

A Structured Doubling Algorithm for the Numerical Solution of Lur'e Equations

F. Poloni^a, T. Reis^{b,*}

^aScuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

^bInstitut für Numerische Simulation, Technische Universität Hamburg-Harburg, Schwarzenbergstraße 95 E, 21073 Hamburg, Germany

Abstract

We introduce a numerical method for the numerical solution of the Lur'e matrix equations

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

that arise, for instance, in linear-quadratic infinite time horizon optimal control. The method is based on the characterization of the solutions in terms of deflating subspaces of a suitable even matrix pencil. Via a Cayley transformation, the problem is transformed to the discrete-time case, and only the infinite eigenvalues of the original problem associated with $\ker \mathcal{E}$ are deflated. This leaves us with a symplectic problem with several Jordan blocks of eigenvalue 1 and even size, which arise from the remaining eigenvalues at infinity of the original problem. For the solution of this modified problem, we use the *structure-preserving doubling algorithm* (SDA), an iterative scheme for the solution of dense continuous- and discrete-time algebraic Riccati equations. Unlike other iterative schemes, this algorithm converges also when the pencil has eigenvalues on the unit circle, as is the case in our problem. Implementation issues such as the choice of the parameter γ in the Cayley transform are discussed. The numerical examples presented confirm the effectiveness of this method.

Keywords: matrix equation, Lur'e equations, structured doubling algorithm, matrix pencil, optimal control

1. Introduction

A classical result of control theory is that for $A, Q \in \mathbb{R}^{n,n}$, $B, C \in \mathbb{R}^{n,m}$, $R \in \mathbb{R}^{m,m}$ with $Q = Q^T$, $R = R^T$, a linear-quadratic optimal control problem of type

Minimize

$$\mathcal{J}(u(\cdot), x_0) = \frac{1}{2} \int_0^\infty \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} Q & C \\ C^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \quad (1a)$$

subject to $u(\cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}^m$ and

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t), & x(0) &= x_0 \\ \lim_{t \rightarrow \infty} x(t) &= 0,\end{aligned} \quad (1b)$$

*Corresponding author

Email addresses: f.poloni@sns.it (F. Poloni), timo.reis@tu-harburg.de (T. Reis)

leads to the *algebraic Riccati equation (ARE)*

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

In particular, the existence of an optimal control is equivalent to solvability of (2), and the optimal control can be constructed by a state feedback that involves a solution X of (2) [1, 2]. Assembling the algebraic Riccati equation however requires the invertibility of the matrix R . The case of singular R is referred to as a *singular optimal control problem*[3, 4, 5, 6] which may cause that the optimal control becomes distributional and/or non-unique. Instead of setting up the ARE, these equations are now formulated as a system

$$\begin{aligned} A^T X + XA + Q &= K^T K, \\ XB + C &= K^T L, \\ R &= L^T L, \end{aligned} \quad (3)$$

which has now 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 p as small as possible. The nonlinear system is called *Lur'e equations* after A.I. LUR'E [7] (see [8] for an historical overview). This type of equations not only occurs in linear-quadratic optimal control, but the can be also used to analyze properties like dissipativity of linear systems [9, 10, 11, 12]. Another important field of application is in balancing-related model reduction, in particular positive real balanced truncation and bounded real balanced truncation [13, 14, 15, 16, 17].

In the case where R is invertible, we have $m = p$, and the matrices K and L can be eliminated by obtaining the algebraic Riccati equation (2). Whereas this type is well-explored both from an analytical and numerical point of view [18, 1, 19], the case of singular R has been treated stepmotherly. However, the singularity of R is often a structural property of the system to be analyzed [20] and can therefore not be excluded by arguments of genericity.

From a theoretical point of view, Lur'e equations have been investigated in [21, 8]. The solution set is completely characterized in [22] via the consideration of the matrix pencil

$$s\mathcal{E} - \mathcal{A} = \begin{bmatrix} 0 & -sI + A & B \\ sI + A^T & Q & C \\ B^T & C^T & R \end{bmatrix}. \quad (4)$$

This pencil has the special property of being *even*, that is \mathcal{E} is skew-Hermitian and \mathcal{A} is Hermitian, and it takes the role of the Hamiltonian matrix for algebraic Riccati equations. Solvability of (3) is characterized via the eigenstructure of this pencil, and furthermore there is a correspondence between the solutions of the Lur'e equations and certain *deflating subspaces* of (4) [22]. Deflating subspaces are a generalization of the concept of invariant subspaces to matrix pencils [23]. Under some slight additional conditions of the pair (A, B) , such as controllability of the system (1b), it is shown in [22] that there exists a so-called *maximal solution* X . Here, maximality means that X is, in terms of semi-definiteness, above all other solutions of the Lur'e equations.

The aim of this work is to present a method for the numerical solution of Lur'e equations. Before our approach is presented, let us briefly review the known approaches for this problem: Basically, these can be divided into elimination and perturbation approaches:

- a) The works [24, 25] 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) The most common approach to the solution of Lur'e equations is the slight 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

$$A^T X_\varepsilon + X_\varepsilon A - (XB + C)(R + \varepsilon I)^{-1}(X_\varepsilon B + C)^T + Q = 0. \quad (5)$$

It is shown in [26, 27] that the corresponding maximal solutions X_ε then converge to the maximal solution of (3).

Whereas the first approach has the great disadvantage that it relies on successive nullspace computations (which may be an arbitrarily ill-conditioned numerical problem), the big problem of the perturbation approach 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 (5) increases drastically as ε tends to 0.

The approach presented in this work is of different nature. Via a Cayley transformation, the problem is transformed to the discrete-time case, and only the m infinite eigenvalues of the original problem associated with $\ker \mathcal{E}$ are deflated. This leaves us with a symplectic problem with several Jordan blocks of eigenvalue 1 and even size, which arise from the remaining eigenvalues at infinity of the original problem. It is indeed shown in [22] that a singular R causes the presence of higher-order Kronecker block of odd size associated with the eigenvalue $\lambda = \infty$. For the solution of this modified problem, we use the *structure-preserving doubling algorithm*, an iterative scheme for continuous- and discrete-time algebraic Riccati equations [28]. It is shown in [29] that, unlike other iterative schemes, this algorithm converges also when the pencil has eigenvalues (of even multiplicity) on the unit circle, as is the case in our problem.

