y. Chiang, E. K.-W. Chu, C.-H. Guo, T.-M. Huang, W.-W. Lin, and S.-F. Xu, *Convergence analysis of the doubling algorithm for several nonlinear matrix equations in the critical case*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 227–247.
[12] E. K.-W. Chu, H.-Y. Fan, and W.-W. Lin, *A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations*, Linear Algebra Appl., 396 (2005), pp. 55–80.
[13] E. K. W. Chu, H.-Y. Fan, W. W. Lin, and C. S. Wang, *Structure-preserving algorithms for periodic discrete-time algebraic Riccati equations*, Internat. J. Control, 77 (2004), pp. 767–788.
[14] M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics*, 2nd ed., Dover Publications, Mineola, NY, 2008.
[15] S. K. Godunov, *Problem of the dichotomy of the spectrum of a matrix*, Siberian Math. J., 27 (1986), pp. 649–660.
[16] C.-H. Guo, *Nonsymmetric algebraic Riccati equations and Wiener-Hopf factorization for M-matrices*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 225–242.
[17] C.-H. Guo, *A new class of nonsymmetric algebraic Riccati equations*, Linear Algebra Appl., 426 (2007), pp. 636–649.
[18] C.-H. Guo, *private communication*, June 2011.
[19] C.-H. Guo and N. Higham, *Iterative solution of a nonsymmetric algebraic Riccati equation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 396–412.
[20] C.-H. Guo, B. Iannazzo, and B. Meini, *On the doubling algorithm for a (shifted) nonsymmetric algebraic Riccati equation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1083–1100.
[21] 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.
[22] X. Guo, W. Lin, and S. Xu, *A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation*, Numer. Math., 103 (2006), pp. 393–412.
[23] J. Juang, *Existence of algebraic matrix Riccati equations arising in transport theory*, Linear Algebra Appl., 230 (1995), pp. 89–100.
[24] J. Juang and W.-W. Lin, *Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 228–243.
[25] W.-W. Lin and S.-F. Xu, *Convergence analysis of structure-preserving doubling algorithms for Riccati-type matrix equations*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 26–39.
[26] A. N. Malyshev, *Computing invariant subspaces of a regular linear pencil of matrices*, Siberian Math. J., 30 (1989), pp. 559–567.
[27] C. D. Meyer, *Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems*, SIAM Rev., 31 (1989), pp. 240–272.
[28] V. Ramaswami, *Matrix analytic methods for stochastic fluid flows*, in Proceedings of the 16th International Teletraffic Congress, Edinburgh, Elsevier Science, New York, 1999, pp. 19–30.
[29] L. Rogers, *Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains*, Ann. Appl. Probab., 4 (1994), pp. 390–413.
[30] R. A. Smith, *Matrix equation $XA + BX = C$*, SIAM J. Appl. Math., 16 (1968), pp. 198–201.
[31] X. Sun and E. S. Quintana-Ortí, *Spectral division methods for block generalized Schur decompositions*, Math. Comp., 73 (2004), pp. 1827–1847.
[32] R. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[33] E. L. Wachspress, *The ADI Model Problem*, www.netlib.org/na-digest-html/96/v96n36.html (1995).
[34] W.-G. Wang, W.-C. Wang, and R.-C. Li, *ADDA: Alternating-Directional Doubling Algorithm for M-Matrix Algebraic Riccati Equations*, Technical report 2011-04, Department of Mathematics, University of Texas at Arlington, May 2011; also available online from http://www.uta.edu/math/preprint/.
[35] J. Xue, S. Xu, and R.-C. Li, *Accurate solutions of M-matrix Sylvester equations*, Numer. Math., to appear.
[36] J. Xue, S. Xu, and R.-C. Li, *Accurate solutions of M-matrix algebraic Riccati equations*, Numer. Math., to appear.