NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 0000; **00**:1–18
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nla

A structure-preserving doubling algorithm for Lur'e Equations

Federico Poloni[1*], Timo Reis[2]

[1] *Dipartimento di Informatica, Università di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy*
[2] *Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, 22083 Hamburg, Germany*

SUMMARY

We introduce a numerical method for the numerical solution of the *Lur'e equations*, a system of matrix equations that arises, for instance, in linear-quadratic infinite time horizon optimal control. We focus on small-scale, dense problems. Via a Cayley transformation, the problem is transformed to the discrete-time case, and the structural infinite eigenvalues of the associated matrix pencil are deflated. The deflated problem is associated to a symplectic pencil with several Jordan blocks of eigenvalue 1 and even size, which arise from the nontrivial Kronecker chains at infinity of the original problem. For the solution of this modified problem, we use the *structure-preserving doubling algorithm* (SDA). Implementation issues such as the choice of the parameter γ in the Cayley transform are discussed. The most interesting feature of this method, with respect to the competing approaches, is the absence of arbitrary rank decisions, which may be ill-posed and numerically troublesome. The numerical examples presented confirm the effectiveness of this method.

Received ...

KEY WORDS: matrix equation; Lur'e equations; structured doubling algorithm; matrix pencil; optimal control

1. INTRODUCTION

Several problems in control theory, such as linear-quadratic optimal control, dissipativity analysis [1–5], model reduction [6–10], $\mathcal{H}_\infty$ control [11], differential games [12], lead to the computation of a special c-semi-stable deflating subspace $\mathcal{V}$ of a matrix pencil of the form

$$\mathcal{A} - s\mathcal{E} = \begin{bmatrix} 0 & A - sI & B \\ A^T + sI & Q & S \\ B^T & S^T & R \end{bmatrix} \tag{1}$$

with $A, Q \in \mathbb{R}^{n,n}$, $B, S \in \mathbb{R}^{n,m}$, $R \in \mathbb{R}^{m,m}$ and $Q = Q^T$, $R = R^T$. The word *c-semi-stable* here means that all the associated eigenvalues are in the closed left half-plane. When R is nonsingular, this problem (under a mild rank assumption) is equivalent to solving the *algebraic Riccati equation* (ARE) [11, 13, 14]

$$A^T X + XA - (XB + S)R^{-1}(XB + S)^T + Q = 0. \tag{2}$$

Its maximal symmetric solution $X_+ = X_+^T \in \mathbb{R}^{n,n}$ is related to the required semi-stable deflating subspace through

$$\mathcal{V} = \operatorname{im} \begin{bmatrix} X_+ & 0 \\ I_n & 0 \\ 0 & I_m \end{bmatrix}.$$

*Correspondence to: Dipartimento di Informatica, Università di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy. Email: fpoloni@di.unipi.it

While there is abundant literature on AREs including textbooks and survey articles [14–17], the case of singular R has been treated more sporadically in journal articles [5, 18–21]. The singularity of R is however a structural property in several applications [22] and can therefore not be excluded by arguments of genericity.

The closest analogue to (2) when R is singular are the *Lur'e equations* [23, 24]

$$\begin{aligned} A^T X + XA + Q &= K^T K, \\ XB + S &= K^T L, \\ R &= L^T L, \end{aligned} \tag{3}$$

to be solved for the triple $(X, K, L) \in \mathbb{R}^{n,n} \times \mathbb{R}^{p,n} \times \mathbb{R}^{p,m}$ with $X = X^T$ and the dimension p as small as possible.

Let us briefly review the known approaches for solving them: basically, these can be divided into elimination, eigenvalue-based and perturbation approaches:

a) The works [25, 26] present an iterative technique for the elimination of variables corresponding to $\ker R$: By performing an orthogonal transformation of R, and an accordant transformation of L, the equations can be divided into a 'regular part' and a 'singular part'. The latter leads to an explicit equation for a part of the matrix K. Plugging this part into (3), one obtains Lur'e equations of slightly smaller size. After a finite number of steps this leads to an algebraic Riccati equation. This also gives an equivalent solvability criterion that is obtained by the feasibility of this iteration.

b) By the introductory comments, there is a one-to-one correspondence between solutions of Lur'e equations and certain deflating subspaces of the matrix pencil (1) (see also Theorem 7). This allows to apply algorithms for computation of deflating subspaces for structured matrix pencils. For instance, the works [27–30] identify the same singular part by computing a staircase form of the pencil (1); this is done by repeatedly identifying the nullspaces of suitable submatrices via SVDs and deflating them with orthogonal transformations. After the staircase and Schur decomposition, another nontrivial issue in this setting is identifying which of the multiple eigenvalues at infinity belong to the correct deflating subspace.

c) In [21] a deflation technique is proposed. A "critical deflating subspace" of the even matrix pencil (1) is determined, using so-called *$\mathcal{E}$-neutral Wong sequences*. Thereafter, matrices which are spanning this critical subspace are used to eliminate certain parts of the Lur'e equation, and a projected algebraic Riccati equation is obtained. The focus of [21] is on large and sparse problems, so the resulting ARE is solved using a Newton-Kleinman iteration. The initial deflation procedure is still based on successive nullspace computations.

d) In the engineering practice, the most common approach to the solution of Lur'e equations is the perturbation of R by εI_m for some $\varepsilon > 0$. Then, by using the invertibility of $R + \varepsilon I$, the corresponding perturbed Lur'e equations are now equivalent to the Riccati equation

$$A^T X_\varepsilon + X_\varepsilon A - (X_\varepsilon B + S)(R + \varepsilon I)^{-1}(X_\varepsilon B + S)^T + Q = 0. \tag{4}$$

It is shown in [31,32] that the corresponding maximal solutions X_ε then converge to the maximal solution of (3).

The big problem of the perturbation approach d) is that, so far, there exist no bounds for the perturbation error $\|X - X_\varepsilon\|$. On top of that, the numerical condition of the Riccati equation (4) increases drastically as ε tends to 0.

The approaches in a), b) and c) also have numerical drawbacks: they rely on successive nullspace computations, which may be an arbitrarily ill-conditioned problem. In all of them it is necessary to identify, for several symmetric matrices M_k (starting from $M_0 = R$), two complementary subspaces $\mathcal{U}_{1,k}$ and $\mathcal{U}_{2,k}$ such that M_k is invertible when restricted to $\mathcal{U}_{1,k}$ and zero when restricted to $\mathcal{U}_{2,k}$. In practice, often this choice is not clear-cut, since the singular values of the matrices M_k may not

DOI: 10.1002/nla

have a large gap in magnitude. One needs to choose an arbitrary threshold under which they are set to zero; so it is possible to end up with matrices that are ill-conditioned on $\mathcal{U}_{1,k}$ and "not quite zero" on $\mathcal{U}_{2,k}$. Consider, as an illustrative example, the case $M_k = \mathrm{diag}(1, 10^{-1}, 10^{-2}, 10^{-3}, \ldots, 10^{-20})$, which has geometrically distributed eigenvalues.

Since all these approaches rely in an essential way on nullspace computations, the reader might be led to think that they are an unavoidable feature of the numerical solution of Lur'e equations, and all possible methods must perform them in one way or the other. We show here that this is not the case.

We present a numerical method based on a modification of the *structure preserving doubling algorithm* (SDA), an iterative scheme for continuous- and discrete-time algebraic Riccati equations [33]. It is shown in [34] that, unlike other iterative schemes, this algorithm has good convergence properties also when the pencil has eigenvalues (of even multiplicity) on the unit circle, as is the case in our problem. The algorithm is tailored to small-scale, dense problems and requires $O(n^3 \log \varepsilon^{-1})$ floating point operations to reach convergence up to an accuracy ε.

The method works directly on the unperturbed problem, without the need for regularization, and has the distinctive advantage that no rank decisions are needed. Hence this feature sets it apart from most algorithms for singular control problems that appeared in the literature.

As a byproduct of this analysis, we obtain some auxiliary results that are interesting in the context of the SDA literature: we derive a convergence result with weaker hypotheses (Theorem 16), and a new formula for its initial values that is more compact than the known ones and simplifies the implementation (Theorem 10). We discuss how this formula can be used to improve the heuristics to choose the parameter γ in the required Cayley transform.

2. CONTROL AND MATRIX THEORETIC PRELIMINARIES

The symbols $\|\cdot\|$, $\|\cdot\|_F$ stand for the spectral and Frobenius matrix norms, respectively. For Hermitian matrices $P, Q \in \mathbb{C}^{n,n}$, we write $P > Q$ ($P \geq Q$) if $P - Q$ is positive (semi-)definite. The symbol $\mathbb{R}(s)$ stands for the field of real rational functions. The symbol $\mathrm{Gl}_n(\mathbb{C})$ denotes the set of invertible matrices in $\mathbb{C}^{n,n}$.