The paper is organized as follows. Section 2 introduces the notation and contains some required control and matrix theoretic background, in particular a normal form for even matrix pencils is introduced. In Section 3 we briefly present some results from [22] which connect the spectral properties of the even pencil (4) to the solvability and the solutions of the Lur'e equations. Section 4 outlines the basics of the structure-preserving doubling algorithm. This method requires the matrices to be in a special structure, the so-called *standard symplectic-like form*, which is for matrix pencils introduced in Section 5. For pencils arising in the analysis of Lur'e equations, this standard symplectic-like form is considered in Section 6. Together with the previously introduced theory, this leads to an iterative method for the numerical solution of Lur'e equations. Further details concerning implementation and, in particular, stopping criteria for the iteration, are discussed in Section 7. The theory and methods introduced in this article are finally illustrated in Section 8 by means of several numerical examples.

2. Control and Matrix Theoretic Preliminaries

Throughout the paper real and complex numbers are denoted by \mathbb{R} and \mathbb{C} , the open left and right half-planes by \mathbb{C}^- and \mathbb{C}^+ , respectively. The symbol i stands for the imaginary unit, and by \bar{z} we denote the complex conjugate of $z \in \mathbb{C}$. Natural numbers excluding 0 are denoted by \mathbb{N} . The space of $n \times m$ complex matrices is denoted by $\mathbb{C}^{n,m}$, and the set of invertible complex $n \times n$ matrices by $\text{Gl}_n(\mathbb{C})$. The matrices A^T and A^* denote, respectively, the transpose and the conjugate transpose of $A \in \mathbb{C}^{n,m}$. We denote by $\text{rank}(A)$ the rank, by $\text{im}(A)$ the image, by $\ker(A)$ the kernel, by $\sigma(A)$ the spectrum of a matrix A . The symbols $\|\cdot\|$, $\|\cdot\|_F$ respectively stand for the spectral and Frobenius matrix norms. For Hermitian matrices $P, Q \in \mathbb{C}^{n,n}$, we write $P > Q$ ($P \geq Q$) if $P - Q$ is positive (semi-)definite.

For a rational matrix-valued function $\Phi : \mathbb{C} \setminus D \rightarrow \mathbb{C}^{n,m}$, where $D \subset \mathbb{C}$ is the finite set of poles, we define the normal rank by $\text{normalrank } \Phi = \max_{s \in \mathbb{C} \setminus D} \text{rank } \Phi(s)$.

With $A_i \in \mathbb{C}^{n_i, m_i}$ with $m_i, n_i \in \mathbb{N} \cup \{0\}$ for $i = 1, \dots, k$, we denote the block diagonal matrix by $\text{diag}(A_1, \dots, A_k)$. An identity matrix of order n is denoted by I_n or simply by I . The zero $n \times m$ ($n \times n$) matrix is denoted by $0_{n,m}$ (resp. 0_n) or simply by 0. Moreover, for $k \in \mathbb{N}$ we define the matrices $J_k, M_k, N_k \in \mathbb{R}^{k,k}$, $K_k, L_k \in \mathbb{R}^{k-1,k}$ as

$$J_k = \begin{bmatrix} & & & 1 \\ & & \ddots & \\ & & & \\ 1 & & & \end{bmatrix}, \quad K_k = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 0 \\ & & & 1 \end{bmatrix}, \quad L_k = \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{bmatrix},$$

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

Definition 1. Let $s\mathcal{E} - \mathcal{A}$ be a matrix pencil with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{m,n}$. Then $s\mathcal{E} - \mathcal{A}$ is called regular if $m = n$ and $\text{normalrank}(s\mathcal{E} - \mathcal{A}) = n$. A pencil $s\mathcal{E} - \mathcal{A}$ 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 $\mathcal{E}J\mathcal{E}^T = \mathcal{A}J\mathcal{A}^T$, where

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

Many properties of a matrix pencil can be characterized in terms of the *Kronecker canonical form (KCF)*.

Type	Size	$C_j(s)$	Parameters
K1	$k_j \times k_j$	$(s - \lambda)I_{k_j} - N_{k_j}$	$k_j \in \mathbb{N}, \lambda \in \mathbb{C}$
K2	$k_j \times k_j$	$sN_{k_j} - I_{k_j}$	$k_j \in \mathbb{N}$
K3	$(k_j - 1) \times k_j$	$sK_{k_j} - L_{k_j}$	$k_j \in \mathbb{N}$
K4	$k_j \times (k_j - 1)$	$sK_{k_j}^T - L_{k_j}^T$	$k_j \in \mathbb{N}$

Table 1: Block types in Kronecker canonical form

Theorem 2. [23] For a matrix pencil $s\mathcal{E} - \mathcal{A}$ with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{n,m}$, there exist matrices $U_l \in \text{Gl}_n(\mathbb{C})$, $U_r \in \text{Gl}_m(\mathbb{C})$, such that

$$U_l(s\mathcal{E} - \mathcal{A})U_r = \text{diag}(C_1(s), \dots, C_k(s)), \quad (6)$$

where each of the pencils $C_j(s)$ is of one of the types presented in Table 1.

The numbers λ appearing in the blocks of type K1 are called the (generalized) eigenvalues of $s\mathcal{E} - \mathcal{A}$. Blocks of type K2 are said to be corresponding to infinite eigenvalues.

A special modification of the KCF for even matrix pencils, the so-called *even Kronecker canonical form (EKCF)* is presented in [30]. Note that there is also a ‘realness-preserving version’ of this result [31].

Type	Size	$\mathcal{D}_j(s)$	Parameters
E1	$2k_j \times 2k_j$	$\begin{bmatrix} 0_{k_j, k_j} & (\lambda - s)I_{k_j} - N_{k_j} \\ (\bar{\lambda} + s)I_{k_j} - N_{k_j}^T & 0_{k_j, k_j} \end{bmatrix}$	$k_j \in \mathbb{N}, \lambda \in \mathbb{C}^+$
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	$k_j \times k_j$	$\epsilon_j(isM_{k_j} + J_{k_j})$	$k_j \in \mathbb{N}, \epsilon_j \in \{-1, 1\}$
E4	$(2k_j - 1) \times (2k_j - 1)$	$\begin{bmatrix} 0_{k_j-1, k_j-1} & -sK_{k_j} + L_{k_j} \\ sK_{k_j}^T + L_{k_j}^T & 0_{k_j, k_j} \end{bmatrix}$	$k_j \in \mathbb{N}$

Table 2: Block types in even Kronecker canonical form

Theorem 3. [30] For an even matrix pencil $s\mathcal{E} - \mathcal{A}$ with $\mathcal{E}, \mathcal{A} \in \mathbb{R}^{n,n}$, there exists a matrix $U \in \text{Gl}_n(\mathbb{C})$ such that

$$U^*(s\mathcal{E} - \mathcal{A})U = \text{diag}(\mathcal{D}_1(s), \dots, \mathcal{D}_k(s)), \quad (7)$$

where each of the pencils $\mathcal{D}_j(s)$ is of one of the types presented in Table 2. The numbers ε_j in the blocks of type E2 and E3 are called the block signatures.

The blocks of type E1 contains pairs $(\lambda, -\bar{\lambda})$ of generalized eigenvalues. Together with realness of \mathcal{E} and \mathcal{A} , this implies that non-imaginary eigenvalues occur in quadruples $(\lambda, \bar{\lambda}, -\lambda, -\bar{\lambda})$. The blocks of type E2 and E3 respectively correspond to the purely imaginary and infinite eigenvalues. Blocks of type E4 consist of a combination of blocks that are equivalent to those of type K3 and K4. Note that regularity of the pencil $s\mathcal{E} - \mathcal{A}$ is equivalent to the absence of blocks of type E4.

Definition 4. A subspace $\mathcal{V} \subset \mathbb{C}^N$ is called (right) deflating subspace for the pencil $s\mathcal{E} - \mathcal{A}$ with $\mathcal{E}, \mathcal{A} \in \mathbb{C}^{M,N}$ if for a matrix $V \in \mathbb{C}^{N,k}$ with full column rank and $\text{im } V = \mathcal{V}$, there exists an $l \leq k$ and matrices $W \in \mathbb{C}^{M,l}$, $\tilde{E}, \tilde{A} \in \mathbb{C}^{l,k}$ with

$$(s\mathcal{E} - \mathcal{A})V = W(s\tilde{E} - \tilde{A}), \quad (8)$$

Definition 5. An eigenvalue λ of a matrix pencil is called *c-stable*, *c-critical* or *c-anti-stable* respectively if $\text{Re}(\lambda)$ 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 hold replacing the prefix *c-* with *d-* if we replace the expression $\text{Re}(\lambda)$ with $|\lambda| - 1$.

Definition 6. Let $M \in \mathbb{C}^{k,k}$ be given. A subspace $\mathcal{V} \subset \mathbb{C}^k$ is called *M-neutral* if $x^* M y = 0$ for all $x, y \in \mathcal{V}$.

Definition 7. Let a pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,m}$ be given. Then

- (i) (A, B) is called *controllable* if $\text{rank}[sI - A, B] = n$ for all $s \in \mathbb{C}$;
- (ii) (A, B) is called *stabilizable* if $\text{rank}[sI - A, B] = n$ for all $s \in \mathbb{C} \setminus \mathbb{C}^-$.

Definition 8. Given $\gamma \in \mathbb{R}$, $\gamma \neq 0$, the Cayley transform of a regular pencil $s\mathcal{E} - \mathcal{A}$ is the pencil

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

In the scalar case, the Cayley transform reads

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

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

Via transformation into (even) Kronecker form, it can be seen that Cayley transform of a matrix pencil preserves left and right eigenvectors and principal vectors, while the eigenvalues λ transform to $C(\lambda)$. In particular, blocks of type K1 (W1) of size k for λ are mapped to blocks of same type and size, but now for $C(\lambda)$.

3. Solvability of Lur'e equations

In this part we collect theoretical results being equivalent for the solvability of Lur'e equations. For convenience, we sometimes refer to the symmetric matrix X alone as a solution of the Lur'e equations, instead of the triple (X, K, L) . Note that, once having found a solution X , the other matrices K and L can be obtained by a full rank factorization

$$\begin{bmatrix} A^T X + XA + Q & XB + C \\ B^T X + C^T & R \end{bmatrix} = \begin{bmatrix} K^T \\ L^T \end{bmatrix} \begin{bmatrix} K & L \end{bmatrix}. \quad (9)$$

Therefore, if X solves the Lur'e equations (3), the matrices L and K are unique up to the multiplication from the right with some orthogonal matrix of size $p \times p$. We now introduce some further concepts which are used to characterize solvability of the Lur'e equations.

Definition 9. For Lur'e equations (3), the spectral density function, also called Popov function, is defined as

$$\Phi(s) = \begin{bmatrix} (-sI - A)^{-1}B \\ I_m \end{bmatrix}^T \begin{bmatrix} Q & C \\ C^T & R \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I_m \end{bmatrix} \quad (10)$$

Definition 10. For Lur'e equations (3), the associated linear matrix inequality (LMI) is defined as

$$\begin{bmatrix} A^T Y + YA + Q & YB + C \\ B^T Y + C^T & R \end{bmatrix} \geq 0, \quad Y = Y^T. \quad (11)$$

The solution set of the LMI is defined as

$$\mathcal{S}_{LMI} = \{Y \in \mathbb{C}^{n,n} : Y \text{ is symmetric and (11) holds true}\}. \quad (12)$$

The LMI (11) is called feasible if $\mathcal{S}_{LMI} \neq \emptyset$.

It can be readily verified that $Y \in \mathcal{S}_{LMI}$ solves the Lur'e equations, if and only if, it minimizes the rank of (11). We now collect some known equivalent solvability criteria Lur'e equations. In the following we require that the pair (A, B) is stabilizable. Note that this assumption can be further weakened by reducing it to *sign-controllability* [22]. This is not considered here in more detail.

Theorem 11 ([22]). Let the Lur'e equations (3) with associated even matrix pencil $s\mathcal{E} - \mathcal{A}$ as in (4) and spectral density function Φ as in (10) be given. Assume that at least one of the claims

- (i) the pair (A, B) is stabilizable and the pencil $s\mathcal{E} - \mathcal{A}$ as in (4) is regular;
- (ii) the pair (A, B) is controllable;

holds true. Then the following statements are equivalent:

1. There exists a solution (X, K, L) of the Lur'e equations.
2. The LMI (11) is feasible
3. For all $\omega \in \mathbb{R}$ with $i\omega \notin \sigma(A)$ holds $\Phi(i\omega) \geq 0$;
4. In the EKCF of $s\mathcal{E} - \mathcal{A}$, all blocks of type E2 have positive signature and even size, and all blocks of type E3 have negative sign and odd size.
5. In the EKCF of $s\mathcal{E} - \mathcal{A}$, all blocks of type E2 have even size, and all blocks of type E3 have negative sign and odd size.

In particular, solutions of the Lur'e equations fulfill $(X, K, L) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,p} \times \mathbb{C}^{m,p}$ with $p = \text{normalrank } \Phi$.

It is shown in [22] that $m - \text{normalrank } \Phi$ is to the number of blocks of type E4 in an EKCF of $s\mathcal{E} - \mathcal{A}$. In particular, the pencil $s\mathcal{E} - \mathcal{A}$ is regular if and only if Φ has full normal rank.