For every positive k, we define the matrices $J_k, M_k, N_k \in \mathbb{R}^{k,k}$ as

$$J_k = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}, \quad M_k = \begin{bmatrix} & & 1 & 0 \\ & \iddots & \iddots & \\ 1 & \iddots & & \\ 0 & & & \end{bmatrix}, \quad N_k = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

Definition 1
Let $\mathcal{A} - s\mathcal{E}$ be a matrix pencil with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{m,n}$. Then $\mathcal{A} - s\mathcal{E}$ is called *regular* if $m = n$ and $\mathrm{rank}_{\mathbb{R}(s)}(\mathcal{A} - s\mathcal{E}) = n$. A pencil $\mathcal{A} - s\mathcal{E}$ is called *even* if $\mathcal{E} = -\mathcal{E}^T$ and $\mathcal{A} = \mathcal{A}^T$. A pencil with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{2n,2n}$ is called *symplectic* if $\mathcal{E}J\mathcal{E}^T = \mathcal{A}J\mathcal{A}^T$, with

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Many properties of a regular matrix pencil can be characterized in terms of the *Weierstrass canonical form (WCF)*.

Theorem 2 ([35])
For any regular matrix pencil $\mathcal{A} - s\mathcal{E}$ with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{n,m}$, there exist matrices $U_l \in \mathrm{Gl}_n(\mathbb{C})$, $U_r \in \mathrm{Gl}_m(\mathbb{C})$, such that

$$U_l(\mathcal{A} - s\mathcal{E})U_r = \mathrm{diag}(\mathcal{C}_1(s), \ldots, \mathcal{C}_k(s)), \tag{5}$$

where each of the pencils $\mathcal{C}_j(s)$ is of one of the types presented in Table I.

DOI: 10.1002/nla

Type	Size	$\mathcal{C}_j(s)$	Parameters
W1	$k_j \times k_j$	$(s-\lambda)I_{k_j} - N_{k_j}$	$k_j \in \mathbb{N}$, $\lambda \in \mathbb{C}$
W2	$k_j \times k_j$	$sN_{k_j} - I_{k_j}$	$k_j \in \mathbb{N}$

Table I. Block types in Weierstrass canonical form

Type	Size	$\mathcal{D}_j(s)$	Parameters
E1	$2k_j \times 2k_j$	$\begin{bmatrix} 0_{k_j,k_j} & (\lambda-s)I_{k_j}-N_{k_j} \\ (\overline{\lambda}+s)I_{k_j}-N_{k_j}^T & 0_{k_j,k_j} \end{bmatrix}$	$k_j \in \mathbb{N}$, $\lambda \in \mathbb{C}^+$
E2	$k_j \times k_j$	$\epsilon_j((-is-\mu)J_{k_j} + M_{k_j})$	$k_j \in \mathbb{N}, \mu \in \mathbb{R}$, $\epsilon_j \in \{-1, 1\}$
E3	$k_j \times k_j$	$\epsilon_j(isM_{k_j} + J_{k_j})$	$k_j \in \mathbb{N}$, $\epsilon_j \in \{-1, 1\}$

Table II. Block types in even Weierstrass canonical form

The numbers λ appearing in the blocks of type W1 are called the *(generalized) eigenvalues* of $\mathcal{A} - s\mathcal{E}$. Blocks of type W2 are said to be corresponding to infinite eigenvalues.

A special modification of the WCF for even matrix pencils, the so-called *even Weierstrass canonical form (EWCF)*, is presented in [36]. Note that there is also a 'realness-preserving version' of this result [37].

Theorem 3 ([36])
For any even matrix pencil $\mathcal{A} - s\mathcal{E}$ with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{n,n}$, there exists a matrix $U \in \mathrm{Gl}_n(\mathbb{C})$ such that

$$U^*(\mathcal{A} - s\mathcal{E})U = \mathrm{diag}(\mathcal{D}_1(s), \dots, \mathcal{D}_k(s)), \tag{6}$$

where each of the pencils $\mathcal{D}_j(s)$ is of one of the types presented in Table II.

The numbers ϵ_j in the blocks of type E2 and E3 are called the *block signatures*. The blocks of type E1 contain pairs $(\lambda, -\overline{\lambda})$ of generalized eigenvalues. Together with realness of $\mathcal{E}$ and $\mathcal{A}$, this implies that non-imaginary eigenvalues occur in quadruples $(\lambda, \overline{\lambda}, -\lambda, -\overline{\lambda})$. The blocks of type E2 and E3 correspond, respectively, to the purely imaginary and infinite eigenvalues.

Definition 4
An eigenvalue λ of a matrix pencil is called *c-stable*, *c-critical* or *c-anti-stable*, respectively, if $\mathrm{Re}(\lambda)$ is smaller than, equal to, or greater than 0. A right deflating subspace is called *c-stable* (resp., *c-anti-stable*) if it contains only c-stable (resp., c-anti-stable) eigenvalues, and *c-semi-stable* (resp., *c-semi-anti-stable*) if it contains only c-stable or c-critical (resp., c-anti-stable or c-critical) eigenvalues. The same definitions with the prefix c- replaced by d- hold if we change the expression $\mathrm{Re}(\lambda)$ to $|\lambda| - 1$.

Definition 5
Let $\mathcal{M} \in \mathbb{C}^{k,k}$ be given. A subspace $\mathcal{V} \subset \mathbb{C}^k$ is called $\mathcal{M}$*-neutral* if $x^*\mathcal{M}y = 0$ for all x, $y \in \mathcal{V}$.

Definition 6
Given $\gamma \in \mathbb{R}$, $\gamma \neq 0$, the *Cayley transform* [14, Section 2.9] of a regular pencil $\mathcal{A} - s\mathcal{E}$ is the pencil

$$\mathcal{A}_\gamma - s\mathcal{E}_\gamma, \qquad \mathcal{E}_\gamma = \mathcal{A} + \gamma\mathcal{E}, \quad \mathcal{A}_\gamma = \mathcal{A} - \gamma\mathcal{E}.$$

This is the extension to matrix pencils of the scalar map

$$\begin{aligned} \mathcal{C} : \quad \mathbb{C} \cup \{\infty\} &\to \mathbb{C} \cup \{\infty\}, \\ \lambda &\mapsto \frac{\lambda - \gamma}{\lambda + \gamma}. \end{aligned}$$

DOI: 10.1002/nla

We have $|\mathcal{C}(\lambda)| = 1$ if, and only if, λ is infinity or on the imaginary axis. Moreover, in the case $\gamma > 0$, we have $|\mathcal{C}(\lambda)| < 1$ if, and only if, $\mathrm{Re}(\lambda) > 0$, whereas, in the case $\gamma < 0$, there holds $|\mathcal{C}(\lambda)| < 1$ if, and only if, $\mathrm{Re}(\lambda) < 0$.

Via transformation into (even) Kronecker form, it can be seen that the Cayley transform of a matrix pencil preserves left and right eigenvectors and Jordan chains, while the eigenvalues are transformed according to $\lambda \mapsto \mathcal{C}(\lambda)$.

We recall from [20] the following theoretical results on Lur'e equations and their solvability that are needed in our article.

Theorem 7 ([20])
Let the Lur'e equations (3) with $A, Q \in \mathbb{R}^{n,n}$, $B, S \in \mathbb{R}^{n,m}$ and $R \in \mathbb{R}^{m,m}$ be given and assume that the associated even pencil (1) is regular and the pair (A, B) is stabilizable. Then, the following statements are equivalent.

(i) There exists a solution (X, K, L) of the Lur'e equations.

(ii) For all $\omega \in \mathbb{R}$ such that $\imath\omega$ is not an eigenvalue of A, it holds $\Phi(\imath\omega) \geq 0$, where

$$\begin{aligned}\Phi(s) = &\begin{bmatrix}(sI - A)^{-1}B \\ I_m\end{bmatrix}^* \begin{bmatrix}Q & S \\ S^T & R\end{bmatrix} \begin{bmatrix}(sI - A)^{-1}B \\ I_m\end{bmatrix} \\ = &R - \begin{bmatrix}B \\ S\end{bmatrix}^T \begin{bmatrix}0 & A - sI \\ A^T - sI & Q\end{bmatrix}^{-1} \begin{bmatrix}B \\ S\end{bmatrix} \in \mathbb{R}(s)^{m,m}.\end{aligned} \tag{7}$$

is the *spectral density function* or *Popov function* of the system.

(iii) In the EWCF of $\mathcal{A} - s\mathcal{E}$, all blocks of type E2 have positive block signature and even size, and all blocks of type E3 have negative sign and odd size.

Moreover, if the above conditions hold,

a) $\Phi(s) \in \mathbb{R}(s)^{m,m}$ is invertible (as a matrix with entries in the field $\mathbb{R}(s)$).

b) Among the solutions there is one, called the *stabilizing* solution and denoted by (X_+, K_+, L_+), such that $X \leq X_+$ for each other solution (X, K, L), in the positive definite ordering.

c) If (X_+, K_+, L_+) is the stabilizing solution, then the matrix

$$\mathcal{V} = \mathrm{im} \begin{bmatrix}X_+ & 0 \\ I_n & 0 \\ 0 & I_m\end{bmatrix} \tag{8}$$

spans the unique $n + m$-dimensional semi-c-stable $\mathcal{E}$-neutral subspace of the pencil (1).

d) If

$$\begin{bmatrix}Q & S \\ S^T & R\end{bmatrix} \geq 0, \tag{9}$$

then X_+ is the only (Hermitian) positive semidefinite solution of (3).

e) Let U be a matrix such that (6) holds, and partition it as $U = [\,U_1\,, \ldots, \,U_k\,]$, with block sizes compatible with the right-hand side of (6). The subspace $\mathcal{V}$ is spanned by

$$V = \begin{bmatrix}V_1 & \ldots & V_k\end{bmatrix} \in \mathbb{C}^{2n+m,n+m} \quad \text{for } V_j = U_j Z_j, \tag{10}$$

where

$$Z_j = \begin{cases}[\,I_{k_j}\,,\, 0_{k_j}\,]^T, & \text{if } \mathcal{D}_j \text{ is of type E1,} \\ [\,I_{k_j/2}\,,\, 0_{k_j/2}\,]^T, & \text{if } \mathcal{D}_j \text{ is of type E2,} \\ [\,I_{(k_j+1)/2}\,,\, 0_{(k_j-1)/2}\,]^T, & \text{if } \mathcal{D}_j \text{ is of type E3.}\end{cases}$$

DOI: 10.1002/nla

In other words, the subspace (8) contains all the vectors belonging to the Kronecker chains relative to c-stable eigenvalues, no vectors from the Kronecker chains relative to c-anti-stable eigenvalues, the first $k_j/2$ vectors from the chains relative to c-critical eigenvalues, and the first $(k_j+1)/2$ from the chains relative to eigenvalues at infinity.

Moreover, in the following we need these two elementary lemmas.

Lemma 8
Let $X, Y, G \in \mathbb{R}^{n,n}$ be symmetric matrices with $0 \le X \le Y$ and $G \le 0$; then

$$X(I - GX)^{-1} \le Y(I - GY)^{-1}.$$

Proof
Let $X_\varepsilon = X + \varepsilon I$ and $Y_\varepsilon = Y + \varepsilon I$; then, $Y_\varepsilon^{-1} \le X_\varepsilon^{-1}$ and both inverses exist. Thus, we have

$$(I - GY_\varepsilon)Y_\varepsilon^{-1} = Y_\varepsilon^{-1} - G \le X_\varepsilon^{-1} - G = (I - GX_\varepsilon)X_\varepsilon^{-1}.$$

Inverting the leftmost and rightmost term of the above inequality and letting $\varepsilon \to 0$ yields the desired result. □

Lemma 9
Let a nonsingular symmetric matrix and its inverse be partitioned as

$$\begin{bmatrix} X & Y \\ Y^T & Z \end{bmatrix}^{-1} = \begin{bmatrix} S & T \\ T^T & U \end{bmatrix}$$

with $X, S \in \mathbb{R}^{n_1,n_1}$, $Z, U \in \mathbb{R}^{n_2,n_2}$ and $Y, T \in \mathbb{R}^{n_1,n_2}$, such that, moreover, $X \le 0$, $Z \ge 0$. Then $S \le 0$, $U \ge 0$.

Proof
In the case where Z is nonsingular, the Schur complement formula yields $S^{-1} = X - YZ^{-1}Y^T \le 0$, and similarly for U if X is nonsingular. As above, a continuity argument can be used to obtain the thesis when these blocks are singular. □

3. THE STRUCTURED DOUBLING ALGORITHM AND ITS CONVERGENCE PROPERTIES

The structure-preserving doubling algorithm (SDA) [33,34,38] is a matrix iteration which computes two special deflating subspaces of a matrix pencil, one semi-stable and one semi-anti-stable. It is directly related to several other types of algorithms based on performing a "repeated squaring" in a matrix pencil setting [39–41].

A pencil $s\mathcal{L} - \mathcal{M}$ with $\mathcal{L}, \mathcal{M} \in \mathbb{R}^{N+M,N+M}$ is said to be in *standard symplectic-like form (SSF)* if

$$\mathcal{L} = \begin{bmatrix} I_N & -G \\ 0 & F \end{bmatrix}, \quad \mathcal{M} = \begin{bmatrix} E & 0 \\ -H & I_M \end{bmatrix}, \tag{11}$$

where the block sizes are chosen such that $E \in \mathbb{R}^{N,N}$ and $F \in \mathbb{R}^{M,M}$. The following result provides an easy method to transform a matrix pencil into one in SSF with the same spectral properties.

Theorem 10
Let $s\mathcal{E} - \mathcal{A}$ be a matrix pencil with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{N+M,N+M}$, and partition both matrices as

$$\mathcal{E} = \begin{bmatrix} \mathcal{E}_1 & \mathcal{E}_2 \end{bmatrix} \quad \mathcal{A} = \begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_2 \end{bmatrix}$$

with $\mathcal{E}_1, \mathcal{A}_1 \in \mathbb{R}^{N+M,N}$ and $\mathcal{E}_2, \mathcal{A}_2 \in \mathbb{R}^{N+M,M}$. A pencil $s\mathcal{L} - \mathcal{M}$ in standard symplectic-like form (11) having the same eigenvalues and right deflating subspaces of the original pencil exists if and only if $\begin{bmatrix} \mathcal{E}_1 & \mathcal{A}_2 \end{bmatrix}$ is nonsingular; in this case, it is unique and it can be computed using the relation

$$\begin{bmatrix} E & -G \\ -H & F \end{bmatrix} = \begin{bmatrix} \mathcal{E}_1 & \mathcal{A}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{A}_1 & \mathcal{E}_2 \end{bmatrix}. \tag{12}$$

DOI: 10.1002/nla

Proof
We are looking for a nonsingular matrix Q such that

$$sQ\begin{bmatrix}\mathcal{E}_1 & \mathcal{E}_2\end{bmatrix} - Q\begin{bmatrix}\mathcal{A}_1 & \mathcal{A}_2\end{bmatrix} = s\begin{bmatrix}I & -G\\ 0 & F\end{bmatrix} - \begin{bmatrix}E & 0\\ -H & I\end{bmatrix}.$$

By taking only some of the blocks from the above equation, we get

$$Q\mathcal{E}_1 = \begin{bmatrix}I\\ 0\end{bmatrix}, \quad Q\mathcal{A}_2 = \begin{bmatrix}0\\ I\end{bmatrix}, \qquad \text{i.e.,} \qquad Q\begin{bmatrix}\mathcal{E}_1 & \mathcal{A}_2\end{bmatrix} = \begin{bmatrix}I & 0\\ 0 & I\end{bmatrix},$$

thus Q must be the inverse of $\begin{bmatrix}\mathcal{E}_1 & \mathcal{A}_2\end{bmatrix}$. Taking the other two blocks we get

$$Q\mathcal{A}_1 = \begin{bmatrix}E\\ -H\end{bmatrix}, \quad Q\mathcal{E}_2 = \begin{bmatrix}-G\\ F\end{bmatrix},$$

which promptly yields (12). □

Notice that the formula in Theorem 10 can be applied also to SDA for continuous-time Riccati equation [33], where it yields (using the notation of [33])

$$\begin{bmatrix}\widehat{A} & \widehat{G}\\ -\widehat{H} & -\widehat{A}^T\end{bmatrix} = \begin{bmatrix}A_\gamma & -G\\ -H & -A_\gamma^T\end{bmatrix}^{-1}\begin{bmatrix}\bar{A}_\gamma & -G\\ -H & -\bar{A}_\gamma^T\end{bmatrix}. \tag{13}$$

Rearranging the blocks gives a system $MX = N$, with M and X $2n \times 2n$ symmetric matrices, whose solution costs $8n^3$ flops [42, Appendix C]. This compares favorably with the formulas in [33, Equations (9)–(11)], which require two LU factorizations, the solution of four linear systems of the form $MX = N$, one product and one explicit inversion, all of them involving unsymmetric $n \times n$ matrices, for a total cost of $(13 + \frac{1}{3})n^3$ flops [42]. Moreover, these formulas are simpler to analyze and can be implemented as a single LAPACK call.

The same trick can be applied, with computational advantage, to SDA for nonsymmetric algebraic Riccati equations [43].

We introduce now the transformation which will be the core of the SDA iteration.