Now we place particular emphasis on the so-called *maximal solution*.

Theorem 12 ([22]). Let the Lur'e equations (3) be given with stabilizable pair (A, B) . Assume that \mathcal{S}_{LMI} as defined in (12) is non-empty. Then there exists a solution X_+ of the Lur'e equations that is maximal in the sense that for all $Y \in \mathcal{S}_{LMI}$ holds

$$Y \leq X_+.$$

The following result states that the maximal solution can be constructed via the c -stable deflating subspace of the associated even matrix pencil $s\mathcal{E} - \mathcal{A}$.

Theorem 13 ([22]). *Let the Lur'e equations (3) be given with stabilizable pair (A, B) . Assume that \mathcal{S}_{LMI} as defined in (12) is non-empty. Then*

$$\begin{bmatrix} X_+ & 0 \\ I_n & 0 \\ 0 & I_m \end{bmatrix} \quad (13)$$

spans the unique $n + m$ -dimensional semi- c -stable \mathcal{E} – neutral subspace of the pencil (4).

The above theorem states that the maximal solution can be expressed in terms of a special deflating subspace of $s\mathcal{E} - \mathcal{A}$. By means of the EKCF, this space can be constructed from the matrix $U \in \mathbb{C}^{2n+m, 2n+m}$ bringing the pencil $s\mathcal{E} - \mathcal{A}$ into even Kronecker form (7). Considering the partitioning $U = [U_1, \dots, U_k]$ according to the block structure of the EKCF, a matrix $V \in \mathbb{C}^{2n+m, n+m}$ spanning the desired deflating subspace can be constructed by

$$V = [V_1 \quad \dots \quad V_k] \quad \text{for } V_j = U_j Z_j, \quad (14)$$

where

$$Z_j = \begin{cases} [I_{k_j}, 0_{k_j}]^T, & \text{if } \mathcal{D}_j \text{ is of type E1,} \\ [I_{k_j/2}, 0_{k_j/2}]^T, & \text{if } \mathcal{D}_j \text{ is of type E2,} \\ [I_{(k_j+1)/2}, 0_{(k_j-1)/2}]^T, & \text{if } \mathcal{D}_j \text{ is of type E3,} \\ [I_{k_j}, 0_{k_j+1}]^T, & \text{if } \mathcal{D}_j \text{ is of type E4.} \end{cases}$$

In particular, the desired subspace 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.

4. Outline of SDA

The structure-preserving doubling algorithm (SDA) is a matrix iteration which computes two special deflating subspaces of a matrix pencil, one semi-stable and one semi-anti-stable. It was introduced by ANDERSON [32] as an algorithm for the solution of a discrete-time algebraic Riccati equation, and later adapted to many other equations and explained in terms of matrix pencils in several papers by WEN-WEI LIN and others [28, 33, 29, 34]. It is strongly related to the sign function method and to the disc method for matrix pencils [35, 19].

Theorem 14 ([35]). *Let $s\mathcal{E} - \mathcal{A}$ be a regular matrix pencil with $\mathcal{A}, \mathcal{E} \in \mathbb{R}^{N+M, N+M}$, and let $s\mathcal{E}_* - \mathcal{A}_*$ be a regular pencil of the same size with $\mathcal{A}_* \mathcal{E} = \mathcal{E}_* \mathcal{A}$. Then*

1. *the pencil $s\mathcal{E}_* \mathcal{E} - \mathcal{A}_* \mathcal{A}$ is regular and has the same right deflating subspaces as $s\mathcal{E} - \mathcal{A}$*
2. *its eigenvalues are the square of the eigenvalues of the original pencil.*

This result is far from surprising in the case in which \mathcal{E} is invertible: in this case, the eigenvalues and right deflating subspaces of $s\mathcal{E} - \mathcal{A}$ correspond to the eigenvalues and right invariant subspaces of $\mathcal{E}^{-1} \mathcal{A}$, and it is simple to check that the conditions imposed on $\mathcal{E}_*, \mathcal{A}_*$ imply $(\mathcal{E}_* \mathcal{E})^{-1} (\mathcal{A}_* \mathcal{A}) = (\mathcal{E}^{-1} \mathcal{A})^2$. Thus the map $(\mathcal{E}, \mathcal{A}) \mapsto (\mathcal{E}_* \mathcal{E}, \mathcal{A}_* \mathcal{A})$ is a way to extend the concept of squaring to matrix pencils.

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

$$\mathcal{E} = \begin{bmatrix} I_N & -G \\ 0 & F \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} E & 0 \\ -H & I_M \end{bmatrix}, \quad (15)$$

where the block sizes are such that $E \in \mathbb{R}^{N, N}$ and $F \in \mathbb{R}^{M, M}$. Note that a pencil in SSF-I is always regular. When a pencil $s\mathcal{E} - \mathcal{A}$ is in SSF-I, a choice of $s\mathcal{E}_* - \mathcal{A}_*$ satisfying the requirements of Theorem 14 is

$$\mathcal{E}_* = \begin{bmatrix} I_N & -E(I_N - GH)^{-1}G \\ 0 & F(I_M - HG)^{-1} \end{bmatrix}, \quad \mathcal{A}_* = \begin{bmatrix} E(I_N - GH)^{-1} & 0 \\ -F(I_M - HG)^{-1}H & I_N \end{bmatrix},$$

and this choice yields a new pencil $s\mathcal{E}_*\mathcal{E} - \mathcal{A}_*\mathcal{A} = s\tilde{\mathcal{E}} - \tilde{\mathcal{A}}$ which is still in SSF-I, as we have

$$\tilde{\mathcal{E}} = \begin{bmatrix} I_N & -(G + E(I_N - GH)^{-1}GF) \\ 0 & F(I_M - HG)^{-1}F \end{bmatrix}, \quad \tilde{\mathcal{A}} = \begin{bmatrix} E(I_N - GH)^{-1}E & 0 \\ -(H + F(I_M - HG)^{-1}HE) & I_N \end{bmatrix}.$$

The only hypothesis needed here is that $I - GH$ and $I - HG$ are nonsingular. In fact, by the Sherman-Morrison formula, they are either both singular or both nonsingular.