Theorem 11
[44] Suppose that $s\mathcal{L} - \mathcal{M}$ is an SSF pencil such that both matrices $I_N - GH$ and $I_M - HG$ are nonsingular. Then, the deflating subspaces of the pencil

$$s\begin{bmatrix}I_N & -G'\\ 0 & F'\end{bmatrix} - \begin{bmatrix}E' & 0\\ -H' & I_M\end{bmatrix}, \tag{14}$$

$$\begin{aligned} E' &= E(I_N - GH)^{-1}E \quad & G' &= G + E(I_N - GH)^{-1}GF\\ F' &= F(I_M - HG)^{-1}F \quad & H' &= H + F(I_M - HG)^{-1}HE \end{aligned} \tag{15}$$

coincide with those of $s\mathcal{L} - \mathcal{M}$, and its eigenvalues are the squares of the corresponding eigenvalues of $s\mathcal{L} - \mathcal{M}$.

The structured doubling algorithm (see [45] for more details) consists in iterating the transformation (15), producing sequences (E_k, F_k, G_k, H_k) from a starting (E_0, F_0, G_0, H_0) defining a pencil in SSF.

The idea behing the convergence of the SDA iteration is that upon repeated squaring, eigenvalues with $|\lambda| < 1$ converge to zero and eigenvalues with $|\lambda| > 1$ to infinity. In fact, under suitable assumptions, convergence happens also in presence of unimodular eigenvalues. We report here a convergence result for the symplectic case, which is the only case needed in our paper. More general convergence results obtained with similar techniques can be found in [45]. We first need a few observations and a definition.

DOI: 10.1002/nla

Remark 1
When $N = M$, a pencil in SSF is symplectic (as defined in Section 2) if and only if $E^T = F$, $G = G^T$ and $H = H^T$. The SDA iteration preserves symplecticity, i.e., if (E_0, F_0, G_0, H_0) define a symplectic pencil (11), then so do (E_k, F_k, G_k, H_k) at each step k, and in particular $E_k^T = F_k$, $G_k = G_k^T$, $H_k = H_k^T$. Some computational savings can be obtained by exploiting this property in the algorithm; namely, one needs to compute only one of E_{k+1} and $F_{k+1} = E_{k+1}^T$ and only one of $I_N - G_k H_k$ and $I_M - H_k G_k = (I_N - G_k H_k)^T$.

Definition 12
Let $s\mathcal{L} - \mathcal{M}$ be a matrix pencil with eigenvalues $\lambda_1, \dots, \lambda_s$ and corresponding partial multiplicities $r_1, \dots, r_s$. Furthermore, assume that all partial multiplicities corresponding to the unimodular eigenvalues are even. Then the *canonical* semi-stable (resp. semi-unstable) subspace is defined as the unique deflating subspace whose associated eigenvalues have partial multiplicities

$$\begin{cases} r_\ell & \text{if } \lambda_\ell \text{ is d-stable (resp. unstable),} \\ r_\ell/2 & \text{if } \lambda_\ell \text{ is unimodular,} \\ 0 & \text{if } \lambda_\ell \text{ is d-unstable (resp. stable).} \end{cases}$$

We are now ready to state the well-definedness and convergence results for SDA that we shall need in the following; recall that $M = N$ in the symplectic case. Well-definedness of the sequence can be proven under suitable hypotheses, which hold true in the optimal control applications.

Theorem 13 ([38])
Let the SDA be applied to a symplectic pencil (11), and suppose that G_0, H_0 are semidefinite, one positive and one negative. Then, SDA is well-defined (i.e., $I - G_k H_k$ and $I - H_k G_k$ are nonsingular), and the sequences $0, G_0, G_1, \dots$ and $0, H_0, H_1, \dots$ are monotonic (in the positive-definite ordering).

Theorem 14 ([34])
Let the SDA be applied to a symplectic pencil (11) such that all its unimodular eigenvalues have even partial multiplicity. Suppose that there exist matrices $G_\infty \in \mathbb{R}^{N,N}$, $H_\infty \in \mathbb{R}^{N,N}$ such that

$$\begin{bmatrix} I_N \\ H_\infty \end{bmatrix}, \qquad \begin{bmatrix} G_\infty \\ I_N \end{bmatrix} \tag{16}$$

span, respectively, the canonical semi-stable and semi-unstable deflating subspaces of (11). Suppose in addition that the sequences $(E_k, F_k = E_k^T, G_k, H_k)$ defined by SDA are well-defined. Then,

- $\|E_k\| = \|F_k\| = \mathcal{O}(2^{-k})$,
- $\|H_\infty - H_k\| = \mathcal{O}(2^{-k})$,
- $\|G_\infty - G_k\| = \mathcal{O}(2^{-k})$.

4. A REDUCED LUR'E PENCIL

Let $s\mathcal{E} - \mathcal{A}$ be the pencil (1) associated to the Lur'e equations (3). Throughout the remaining part, we employ the following assumptions.

A1 The Lur'e equations (3) are solvable.
A2 The pencil (1) is regular.
A3 The pair (A, B) is stabilizable.

Let $\gamma > 0$ be such that both $\Phi(\gamma)$ as in (7) and $A - \gamma I$ are nonsingular (there exist at least one such γ, since $\operatorname{rank}_{\mathbb{R}(s)} \Phi(s) = m$ by assumption), and define

$$T := \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}, \qquad \mathcal{T} := \begin{bmatrix} T & 0 \\ 0 & 0_m \end{bmatrix}, \qquad \mathcal{B}_\gamma := \mathcal{A} - \gamma \mathcal{T}.$$

DOI: 10.1002/nla

We apply Theorem 10 to the Cayley transform $\mathcal{A}_\gamma - s\mathcal{E}_\gamma$ of $\mathcal{A} - s\mathcal{E}$, in order to obtain a pencil with the same eigenvalues and right deflating subspaces, which we denote by

$$s\widehat{\mathcal{L}} - \widehat{\mathcal{M}} = s\begin{bmatrix} I_n & -\widehat{G} \\ 0 & \widehat{F} \end{bmatrix} - \begin{bmatrix} \widehat{E} & 0 \\ -\widehat{H} & I_{n+m} \end{bmatrix}. \tag{17}$$

The expression which gives its blocks is

$$\begin{bmatrix} \widehat{E} & -\widehat{G} \\ -\widehat{H} & \widehat{F} \end{bmatrix} = \mathcal{B}_\gamma^{-1}\mathcal{B}_{-\gamma} \tag{18}$$

Notice that the inverse of $\mathcal{B}_\gamma$ exists, since both its leading $2n \times 2n$ principal block

$$\begin{bmatrix} 0 & A - \gamma I \\ A^T - \gamma I & Q \end{bmatrix}$$

and its Schur complement, which is exactly $\Phi(\gamma)$ as defined in Theorem 7, are nonsingular. An alternative expression for the quantity in (18) is

$$\begin{bmatrix} \widehat{E} & -\widehat{G} \\ -\widehat{H} & \widehat{F} \end{bmatrix} = \mathcal{B}_\gamma^{-1}\left(\mathcal{B}_\gamma + 2\gamma\mathcal{T}\right) = I_{2n+m} + 2\gamma\mathcal{B}_\gamma^{-1}\mathcal{T} = \begin{bmatrix} I_{2n} + 2\gamma ZT & 0 \\ * & I_m \end{bmatrix}, \tag{19}$$

where we denote by Z the leading principal $2n \times 2n$ block of $\mathcal{B}_\gamma^{-1}$. This form reveals a structure in the blocks: namely, we can set

$$\widehat{E} = E, \qquad \widehat{F} = \begin{bmatrix} F & 0 \\ * & I_m \end{bmatrix}, \qquad \widehat{G} = \begin{bmatrix} G & 0 \end{bmatrix}, \qquad \widehat{H} = \begin{bmatrix} H \\ * \end{bmatrix}, \tag{20}$$

so that the smaller blocks E, F, G, H have all size $n \times n$. Moreover, as Z is Hermitian in (19) and T has the only effect of swapping its block columns, it follows that $E^T = F$ and $G = G^T$, $H = H^T$, that is, the pencil

$$s\mathcal{L} - \mathcal{M} = s\begin{bmatrix} I_n & -G \\ 0 & F \end{bmatrix} - \begin{bmatrix} E & 0 \\ -H & I_n \end{bmatrix} \tag{21}$$

is symplectic.

What can we say about the spectral properties of $s\widehat{\mathcal{L}} - \widehat{\mathcal{M}}$ and $s\mathcal{L} - \mathcal{M}$? By the properties of the Cayley transform for $\gamma > 0$, c-(semi-) stable deflating subspaces of $\mathcal{A} - s\mathcal{E}$ become d-(semi-) unstable deflating subspaces of $\mathcal{A}_\gamma - s\mathcal{E}_\gamma$, and hence also of $s\widehat{\mathcal{L}} - \widehat{\mathcal{M}}$, and similarly c-(semi-) unstable deflating subspaces become d-(semi-)stable.

Moreover, the block structure in (20) implies that the pencil $s\widehat{\mathcal{L}} - \widehat{\mathcal{M}}$ is block lower triangular, with the leading $2n \times 2n$ diagonal block equal to $s\mathcal{L} - \mathcal{M}$ and the trailing $m \times m$ one equal to $sI_m - I_m$. Indeed, the column space of

$$\begin{bmatrix} 0_{2n\times m} \\ I_m \end{bmatrix},$$

is a deflating subspace of $\mathcal{A} - s\mathcal{E}$ with associated eigenvalue ∞ (with algebraic and geometric multiplicity m), and this is also a deflating subspace for $s\widehat{\mathcal{L}} - \widehat{\mathcal{M}}$ with eigenvalue $1 = \mathcal{C}_\gamma(\infty)$.