The structured doubling algorithm, outlined as Algorithm 1, consists in repeating this transformation.

input : E, F, G and H defining the pencil $s\mathcal{E} - \mathcal{A}$ with \mathcal{E}, \mathcal{A} as in (15)

output: H_∞, G_∞ so that the subspaces in (16) are respectively the canonical semi-d-stable and semi-d-anti-stable deflating subspaces of the pencil $s\mathcal{E} - \mathcal{A}$

$E_0 = E;$

$F_0 = F;$

$G_0 = G;$

$H_0 = H;$

for $k = 0, \dots, k_{\max}$ **do**

$E_* = E_k(I_N - G_k H_k)^{-1};$

$F_* = F_k(I_M - H_k G_k)^{-1};$

$G_{k+1} = G_k + E_* G_k F_k;$

$H_{k+1} = H_k + F_* H_k E_k;$

$E_{k+1} = E_* E_k;$

$F_{k+1} = F_* F_k;$

end

$H_\infty = H_{k_{\max}};$

$G_\infty = G_{k_{\max}};$

Algorithm 1: SDA-I

For k_{\max} large enough, the algorithm produces two matrices X and Y whose ranges are approximations of, respectively, a semi-d-stable and a semi-d-anti-stable deflating subspaces of the given pencil. Each step of the algorithm costs $\frac{14}{3}(M^3 + N^3) + 6MN(M + N)$ floating point operations. This reduces to $\frac{64}{3}N^3$ in the case $M = N$.

The following convergence result is proved in [29] for the symplectic case and in [34] for several specific matrix equations, but its proof works without changes for our slightly more general case. We call a pencil *weakly d-split* if there exists an r such that:

- the lengths of the Kronecker chains relative to d-stable eigenvalues sum up to $N - r$;
- the lengths of the Kronecker chains relative to d-anti-stable eigenvalues sum up to $M - r$;
- the lengths k_j of the Kronecker chains relative to d-critical eigenvalues (which must sum up to $2r$ if the two previous properties hold) are all even.

In this case, we define the *canonical* d-semi-stable (resp. d-semi-anti-stable) subspace as the invariant subspace spanned by all the Kronecker chains relative to d-stable (resp. d-anti-stable) eigenvalues, plus the first $k_j/2$ vectors from each critical chain.

Theorem 15. *Let the pencil (15) be weakly d-split, and suppose that there are two matrices in the form*

$$\begin{bmatrix} I_N \\ H_\infty \end{bmatrix}, \quad \begin{bmatrix} G_\infty \\ I_M \end{bmatrix}, \quad (16)$$

spanning respectively the canonical d-stable and d-anti-stable deflating subspace. Then for Algorithm 1 it holds that

- $\|E_k\| = O(2^{-k})$,

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

Notice that, when $N = M$, a pencil in SSF-I is symplectic if and only if $E^T = F$, $G = G^T$ and $H = H^T$. In this case, all the pencils generated by the successive steps of SDA are symplectic, i.e., at each step k we have $E_k^T = F_k$, $G_k = G_k^T$, $H_k = H_k^T$. The implementation can be slightly simplified, since there is no need to compute E_{k+1} and F_{k+1} separately, nor to invert both $I_N - G_k H_k$ and $I_M - H_k G_k$, as the second matrix of both pairs is the transposed of the first.

5. A method to compute the SSF-I of a pencil

One can transform a regular pencil into SSF-I easily using the following result.

Theorem 16. *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 SSF-I pencil having the same eigenvalues and right deflating subspaces of the original pencil exists if and only if

$$\begin{bmatrix} \mathcal{E}_1 & \mathcal{A}_2 \end{bmatrix} \tag{17}$$

is nonsingular; in this case, it holds

$$\begin{bmatrix} E & -G \\ -H & F \end{bmatrix} = \begin{bmatrix} \mathcal{E}_1 & \mathcal{A}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{A}_1 & \mathcal{E}_2 \end{bmatrix}. \tag{18}$$

PROOF. We are looking for a 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

$$Q\mathcal{E}_1 = \begin{bmatrix} I \\ 0 \end{bmatrix}, \quad Q\mathcal{A}_2 = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

i.e.,

$$Q \begin{bmatrix} \mathcal{E}_1 & \mathcal{A}_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$

thus Q must be the inverse of the matrix in (17).

On the other hand, 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 (18).

This formula is strictly related to the principal pivot transform (PPT) [36].

We point out an interesting application of Theorem 16, which is not related to the rest of the paper. SDA is often used in the solution of nonsymmetric algebraic Riccati equations (NARE) [37], where it is applied to the Cayley

transform of the matrix

$$\mathcal{H} = \begin{bmatrix} D & -C \\ B & -A \end{bmatrix},$$

with A, B, C, D blocks of suitable size associated with the coefficients of the problem. In pencil form, its Cayley transform given by $s(\mathcal{H} + \gamma I) - (\mathcal{H} - \gamma I)$, thus (18) becomes

$$\begin{bmatrix} E & -G \\ -H & F \end{bmatrix} = \begin{bmatrix} D + \gamma I & -C \\ B & -A - \gamma I \end{bmatrix}^{-1} \begin{bmatrix} D - \gamma I & -C \\ B & -A + \gamma I \end{bmatrix}. \quad (19)$$

This formula is more compact to write and more computationally effective than the one suggested by GUO ET AL. [38]. In fact, their expressions for the starting blocks require the inversion of two $N \times N$ and two $M \times M$ matrices, which are indeed the (1, 1) and (2, 2) blocks of the matrix to be inverted in (19) and their Schur complements. Clearly, two of these four inversions are redundant if we use (19), which requires more or less half of the computational cost with respect to the original formulas in [38].

6. A reduced Lur'e pencil

Let $s\mathcal{E} - \mathcal{A}$ be pencil (4) 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 (4) is regular.

A3 The pair (A, B) is stabilizable.

Note that, according to Theorem 11, assumptions **A1** and **A2** are equivalent to the spectral density function Φ as in 10 having full normal rank and being pointwisely positive semi-definite on the imaginary axis. Note that, for the optimal control problem (1), this means that the optimal control is unique for any $x_0 \in \mathbb{R}^n$. Using Theorem 12, assumption **A3** implies the existence of a maximal solution of the Lur'e equations (3)

By assumption **A2**, there exists some $\gamma \in \mathbb{R}$ such that $\gamma\mathcal{E} - \mathcal{A}$ is invertible. Hence, we may build the SSF-I form of its Cayley transform (assuming $M = n + m, N = n$) which is, by Theorem 16, given by

$$\begin{bmatrix} E & -G \\ -H & F \end{bmatrix} = \begin{bmatrix} 0 & A - \gamma I & B \\ A^T - \gamma I & Q & C \\ B^T & C^T & R \end{bmatrix}^{-1} \begin{bmatrix} 0 & A + \gamma I & B \\ A^T + \gamma I & Q & C \\ B^T & C^T & R \end{bmatrix}. \quad (20)$$

Let now

$$\tilde{A} := \begin{bmatrix} 0 & A - \gamma I \\ A^T - \gamma I & Q \end{bmatrix}, \quad \tilde{B} := \begin{bmatrix} B \\ C \end{bmatrix}, \quad T := \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$

Choosing $\gamma \in \mathbb{R} \setminus \{0\}$ is a way that it is not an eigenvalue of A , the matrix \tilde{A} is invertible. Together with the invertibility of the overall matrix, we also also have the invertibility of its Schur complement

$$\begin{aligned} \mathcal{S}(\gamma) &= R - \begin{bmatrix} B^T & C^T \end{bmatrix} \tilde{A}^{-1} \begin{bmatrix} B \\ C \end{bmatrix} \\ &= R - \begin{bmatrix} B^T & C^T \end{bmatrix} \begin{bmatrix} -(A^T - \gamma I)^{-1} Q (A - \gamma I)^{-1} & (A^T - \gamma I)^{-1} \\ (A^T - \gamma I)^{-1} & 0 \end{bmatrix} \begin{bmatrix} B \\ C \end{bmatrix} \\ &= \begin{bmatrix} (\gamma I - A)^{-1} B \\ I_m \end{bmatrix}^T \begin{bmatrix} Q & -C \\ -C^T & R \end{bmatrix} \begin{bmatrix} (\gamma I - A)^{-1} B \\ I_m \end{bmatrix}. \end{aligned}$$

Note that $\mathcal{S}(\gamma)$ is closely related to the spectral density function Φ in (10).

We can perform the inversion in (20) explicitly with the help of a block LDU factorization. We have

$$\begin{bmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{B}^* & R \end{bmatrix} = \begin{bmatrix} I_{2n} & 0 \\ \widetilde{B}^* \widetilde{A}^{-1} & I_m \end{bmatrix} \begin{bmatrix} \widetilde{A} & 0 \\ 0 & \mathcal{S}(\gamma) \end{bmatrix} \begin{bmatrix} I_{2n} & \widetilde{A}^{-1} \widetilde{B} \\ 0 & I_m \end{bmatrix},$$

thus

$$\begin{aligned} \begin{bmatrix} E & -G \\ -H & F \end{bmatrix} &= \begin{bmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{B}^T & R \end{bmatrix}^{-1} \begin{bmatrix} \widetilde{A} + 2\gamma T & \widetilde{B} \\ \widetilde{B}^T & R \end{bmatrix} \\ &= \begin{bmatrix} I_{2n} & -\widetilde{A}^{-1} \widetilde{B} \\ 0 & I_m \end{bmatrix} \begin{bmatrix} \widetilde{A}^{-1} & 0 \\ 0 & \mathcal{S}(\gamma)^{-1} \end{bmatrix} \begin{bmatrix} I_{2n} & 0 \\ -\widetilde{B}^T \widetilde{A}^{-1} & I_m \end{bmatrix} \begin{bmatrix} \widetilde{A} + 2\gamma T & \widetilde{B} \\ \widetilde{B}^T & R \end{bmatrix} \\ &= \begin{bmatrix} \widehat{A} & 0 \\ -2\gamma \mathcal{S}(\gamma)^{-1} \widetilde{B}^T \widetilde{A}^{-1} T & I_m \end{bmatrix} \end{aligned}$$

where

$$\widehat{A} = I + 2\gamma \widetilde{A}^{-1} T + 2\gamma \widetilde{A}^{-1} \widetilde{B} \mathcal{S}(\gamma)^{-1} \widetilde{B}^T \widetilde{A}^{-1} T. \quad (21)$$

This means that we can identify an additional structure in the blocks used in SSF-I:

$$F = \begin{bmatrix} \widehat{F} & 0 \\ * & I_m \end{bmatrix}, \quad G = \begin{bmatrix} \widehat{G} & 0 \end{bmatrix}, \quad H = \begin{bmatrix} \widehat{H} \\ * \end{bmatrix},$$

where the smaller blocks \widehat{F} , \widehat{G} , \widehat{H} have size $n \times n$. It follows that a special right deflating subspace of this pencil is

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

whose only eigenvalue is 1 with algebraic and geometric multiplicity m , while the deflating subspaces relative to the other eigenvalues are in the form

$$\begin{bmatrix} V \\ * \end{bmatrix},$$

where V has $2n$ rows and is a deflating subspace of the reduced pencil

$$s \begin{bmatrix} I_n & -\widehat{G} \\ 0 & \widehat{F} \end{bmatrix} - \begin{bmatrix} E & 0 \\ -\widehat{H} & I_n \end{bmatrix}. \quad (22)$$

Using (21) and the fact that \widetilde{A} and $\mathcal{S}(\gamma)$ are symmetric, one sees that $\widehat{A}T$ is symmetric, too. This means that $E^T = F$ and $G = G^T$, $H = H^T$, that is, the pencil (22) is symplectic.

The pencil (22) is given by $P^T(s\mathcal{E} - \mathcal{A})P$, where P is the projection on

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

With this characterization, it is easy to derive the KCF of (22) from that of the Cayley transform of (4). We see that $\ker \mathcal{E}$ is the space spanned by the first column of each block of type K2 (as a corollary, we see that there are exactly $m = \dim \ker E$ such blocks). These blocks are transformed into blocks of type K1 with $\lambda = 1$ by the Cayley transform. Thus projecting on their orthogonal complement corresponds to dropping the first row and column from each of the blocks of type K1 corresponding to $\lambda = 1$. In particular, it follows that if the criteria in Theorem 11 hold, then for the KCF of the pencil (22), every block of type K1 corresponding to an eigenvalue on the unit circle has even size. Therefore, the reduced pencil (22) is weakly d-split. By considering which vectors are needed from each vector chain

corresponding to blocks in the KCF to form the subspace in (14) and, furthermore, incorporating assumption **A3**, we get therefore the following result.

Theorem 17. *Let \mathcal{V} be the unique $(n+m)$ -dimensional c -semi-stable \mathcal{E} -neutral deflating subspace of (4). Then, there is a matrix $V_2 \in \mathbb{C}^{n,m}$ such that*

$$\mathcal{V} = \text{span} \begin{bmatrix} V_1 & 0 \\ V_2 & I_m \end{bmatrix},$$

where V_1 spans the canonical d -semi-anti-stable subspace of the pencil (22). Moreover, if $\text{span}(V_1)$ admits a basis in the form

$$\begin{bmatrix} X_+ \\ I_n \end{bmatrix},$$

then X_+ is the maximal solution of the Lur'e equation (3).