The pencil (21) is given by $P^T(s\mathcal{E} - \mathcal{A})P$, where P is the projection on

$$\left(\operatorname{im}\left(\begin{bmatrix} 0 \\ 0 \\ I_m \end{bmatrix}\right)\right)^\perp = (\ker \mathcal{E})^\perp.$$

With this characterization, it is easy to derive the WCF of $s\mathcal{L} - \mathcal{M}$ from that of $s\widehat{\mathcal{L}} - \widehat{\mathcal{M}}$: the projection corresponds to dropping the first row and column from each of the m blocks associated

DOI: 10.1002/nla

with 1. If the criteria in Theorem 7 hold, those blocks have odd length, hence after dropping one entry they become blocks of even length. Similarly, the blocks associated to the c-critical eigenvalues of $\mathcal{A} - s\mathcal{E}$, which must have even length, become d-critical eigenvalues of $s\mathcal{L} - \mathcal{M}$ (different from 1) of even length. So, all the unimodular eigenvalues of the reduced pencil $s\mathcal{L} - \mathcal{M}$ have even partial multiplicities.

As a consequence of this analysis of the Kronecker chains, we get the following result.

Theorem 15
Suppose that for $A, Q \in \mathbb{R}^{n,n}$, $B, S \in \mathbb{R}^{n,m}$ and $R \in \mathbb{R}^{m,m}$ the Lur'e equations (3) satisfy Assumptions **A1**, **A2** and **A3**. Let V span an deflating subspace of the reduced pencil (21) with E, F, G and H as in (19), (20). Then,

$$\widetilde{V} = \begin{bmatrix} V & 0 \\ 0 & I_m \end{bmatrix} \tag{22}$$

spans a deflating subspace of (1). In particular, if V spans the canonical n-dimensional d-semi-unstable deflating subspace of (21), then $\widetilde{V}$ spans the subspace $\mathcal{V}$ in (8). Moreover, the Lur'e equations have a stabilizing solution (8), and thus $X_+ = G_\infty$ exists in (16).

Remark 2
From the above discussion, one also obtains that the matrix X_+ appearing in (8) is the canonical weakly stabilizing solution of the discrete-time algebraic Riccati equation (DARE)

$$X = EX(I - HX)^{-1}E^T + G. \tag{23}$$

If the matrix H_∞ in (16) exists as well, then we can apply Theorem 14 to show that the sequence G_k generated by SDA converges to X_+. Unfortunately, this hypothesis is too restrictive for an important class of Lur'e equations, namely those corresponding to optimal control problems with positive semidefinite cost functional, i.e., when (9) holds true. Indeed, in all cases in which R is singular but the even pencil (1) is regular, there is an E3 block of length $k_j \geq 3$ in the EWCF of (1), and thus using [20, Lemma A.2], one can show that all solutions to the Lur'e equations are singular. In particular, the canonical anti-stabilizing subspace of the pencil is spanned by

$$\begin{bmatrix} X_- \\ I \end{bmatrix},$$

for a suitable solution X_-, and thus the topmost block is singular.

However, in numerical experiments, we observe that G_k converges to X_+ nevertheless, while H_k diverges and $G_k H_k$ and $H_k G_k$ are bounded. The same phenomenon was observed also in [34, Example 5.5] without a full proof.

We prove here a set of SDA convergence results with weaker hypotheses, which covers the cases in which only one of the two matrices G_∞ and H_∞ exists.

Theorem 16
Suppose that $G_0 \leq 0$ and $H_0 \geq 0$ in SDA, and that there exists at least one X satisfying

$$X \geq 0, \text{ and } \begin{bmatrix} I \\ X \end{bmatrix} \text{ is a deflating subspace of (11).} \tag{24}$$

Then, there is a minimal X_* satisfying (24) (i.e., one such that $X_* \leq X$ for each other X satisfying it), and the sequence H_k converges to X_*.

Proof
An early result in the theory of doubling methods [39] shows that $H_k = X_{2^k}$, where X_k is the sequence defined by

$$X_0 = 0, \qquad X_{k+1} = H_0 + E_0^T X_k (I - G_0 X_k)^{-1} E_0. \tag{25}$$

DOI: 10.1002/nla

Therefore, we may reduce the problem to computing the limit of (25). Notice that this is a fixed-point iteration for the DARE associated with the pencil (11). Using Lemma 8, we can easily prove by induction that $X_{k+1} \geq X_k$ and that $X - X_k \geq 0$ for each positive semidefinite solution X of the Riccati equation. The sequence X_k is bounded and increasing, and therefore it converges; its limit X_* is a positive semidefinite solution of the DARE, as obtained by passing (25) to the limit, and satisfies $X_* \leq X$ for every other solution $X \geq 0$. □

Remark 3
The same results hold with all the inequalities reversed (proof: if we change the signs of G_k and H_k for each k, the formulas in (15) are unchanged).

Remark 4
A corresponding result holds for G_k, namely: suppose that $G_0 \geq 0$, $H_0 \leq 0$, and there exists at least one Y such that

$$Y \geq 0, \text{ and } \begin{bmatrix} Y \\ I \end{bmatrix} \text{ is a deflating subspace of (11).} \tag{26}$$

Then, there is a minimal Y_* satisfying it, and $G_k \to Y_*$ (proof: apply the previous remark to the dual equation $Y = G_0 + E_0 Y (I - H_0 Y)^{-1} E_0^T$). As above, we may also reverse all inequalities and replace "minimal" with "maximal".

Remark 4 is the one that applies to our setting. We can prove the following convergence result by showing that its hypotheses are satisfied for the SSF pencil produced by Lur'e equations under condition (9).

Theorem 17
Let the solvable Lur'e equations (3) with $A, Q \in \mathbb{R}^{n,n}$, $B, S \in \mathbb{R}^{n,m}$ and $R \in \mathbb{R}^{m,m}$ be given and assume that the associated even pencil (1) is regular and the pair (A, B) is stabilizable. Furthermore, assume that

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \geq 0 \tag{27}$$

and let $\gamma > 0$ be such that $\Phi(\gamma)$ with $\Phi(s) \in \mathbb{R}(s)^{m,m}$ as in (7) and $\gamma I - A$ are nonsingular. Then, for the matrices E, F, G and H as in (19), (20), the SDA iteration is well-defined and the sequence G_k converges to the maximal solution X_+ of the Lur'e equations (3).

Proof
By Theorem 7 e), there is exactly one positive semidefinite solution X_+ to the Lur'e equations, and thus, by Theorem 13, there is exactly one $Y = X_+$ satisfying (26). In view of the modification of Theorem 16 given in Remark 4, we now only need to show that the matrices E, F, G and H as in (19), (20) fulfill $G \geq 0$ and $H \leq 0$. The former statement follows by Lemma 9. For the latter one, we first prove positive semidefiniteness by additionally assuming that R is nonsingular, and then invoke a continuity argument again as in the proof of Lemmas 8 and 9. When R is invertible, the leading $2n \times 2n$ block Z of $\mathcal{B}_\gamma^{-1}$ is the inverse of the Schur complement of R:

$$\begin{aligned} Z^{-1} &= \begin{bmatrix} 0 & A - \gamma I \\ A^T - \gamma I & Q \end{bmatrix} - \begin{bmatrix} B \\ S \end{bmatrix} R^{-1} \begin{bmatrix} B^T & S^T \end{bmatrix} \\ &= \begin{bmatrix} -BR^{-1}B^T & A - BR^{-1}S^T - \gamma I \\ (A - BR^{-1}S^T - \gamma I)^T & Q - CR^{-1}S^T \end{bmatrix}. \end{aligned} \tag{28}$$

Notice that $-BR^{-1}B^T \leq 0$ and $Q - SR^{-1}S^T \geq 0$, as the latter is a Schur complement in a positive semidefinite matrix, and thus Z^{-1} satisfies the hypotheses of Lemma 9. Applying that lemma, we obtain that $-G$ and $-H$, which are the diagonal blocks of Z, are nonpositive and nonnegative definite, respectively. □

Remark 5
Theorem 17 requires the positive semidefiniteness condition (27). This is a reasonable assumption

DOI: 10.1002/nla

in optimal control [5], which is one of the main motivations for the study of Lur'e equations (see Section 1). However, there are a lot of application areas where Lur'e equations arise in which the matrix $\left[\begin{smallmatrix} Q & S \\ S^T & R \end{smallmatrix}\right]$ is indefinite, such as in dissipativity analysis [1–5]. Indeed, the indefinite case can be led back to the semidefinite one in the following way: Assume that $Y \in \mathbb{R}^{n,n}$ is symmetric and fulfills the *linear matrix inequality (LMI)*

$$\begin{bmatrix} A^TY + YA + Q & YB + S \\ B^TY + S^T & R \end{bmatrix} \geq 0. \tag{29}$$

By [20, Lem. 17], (X, K, L) is a (stabilizing) solution of the Lur'e equations (3) if, and only if, $(X_Y, K_Y, L_Y) := (X - Y, K, L)$ is a (stabilizing) solution of the Lur'e equations

$$\begin{aligned} A^TX_Y + X_YA + Q_Y &= K_Y^TK_Y, \\ X_YB + S_Y &= K_Y^TL_Y, \\ R_Y &= L_Y^TL_Y \end{aligned} \tag{30}$$

with

$$\begin{bmatrix} Q_Y & S_Y \\ S_Y^T & R_Y \end{bmatrix} := \begin{bmatrix} A^TY + YA + Q & YB + S \\ B^TY + S^T & R \end{bmatrix}.$$

The latter matrix is positive semidefinite by (29); SDA will be convergent in this case, and we can solve the Lur'e equations (30) for $(X_Y, K_Y, L_Y) = (X - Y, K, L)$.
We note that the numerical solution of LMIs is hard from a numerical point of view [46]. However, a symmetric matrix $Y \in \mathbb{R}^{n,n}$ fulfilling the LMI (29) can often be obtained a priori by "physical knowledge" [4]. For instance, in model reduction of passivity [47, 48], Lur'e equations occur with $R = 0$, $B = -S$ and $A + A^T \geq 0$. In this case, the identity matrix $Y = I$ solves the LMI (29), and we can later on solve the modified Lur'e equations (30) for $(X_Y, K_Y, L_Y) = (X - I, K, L)$.

5. IMPLEMENTATION OF SDA FOR LUR'E EQUATIONS

Based on the results of the previous sections, we can use the SDA iteration to compute the solution to a Lur'e equation. The resulting algorithm is reported as Algorithm 1.

As we saw in Section 3, the symplecticity of the pencil is preserved during the SDA iterations, and helps reducing the computational cost of the iteration. Moreover, in this way we can preserve the eigenvalue symmetry of the original pencil along the iteration.

Algorithm 1: A structured doubling algorithm for the maximal solution of a Lur'e equation

input : A, B, Q, R, S defining Lur'e equations (3) fulfilling **A1–A3**
output: An approximation of the maximal solution X_+
Choose a suitable $\gamma > 0$;
Compute

$$M \longleftarrow \begin{bmatrix} 0 & A - \gamma I & B \\ A^T - \gamma I & Q & S \\ B^T & S^T & R \end{bmatrix}^{-1} \begin{bmatrix} 0 & A + \gamma I & \\ A^T + \gamma I & Q & \\ B^T & S^T & \end{bmatrix};$$

Partition

$$M = \begin{bmatrix} E & -G \\ -H & E^T \\ * & * \end{bmatrix};$$

Use SDA on E, $F = E^T$, G,H to compute G_∞, H_∞;
Return $X_+ = G_\infty$;

DOI: 10.1002/nla

Algorithm 1 produces a sequence G_k of approximations of the maximal solution X. Corresponding sequences K_k, L_k of L and K satisfying (3) can be constructed by performing an eigenvalue decomposition

$$\begin{bmatrix} A^TX_k + X_kA + Q & X_kB + S \\ B^TX_k + S^T & R \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix}^T,$$
$$\Sigma_1 \in \mathbb{R}^{m,m}, \quad U_1 \in \mathbb{R}^{m+n,m}, \quad \Sigma_2 \in \mathbb{R}^{n,n}, \quad U_2 \in \mathbb{R}^{m+n,n},$$

ordered such that Σ_1 contains the largest diagonal elements, and taking

$$\begin{bmatrix} K_k & L_k \end{bmatrix} = \Sigma_1^{1/2} U_1^T.$$

Notice, though, that K and L are non-unique and are typically not needed in applications; we use them here only to check the residual of the Lur'e equations *a posteriori*. Namely, with this choice of K_k and L_k, we can define the *relative Lur'e residual* as

$$\frac{\left\| \begin{bmatrix} A^TX_k + X_kA + Q & X_kB + S \\ B^TX_k + S^T & R \end{bmatrix} - \begin{bmatrix} K_k^T \\ L_k^T \end{bmatrix} \begin{bmatrix} K_k & L_k \end{bmatrix} \right\|_F}{\left\| \begin{bmatrix} A^TX_k + X_kA + Q & X_kB + S \\ B^TX_k + S^T & R \end{bmatrix} \right\|_F}. \tag{31}$$

A delicate choice which affects the accuracy of the computed solution is the choice of γ in the Cayley transform. A heuristic strategy to this purpose is presented in [33]. The authors perform an error analysis in the ∞ norm for their version of the formulas that give the initial values E_0, G_0, H_0 of SDA, obtaining a first-order upper bound $F(\gamma)$ for the absolute error, and then apply an univariate optimization method to approximate $\arg\min F(\gamma)$. This heuristic is not always satisfactory, as it minimizes the error in the first step of the algorithm only; in particular, the objective function $F(\gamma)$ has a qualitatively different behavior from the actual error attainable by SDA in the limit $\gamma \to 0$: the former typically converges to a finite limit, while the latter diverges. However, up to our knowledge, it is the only such heuristic available.

The simpler expression for the SDA initial values given in (13) allows one to apply the standard accuracy theory for linear systems in order to give a simpler error bound for their computation; namely, the forward error is bounded by

$$\hat{F}(\gamma) = \kappa_\infty \left(\begin{bmatrix} A_\gamma & -G \\ -H & -A_\gamma^T \end{bmatrix} \right) \left\| \begin{bmatrix} \bar{A}_\gamma & -G \\ -H & -\bar{A}_\gamma^T \end{bmatrix} \right\|_\infty.$$

This formula gives a tighter bound than the one in [33], for instance in cases in which the block A_γ is ill-conditioned but the full matrix is well-conditioned. The new approach can be extended easily to Lur'e equations: the equation for the initial values is (18), and thus we have the error estimate

$$f(\gamma) = \kappa_\infty(\mathcal{A}_\gamma) \left\| \mathcal{A}_{-\gamma} \right\|_\infty.$$

Hence in our experiments we use the same optimization method as [33] (Fibonacci search), but with this new objective function $f(\gamma)$.

6. NUMERICAL EXPERIMENTS

We have implemented Algorithm 1 (SDA-L) using MATLAB, and applied it to the following test problems.

P1 a Lur'e equation with a random stable matrix $A \in \mathbb{R}^{n,n}$, a random $S = B$, $Q = 0$ and R the $m \times m$ matrix with all the entries equal to 1, with $\operatorname{rank}(R) = 1$. Namely, B was generated with the command

DOI: 10.1002/nla

Figure 1. Relative residual for **P1**

n	m	SDA-L	R+S $\varepsilon = 10^{-6}$	R+S $\varepsilon = 10^{-8}$	R+S $\varepsilon = 10^{-12}$	R+N $\varepsilon = 10^{-8}$
10	3	$5 \cdot 10^{-15}$	$2 \cdot 10^{-08}$	$8 \cdot 10^{-10}$	$10 \cdot 10^{-06}$	$3 \cdot 10^{-10}$
50	5	$4 \cdot 10^{-15}$	$8 \cdot 10^{-09}$	$2 \cdot 10^{-08}$	$2 \cdot 10^{-04}$	$4 \cdot 10^{-10}$
500	10	$2 \cdot 10^{-14}$	$8 \cdot 10^{-10}$	$2 \cdot 10^{-08}$	$2 \cdot 10^{-04}$	$8 \cdot 10^{-10}$

Figure 2. Relative residual for **P2**

Problem #	SDA-L	R+S $\varepsilon = 10^{-6}$	R+S $\varepsilon = 10^{-8}$	R+S $\varepsilon = 10^{-12}$	R+N $\varepsilon = 10^{-8}$
3	$2 \cdot 10^{-15}$	$6 \cdot 10^{-02}$	$6 \cdot 10^{-02}$	$6 \cdot 10^{-02}$	$1 \cdot 10^{-09}$
4	$4 \cdot 10^{-15}$	$6 \cdot 10^{-07}$	$6 \cdot 10^{-09}$	$9 \cdot 10^{-08}$	$6 \cdot 10^{-09}$
5	$7 \cdot 10^{-13}$	$3 \cdot 10^{-07}$	$1 \cdot 10^{-09}$	$2 \cdot 10^{-08}$	$1 \cdot 10^{-09}$
6	$1 \cdot 10^{-15}$	$7 \cdot 10^{-12}$	$2 \cdot 10^{-13}$	$4 \cdot 10^{-13}$	$2 \cdot 10^{-09}$