In other words, X is the canonical weakly stabilizing solution of the DARE

$$X = EX(I - \widehat{H}X)^{-1}E^T + \widehat{G}. \quad (23)$$

7. Implementation of SDA for Lur'e Equations

Based on the results of the previous sections, we can use the SDA-I algorithm to compute the solution to a Lur'e equation. The resulting algorithm is reported as Algorithm 2.

As we saw in Section 4, 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.

The explicit computation of (a possible choice of) K and L is typically not needed in the applications of the Lur'e equations. If they are needed, they can be computed by a full rank factorization (9).

input : A, B, C, Q, R defining Lur'e equations (3) fulfilling **A1–A3**

output: An approximation of the maximal solution X

Choose a suitable $\gamma > 0$;

Compute

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

Partition

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

Use SDA-I on $E, F = E^T, G, H$ to compute G_∞, H_∞ ;

Return $X = G_\infty$;

Algorithm 2: A structured doubling algorithm for the maximal solution of a Lur'e equation

Algorithm 2 produces a sequence $(G_k)_k$ of approximations of the maximal solution X . Corresponding sequences K_k, L_k of L and K satisfying (3) can be constructed by performing a singular value decomposition

$$\begin{bmatrix} A^T X_k + X_k A + Q & X_k B + C \\ B^T X_k + C^T & R \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}, \quad (24)$$

$$\Sigma_1 \in \mathbb{R}^{m,m}, \quad V_{11}, U_{11}^T \in \mathbb{R}^{n,m}, \quad V_{21}, U_{12}^T \in \mathbb{R}^{m,m}.$$

With approximations $K_k = \Sigma_1^{1/2} V_{11}$, $L_k = \Sigma_1^{1/2} V_{12}$, the *absolute Lur'e residual* can be defined as

$$\left\| \begin{bmatrix} A^T X_k + X_k A + Q & X_k B + C \\ B^T X_k + C^T & R \end{bmatrix} - \begin{bmatrix} K_k^T \\ L_k^T \end{bmatrix} \begin{bmatrix} K_k & L_k \end{bmatrix} \right\|_F. \quad (25)$$

We define the *relative Lur'e residual* as

$$\frac{\left\| \begin{bmatrix} A^T X_k + X_k A + Q & X_k B + C \\ B^T X_k + C^T & R \end{bmatrix} - \begin{bmatrix} K_k^T \\ L_k^T \end{bmatrix} \begin{bmatrix} K_k & L_k \end{bmatrix} \right\|_F}{\left\| \begin{bmatrix} A^T X_k + X_k A + Q & X_k B + C \\ B^T X_k + C^T & R \end{bmatrix} \right\|_F}. \quad (26)$$

Adequate stopping criteria can be designed on the basis of the above defined two quantities.

The accuracy of the computed solution depends also on an appropriate choice of γ . Following [28], we try to minimize the error in the initial values E, F, G, H ; however, notice that, thanks to Theorem 18 and the nicer expression (20) which it provides for the initial values, the criterion exposed there can be substantially simplified and reduced to minimizing the function

$$f(\gamma) = \text{condest} \left(\begin{bmatrix} \mathcal{E}_1 & \mathcal{A}_2 \end{bmatrix} \right)$$

where $\text{condest}(\cdot)$ is the condition number estimate given by `MATLAB`[®]. As an analytical solution is not readily available, following again the strategy in [28], we perform five steps of the golden section search method [39] on $f(\gamma)$ in order to get a reasonably good value of the objective function without devoting too much time to this ancillary computation.

8. Numerical experiments

We have implemented Algorithm 2 (SDA-L) using `MATLAB`[®], and tested it on the following test problems.

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

```
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` [40]. Namely, we took Examples 3 to 6 (the real-world applicative problems) of this 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 high-index problem with $m = 1$, $A = I_n + N_n$, $B = e_n$ (the last column of the $n \times n$ identity matrix), $C = -B$, $R = 0$ and

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

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	5E-15	2E-8	4E-10	5E-6	2E-10
50	5	1E-14	8E-9	2E-8	6E-3	4E-10
500	10	2E-14	8E-9	8E-7	1E-1	★

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	6E-15	5E-2	5E-2	5E-2	9E-10
4	4E-15	6E-7	5E-9	1E-7	5E-9
5	2E-10	3E-7	1E-9	3E-8	1E-9
6	2E-15	6E-12	2E-13	1E-12	2E-13

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 invariant subspace problem in this case is $\varepsilon^{1/(2n+1)}$, for an unstructured perturbation of the input data of the order of the machine precision ϵ [41, section 16.5].

The results of SDA-L are compared to those of a regularization method as the one described in (5), for different values of the regularization parameter ε . After the regularization, the equations are solved using Algorithm 1 after a Cayley transform with the same parameter γ (R+S), or with the matrix sign method with determinant scaling [42, 43] (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 based on a pencil in a form equivalent to (4). In fact, 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 the following 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 the values of the forward error

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

For **P1** and **P2**, for which no analytical solution is available, we computed instead the relative Lur'e residual (26). A star ★ in the data denotes convergence failure.

We see that in all the experiments our solution method obtains a better result than the ones based on regularization.

9. 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 equation without introducing any artificial perturbation.

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	1E-8	1E-3	1E-4	1E-6	1E-4
2	5E-5	3E-2	1E-2	3E-2	★
3	2E-3	1E-1	5E-2	1E+1	★
4	1E-2	4E-1	1E-1	5E-1	★
5	4E-2	1E+0	4E-1	2E+0	★

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 behaviour in presence of eigenvalues on the unit circle, as proved by the convergence results in [29]. Direct methods, such as the symplectic eigensolvers presented in [44], could also be used for the solution of the deflated DARE.

The numerical experiments confirm the effectiveness of our new approach for regular matrix pencils. It is not clear whether the same method can be adapted to work in cases in which the pencil (4) is singular, 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.

10. Acknowledgments

The algorithm exposed in this paper was developed while F. Poloni, was visiting the Numerical Mathematics group of the Technische Universität in Berlin. The hospitality of the research group and of TU Berlin and the support of Scuola Normale Superiore for the visit are gratefully acknowledged.

References

- [1] J. Willems, Least squares stationary optimal control and the algebraic Riccati equation, *IEEE Trans. Automat. Control* 16 (1971) 621–634.
- [2] K. Zhou, J. Doyle, K. Glover, *Robust and Optimal Control*, Prentice-Hall, Princeton, 1996.
- [3] D. Clements, B. Anderson, Matrix inequality solution to linear-quadratic singular control problems, *IEEE Trans. Automatic Control AC-22* (1977) 55–57.
- [4] D. Clements, B. Anderson, Transformational solution of singular linear-quadratic control problems, *IEEE Trans. Automatic Control AC-22* (1977) 57–60.
- [5] D. Clements, B. Anderson, Singular optimal control: the linear-quadratic problem, volume 5 of *Lecture Notes in Control and Information Sciences*, Springer-Verlag, Berlin-New York, 1978.
- [6] V. Yakubovich, Singular problem in optimal control of linear stationary system with quadratic functional, *Siberian Math. J.* 26 (1985) 148–158.
- [7] A. Lur'e, *Certain Nonlinear Problems in the Theory of Automatic Control*, Gostekhzizat, Moscow, Leningrad, 1951. Translated into English, H.M. Stationery, 1957.
- [8] N. Barabanov, Kalman-Yakubovich lemma in general finite-dimensional case, *Int. J. Robust Nonlinear Control* 17 (2007) 369–386.
- [9] B. Anderson, Algebraic description of bounded real matrices, *Electronics Letters* 2 (1966) 464–465.
- [10] B. Anderson, R. Newcomb, Impedance synthesis via state-space techniques, *Proc. IEE* 115 (1968) 928–936.
- [11] B. Anderson, S. Vongpanitlerd, *Network Analysis and Synthesis*, Prentice Hall, Englewood Cliffs, NJ, 1973.
- [12] J. Willems, Dissipative dynamical systems. II: Linear systems with quadratic supply rates, *Arch. Ration. Mech. Anal.* 45 (1972) 352–393.
- [13] X. Chen, J. Wen, Positive realness preserving model reduction with \mathcal{H}_∞ norm error bounds, *Systems Control Lett.* 54 (2005) 361–374.
- [14] S. Gugercin, A. Antoulas, A survey of model reduction by balanced truncation and some new results, *Internat. J. Control* 77 (2004) 748–766.
- [15] P. Odenacker, E. Jonckheere, A contraction mapping preserving balanced reduction scheme and its infinity norm error bounds, *IEEE Trans. Circuits Syst. I Regul. Pap.* 35 (1988) 184–189.
- [16] J. Phillips, L. Daniel, L. Silveira, Guaranteed passive balancing transformations for model order reduction, *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.* 22 (2003) 1027–1041.
- [17] T. Reis, T. Stykel, Passive and bounded real balancing for model reduction of descriptor systems, *Internat. J. Control* 83 (2010) 74–88.
- [18] P. Lancaster, L. Rodman, *Algebraic Riccati equations*, Clarendon Press, Oxford, 1995.
- [19] P. Benner, *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.
- [20] T. Reis, T. Stykel, Pabtec: Balanced truncation for electrical circuits, *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.* 29 (2010) 1354–1367.
- [21] D. Clements, B. Anderson, A. Laub, L. Matson, Spectral factorization with imaginary-axis zeros, *Linear Algebra Appl.* 250 (1997) 225–252.
- [22] T. Reis, Lur'e equations and even matrix pencils, *Linear Algebra Appl.* 434 (2010) 152–173.
- [23] F. Gantmacher, *Theory of Matrices*, Chelsea Publishing Company, New York, 1959.
- [24] D. Jacobson, J. Speyer, Necessary and sufficient conditions for optimality for singular control problems; a transformation approach, *J. Math. Anal. Appl.* 33 (1971) 163–187.
- [25] H. Weiss, Q. Wang, J. Speyer, System characterization of positive real conditions, *IEEE Trans. Automat. Control* 39 (1994) 540–544.
- [26] D. Jacobson, J. Speyer, Necessary and sufficient conditions for singular control problems: a limit approach, *J. Math. Anal. Appl.* 34 (1971) 239–266.
- [27] H. Trentelman, Families of linear-quadratic problems: Continuity properties, *IEEE Trans. Automatic Control AC-32* (1987) 323–329.
- [28] E. K.-W. Chu, H.-Y. Fan, W.-W. Lin, A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations, *Linear Algebra Appl.* 396 (2005) 55–80.

- [29] T.-M. Huang, W.-W. Lin, Structured doubling algorithms for weakly stabilizing Hermitian solutions of algebraic Riccati equations, *Linear Algebra Appl.* 430 (2009) 1452–1478.
- [30] R. Thompson, The characteristic polynomial of a principal subpencil of a Hermitian matrix pencil, *Linear Algebra Appl.* 14 (1976) 135–177.
- [31] R. Thompson, Pencils of complex and real symmetric and skew matrices, *Linear Algebra Appl.* 147 (1991) 323–371.
- [32] B. Anderson, Second-order convergent algorithms for the steady-state Riccati equation, *Internat. J. Control* 28 (1978) 295–306.
- [33] W.-W. Lin, S.-F. Xu, Convergence analysis of structure-preserving doubling algorithms for Riccati-type matrix equations, *SIAM J. Matrix Anal. Appl.* 28 (2006) 26–39 (electronic).
- [34] C.-Y. Chiang, E. K.-W. Chu, C.-H. Guo, T.-M. Huang, W.-W. Lin, 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) 227–247.
- [35] Z. Bai, J. Demmel, M. Gu, An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems, *Numer. Math.* 76 (1997) 279–308.
- [36] M. Tsatsomeros, Principal pivot transforms: properties and applications, *Linear Algebra Appl.* 307 (2000) 151–165.
- [37] C.-H. Guo, A. Laub, On a Newton-like method for solving algebraic Riccati equations, *SIAM J. Matrix Anal. Appl.* 21 (1999) 694–698 (electronic).
- [38] X.-X. Guo, W.-W. Lin, S. Xu, A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation, *Numer. Math.* 103 (2006) 393–412.
- [39] D. Luenberger, Y. Ye, *Linear and nonlinear programming*, International Series in Operations Research & Management Science, 116, Springer, New York, third edition, 2008.
- [40] P. Benner, A. Laub, V. Mehrmann, 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.
- [41] I. Gohberg, P. Lancaster, L. Rodman, Invariant subspaces of matrices with applications, volume 51 of *Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Reprint of the 1986 original.
- [42] V. Mehrmann, The autonomous linear quadratic control problem, volume 163 of *Lecture Notes in Control and Information Sciences*, Springer-Verlag, Berlin, 1991. Theory and numerical solution.
- [43] N. Higham, *Functions of matrices*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. Theory and computation.
- [44] H. Fassbender, *Symplectic methods for the symplectic eigenproblem*, Kluwer Academic/Plenum Publishers, New York, 2000.