Figure 3. Forward error for **P3**

n	SDA-L	R+S $\varepsilon = 10^{-6}$	R+S $\varepsilon = 10^{-8}$	R+S $\varepsilon = 10^{-12}$	R+N $\varepsilon = 10^{-8}$
1	$1 \cdot 10^{-08}$	$1 \cdot 10^{-03}$	$10 \cdot 10^{-05}$	$1 \cdot 10^{-06}$	$1 \cdot 10^{-04}$
2	$8 \cdot 10^{-05}$	$3 \cdot 10^{-02}$	$1 \cdot 10^{-02}$	$4 \cdot 10^{-02}$	$1 \cdot 10^{-02}$
3	$4 \cdot 10^{-03}$	$1 \cdot 10^{-01}$	$6 \cdot 10^{-02}$	$1 \cdot 10^{+01}$	$6 \cdot 10^{-02}$
4	$3 \cdot 10^{-02}$	$4 \cdot 10^{-01}$	$2 \cdot 10^{-01}$	$10 \cdot 10^{-01}$	$5 \cdot 10^{-01}$
5	$8 \cdot 10^{-02}$	$1 \cdot 10^{+00}$	$5 \cdot 10^{-01}$	$2 \cdot 10^{+00}$	$6 \cdot 10^{-01}$

```
B=rand(n,m);
```

To generate a stable A, we used the following sequence of commands:

```
V=randn(n);
W=randn(n);
A=-V*V'-W+W';
```

P2 a set of problems motivated from real-world examples, taken with some modifications from the benchmark set CAREX [49]. Namely, we took Examples 3 to 6 (the real-world applicative problems) of that paper, which are a set of real-world problems varying in size and numerical characteristics, and changed the value of R to get a singular problem. In the original versions of all examples, R is the identity matrix of appropriate size; we simply replaced its $(1,1)$ entry with 0, in order to get a singular problem.

P3 a highly ill-conditioned problem with larger Kronecker blocks: with $m = 1$, $A = I_n + N_n$, $B = e_n$ (the last column of the $n \times n$ identity matrix), $S = -B$, $R = 0$ and

$$Q = \begin{bmatrix} -2 & -1 & & & \\ -1 & -2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & -2 & -1 \\ & & & -1 & -2 \end{bmatrix}.$$

Such a problem corresponds to a Kronecker chain of length $2n+1$ associated to an infinite eigenvalue, and its canonical semi-stable solution is $X = I$. Notice that the conditioning of the deflating subspace problem in this case is $\epsilon^{1/(2n+1)}$, for an unstructured perturbation of the input data of the order of the machine precision ϵ [50, section 16.5].

The results of SDA-L are compared to those of a regularization method as the one described in (4), for different values of the regularization parameter ε. After the regularization, the equations are

DOI: 10.1002/nla

solved using SDA after a Cayley transform with the same parameter γ (R+S), or with the matrix sign method with norm scaling [42, 51] (R+N). We point out that the control toolbox of MATLAB contains a command `gcare` that solves a so-called *generalized continuous-time algebraic Riccati equation*; this is equivalent to finding X_+ for a pencil in the form (1). However, this command is not designed to deal with a singular R, nor with eigenvalues numerically on the imaginary axis. Therefore, when applied to nearly all these experiments, this command fails reporting the presence of eigenvalues too close to the imaginary axis.

For the problem **P3**, where an analytical solution $X = I$ is known, we reported in Figure 3 the values of the relative forward error

$$\frac{\left\|\tilde{X} - X\right\|_F}{\|X\|_F}.$$

For **P1** and **P2**, for which no analytical solution is available, we computed instead the relative Lur'e residual (31), which are in Figures 1 and 2, respectively.

We see that in all the experiments our solution method obtains a better result than the ones based on regularization. The reader may wonder why the residual for problem 5 in **P2** is two order of magnitude larger than for the other problems. It turns out that the culprit is the choice of γ in the Cayley transform: with a hand-picked value, the error drops to $9 \cdot 10^{-16}$. This shows that the heuristic for the choice of γ is still not perfect; as far as we know, finding the optimal value of the parameter γ is still an open problem in all applications of Cayley transforms.

Finally, we wish to show on an example that algorithms based on successive projections and rank decisions may encounter trouble in cases in which determining the kernel of R is an ill-posed problem. We choose a different test problem that belongs to this category.

P4 Defined exactly as **P1**, with dimensions $n = 20$, $m = 16$, but R is replaced with a matrix with geometrically distributed singular values, generated with the MATLAB command

```
R=gallery('randsvd', m, 1e8, 3); R = R*R';
```

The eigenvalues of this matrix R are

```
1.2482e-16 1.1744e-15 1.3583e-14 1.5856e-13 1.8479e-12
2.1544e-11 2.5119e-10 2.9286e-09 3.4145e-08 3.9811e-07
4.6416e-06 5.4117e-05 6.3096e-04 7.3564e-03 8.5770e-02
1.0000e+00.
```

Note, indeed, that there is not a spectral gap that allows to separate clearly the nullspace and the image of R; the eigenvalues are almost equispaced between 1 and 10^{-16}.

We compare the results obtained by SDA-L with those obtained after deflating the subspace $\mathcal{W}$ corresponding to the defective eigenvalues at infinity with the method suggested in [21, Algorithm 2] (with different choices of the threshold τ), and solving the projected problem (which has a nonsingular coefficient R) with the MATLAB function `care`. Other projection-based methods would essentially have to do the same nullspace computations, possibly embedded in a staircase form calculation. The results are in Figure 4.

The results show that the obtained accuracy is very sensitive to the chosen value of the rank decision threshold τ. Different choices lead to different subspaces being identified as the subspace at infinity, and the accuracy of the computed solution varies in a difficult-to-predict way. Our proposed method SDA-L, in contrast, does not require to pick a threshold arbitrarily, and gets better results even compared to the best-performing choice of τ.

7. CONCLUSION AND OPEN ISSUES

In this work we have introduced a new numerical method for the solution of Lur'e matrix equations. Unlike previous methods based on regularization, this approach allows one to solve the original

DOI: 10.1002/nla

Figure 4. Relative residuals for the projection method in [21] with different thresholds and for SDA-L.

Threshold τ	Residual	$\dim \mathcal{W}$
$1 \cdot 10^{-06}$	$2 \cdot 10^{-08}$	14
$1 \cdot 10^{-07}$	$3 \cdot 10^{-09}$	13
$1 \cdot 10^{-08}$	$4 \cdot 10^{-10}$	12
$1 \cdot 10^{-09}$	$6 \cdot 10^{-11}$	11
$1 \cdot 10^{-10}$	$1 \cdot 10^{-12}$	9
$1 \cdot 10^{-11}$	$2 \cdot 10^{-13}$	8
$1 \cdot 10^{-12}$	$3 \cdot 10^{-14}$	7
$1 \cdot 10^{-13}$	$4 \cdot 10^{-13}$	6
$1 \cdot 10^{-14}$	$1 \cdot 10^{-14}$	5
$1 \cdot 10^{-15}$	$8 \cdot 10^{-3}$	3
$1 \cdot 10^{-16}$	$1 \cdot 10^{-2}$	2
$1 \cdot 10^{-17}$	$1 \cdot 10^{-2}$	1
$1 \cdot 10^{-18}$	$6 \cdot 10^{-3}$	0
SDA-L	$3 \cdot 10^{-15}$	—

equation without introducing any artificial perturbation and without relying on possibly ill-posed rank determinations.

The first step of this approach is applying a Cayley transform to convert the problem to an equivalent discrete-time pencil. In this new form, the infinite eigenvalues can be easily deflated, reducing the problem to a discrete-time algebraic Riccati equation with eigenvalues on the unit circle. For the solution of this latter equation, the structured-preserving doubling algorithm was chosen, due to its good convergence properties in presence of eigenvalues on the unit circle, as proved in [34]. Direct methods, such as the symplectic eigensolvers presented in [52], can also be used for the solution of the deflated DARE.

Moreover, we derive a novel, simpler formula (12) for the initial values of SDA, and, correspondingly, a simplification of the heuristic criterion in [33] for the choice of the parameter γ of the Cayley transform.

The numerical experiments confirm the effectiveness of our new approach for regular matrix pencils. It is not clear whether a similar method can be adapted to work in cases in which the pencil (1) is singular, a situation which may indeed happen in the context of Lur'e equations. Another issue is finding a method to exploit the low-rank structure of Q (when present). These further developments are currently under our investigation.

Acknowledgment F. Poloni acknowledges the support of the PRA project "Mathematical models and computational methods for complex networks" of the University of Pisa.

REFERENCES

1. Anderson B. Algebraic description of bounded real matrices. *Electronics Letters* 1966; **2**(12):464–465.
2. Anderson B, Newcomb R. Impedance synthesis via state-space techniques. *Institution of Electrical Engineering Proceedings* 1968; **115**(7):928–936.
3. Anderson B, Vongpanitlerd S. *Network Analysis and Synthesis*. Prentice Hall: Englewood Cliffs, NJ, 1973.
4. Willems J. Dissipative dynamical systems. II: Linear systems with quadratic supply rates. *Archive for Rational Mechanics and Analysis* 1972; **45**:352–393.
5. Clements D, Anderson B. *Singular optimal control: the linear-quadratic problem*, *Lecture Notes in Control and Information Sciences*, vol. 5. Springer-Verlag: Berlin-New York, 1978.
6. Chen X, Wen J. Positive realness preserving model reduction with $\mathcal{H}_\infty$ norm error bounds. *Systems & Control Letters* 2005; **54**(4):361–374.
7. Gugercin S, Antoulas A. A survey of model reduction by balanced truncation and some new results. *International Journal of Control* 2004; **77**(8):748–766.
8. Opdenacker P, Jonckheere E. A contraction mapping preserving balanced reduction scheme and its infinity norm error bounds. *IEEE Transactions on Circuits and Systems I: Regular Papers* 1988; **35**(2):184–189.

DOI: 10.1002/nla

9. Phillips J, Daniel L, Silveira L. Guaranteed passive balancing transformations for model order reduction. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 2003; **22**(8):1027–1041.
10. Reis T, Stykel T. Passive and bounded real balancing for model reduction of descriptor systems. *International Journal of Control* 2010; **83**(1):74–88.
11. Zhou K, Doyle J, Glover K. *Robust and Optimal Control*. Prentice-Hall: Princeton, 1996.
12. Basar T, Bernhard P. *$\mathcal{H}_\infty$-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Birkhäuser: Boston-Basel-Berlin, 1995.
13. Willems J. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Transactions on Automatic Control* 1971; **16**:621–634.
14. Lancaster P, Rodman L. *Algebraic Riccati equations*. Clarendon Press: Oxford, 1995.
15. Abou-Kandil H, Freiling G, Ionescu V, Jank G. *Matrix Riccati equations. In control and systems theory*. Systems & Control: Foundations & Applications, Birkhäuser Verlag: Basel, 2003.
16. Benner P, Saak J. Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: a state of the art survey. *GAMM-Mitt.* 2013; **36**(1):32–52.
17. Bini D, Iannazzo B, Meini B. *Numerical solution of algebraic Riccati equations*, *Fundamentals of Algorithms*, vol. 9. SIAM publishing: Philadelphia, PA, 2012.
18. Yakubovich V. Singular problem in optimal control of linear stationary system with quadratic functional. *Siberian Mathematical Journal* 1985; **26**(1):148–158.
19. Ionescu V, Oară C, Weiss M. *Generalized Riccati theory and robust control. A Popov function approach*. John Wiley & Sons Ltd.: Chichester, 1999.
20. Reis T. Lur'e equations and even matrix pencils. *Linear Algebra and its Applications* 2010; **434**(1):152–173.
21. Poloni F, Reis T. A deflation approach for large-scale Lur'e equations. *SIAM Journal on Matrix Analysis and Applications* 2012; **33**(4):1339–1368.
22. Reis T, Stykel T. Pabtec: Balanced truncation for electrical circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 2010; **29**(9):1354–1367.
23. Lur'e A. *Certain Nonlinear Problems in the Theory of Automatic Control*. Gostekhizat: Moscow, Leningrad, 1951. Translated into English, H.M. Stationery, 1957.
24. Barabanov N. Kalman-Yakubovich lemma in general finite-dimensional case. *International Journal of Robust and Nonlinear Control* 2007; **17**:369–386.
25. Jacobson D, Speyer J. Necessary and sufficient conditions for optimality for singular control problems: a transformation approach. *Journal of Mathematical Analysis and Applications* 1971; **33**:163–187.
26. Weiss H, Wang Q, Speyer J. System characterization of positive real conditions. *IEEE Transactions on Automatic Control* 1994; **39**(3):540–544.
27. Byers R, Mehrmann V, Xu H. A structured staircase algorithm for skew-symmetric/symmetric pencils. *Electronic Journal of Linear Algebra* 2007; **26**:1–33.
28. Van Dooren P. The computation of Kronecker's canonical form of a singular pencil. *Linear Algebra and its Applications* 1979; **27**:103–140, doi:10.1016/0024-3795(79)90035-1. URL http://dx.doi.org/10.1016/0024-3795(79)90035-1.
29. Van Dooren PM. The generalized eigenstructure problem in linear system theory. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control* 1981; **26**(1):111–129, doi:10.1109/TAC.1981.1102559. URL http://dx.doi.org/10.1109/TAC.1981.1102559.
30. Van Dooren P. A generalized eigenvalue approach for solving Riccati equations. *Society for Industrial and Applied Mathematics. Journal on Scientific and Statistical Computing* 1981; **2**(2):121–135, doi:10.1137/0902010. URL http://dx.doi.org/10.1137/0902010.
31. Jacobson D, Speyer J. Necessary and sufficient conditions for singular control problems: a limit approach. *Journal of Mathematical Analysis and Applications* 1971; **34**:239–266.
32. Trentelman H. Families of linear-quadratic problems: Continuity properties. *IEEE Transactions on Automatic Control* 1987; **AC-32**(4):323–329.
33. Chu EKW, Fan HY, Lin WW. A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations. *Linear Algebra and its Applications* 2005; **396**:55–80, doi:10.1016/j.laa.2004.10.010.
34. Huang TM, Lin WW. Structured doubling algorithms for weakly stabilizing Hermitian solutions of algebraic Riccati equations. *Linear Algebra and its Applications* 2009; **430**(5-6):1452–1478, doi:10.1016/j.laa.2007.08.043.
35. Gantmacher F. *Theory of Matrices*. Chelsea Publishing Company: New York, 1959.
36. Thompson R. The characteristic polynomial of a principal subpencil of a Hermitian matrix pencil. *Linear Algebra and its Applications* 1976; **14**:135–177.
37. Thompson R. Pencils of complex and real symmetric and skew matrices. *Linear Algebra and its Applications* 1991; **147**:323–371.
38. Lin WW, Xu SF. Convergence analysis of structure-preserving doubling algorithms for Riccati-type matrix equations. *SIAM Journal on Matrix Analysis and Applications* 2006; **28**(1):26–39 (electronic), doi:10.1137/040617650.
39. Anderson B. Second-order convergent algorithms for the steady-state Riccati equation. *International Journal of Control* 1978; **28**(2):295–306.
40. Bai Z, Demmel J, Gu M. An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems. *Numerische Mathematik* 1997; **76**(3):279–308, doi:10.1007/s002110050264.
41. Benner P. Contributions to the numerical solution of algebraic Riccati equations and related eigenvalue problems. Doctoral dissertation, Fakultät für Mathematik, TU Chemnitz-Zwickau, Chemnitz 1997. Published by Logos-Verlag, Berlin.
42. Higham N. *Functions of matrices*. Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, 2008. Theory and computation.
43. Guo XX, Lin WW, Xu S. A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation. *Numerische Mathematik* 2006; **103**(3):393–412, doi:10.1007/s00211-005-0673-7.

44. Chu EW, Fan HY, Lin WW, Wang CS. Structure-preserving algorithms for periodic discrete-time algebraic Riccati equations. *International Journal of Control* 2004; **77**(8):767–788.
45. Chiang CY, Chu EKW, Guo CH, Huang TM, Lin WW, Xu SF. Convergence analysis of the doubling algorithm for several nonlinear matrix equations in the critical case. *SIAM Journal on Matrix Analysis and Applications* 2009; **31**(2):227–247, doi:10.1137/080717304.
46. Boyd S, El Ghaoui L, Feron E, Balakrishnan V. *Linear Matrix Inequalities in System and Control Theory*, *Studies in Applied Mathematics*, vol. 15. Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, 1994.
47. Reis T, Willems J. A balancing approach to the realization of systems with internal passivity and reciprocity. *Systems & Control Letters* 2011; **60**(1):69–74.
48. Guiver C, Opmeer MR. Error bounds in the gap metric for dissipative balanced approximations. *Linear Algebra and its Applications* 2013; **439**:3659–3698.
49. Benner P, Laub A, Mehrmann V. A collection of benchmark examples for the numerical solution of algebraic Riccati equations I: the continuous-time case. *Technical Report SPC 95-22*, Forschergruppe 'Scientific Parallel Computing', Fakultät für Mathematik, TU Chemnitz-Zwickau 1995. Version dated February 28, 1996.
50. Gohberg I, Lancaster P, Rodman L. *Invariant subspaces of matrices with applications*, *Classics in Applied Mathematics*, vol. 51. Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, 2006. Reprint of the 1986 original.
51. Mehrmann V. *The autonomous linear quadratic control problem: theory and numerical solution*, *Lecture Notes in Control and Information Sciences*, vol. 163. Springer-Verlag: Berlin, 1991, doi:10.1007/BFb0039443.
52. Fassbender H. *Symplectic methods for the symplectic eigenproblem*. Kluwer Academic/Plenum Publishers, New York, 2000.

DOI: 10.1002/nla