The extended symplectic pencil and the finite-horizon LQ problem with two-sided boundary conditions

Augusto Ferrante and Lorenzo Ntogramatzidis*†‡

September 11, 2013

Abstract

This note introduces a new approach to the solution of a very general class of finite-horizon optimal control problems for discrete-time systems. This approach provides a parametric expression for the optimal control sequences, as well as the corresponding optimal state trajectories, by exploiting a new decomposition of the so-called extended symplectic pencil. This decomposition provides an original strategy for a more direct solution of the problem with no need of the system-theoretic hypotheses (including regularity of the symplectic pencil) that have always been assumed in the literature so far.

1 Introduction

This paper focuses on a very general class of finite-horizon linear-quadratic (LQ) problems with affine constraints at the end-points. These problems are not just important *per se*. In fairly recent literature it has been shown that LQ problems are becoming increasingly useful as building blocks to solve complex optimisation problems, broken down into two or more LQ subproblems, each one with constraints at the end-points. In particular, finite-horizon LQ problems with constraints at the end-points [13, 14] are intermediate steps in the solution of

*A. Ferrante is with the Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo, 6/B – I-35131 Padova, Italy, E-mail: `augusto@dei.unipd.it`.

†L. Ntogramatzidis is with Department of Mathematics and Statistics, Curtin University, Perth WA 6845, Australia. E-mail: `L.Ntogramatzidis@curtin.edu.au`.

‡Manuscript received April 19, 2005; revised January 11, 2007.

H_2 receding-horizon problems and the minimisation of regulation transients in switching linear plants.

The aim of this paper is to present a method to solve the most general class of finite-horizon LQ optimal control problems in the discrete time with positive semi-definite cost index and affine constraints at the end-points. The proposed solution is based on a procedure for the parameterisation of the set of trajectories generated by the so-called extended symplectic difference equation (ESDE). The idea of solving finite-horizon LQ problems by exploiting expressions of the trajectories generated by the ESDE originated in the papers [3] and [4]. In the past literature, however, the problem solution was always essentially based on two "opposite" solutions of the associated discrete algebraic Riccati equation (with some extra tricks to deal with the case when the closed-loop matrix is singular and hence no pairs of completely opposite solutions exist). This point of view always requires some controllability-type assumption and the extended symplectic pencil [17] to be regular and devoid of generalised eigenvalues on the unit circle. The goal of this paper is to propose a new point of view aimed at a more direct and simple solution to this problem, without requiring system-theoretic assumptions. The technique presented here only requires a solution of the so-called generalised discrete-time algebraic Riccati equation, which may exist even when the symplectic pencil is not regular (in which case the standard discrete algebraic Riccati equation does not admit solutions, let alone pairs of "opposite" solutions). Such solution is used to derive a decomposition of the extended symplectic pencil that yields a natural parameterisation of the solutions of the symplectic difference equation. Thus, while for practical purposes our paper simply provides a generalisation (yet in three different directions) with respect to the existing literature, its different point of view casts a new light on the theoretical comprehension of this problem and on its connections to the classical cornerstones of linear systems theory.

For a better description of the features and the generality of our framework, we illustrate all our results in a running example in which the underlying system is not modulus controllable, and the extended symplectic pencil is not regular (so that the methods in previous literature cannot be used).

2 Statement of the problem

Consider the linear time-invariant discrete-time system governed by the difference equation

$$x(t+1) = A\,x(t) + B\,u(t), \tag{1}$$

where, for all $t \in \mathbb{N}$, $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the control input, $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$. Let $N \in \mathbb{N} \setminus \{0\}$ be the length of the time horizon. Let $V_0, V_N \in \mathbb{R}^{q \times n}$

and $v\in\mathbb{R}^q$; consider

$$V_0\,x(0)+V_N\,x(N)=v, \tag{2}$$

which represents a two-point boundary-value affine constraint on the states at the end-points. With no loss of generality, we can consider $V \triangleq [\,V_0 \;\; V_N\,]$ to be of full row-rank. In the case where $q=0$, the matrices V_0, V_N, V and the vector v are considered to be void.

Let $\Pi = \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} = \Pi^\top \geq 0$ be a square $(n+m)$-dimensional matrix with $Q\in\mathbb{R}^{n\times n}$, $S\in\mathbb{R}^{n\times m}$ and $R\in\mathbb{R}^{m\times m}$ (note that we do not assume the non-singularity of R). We denote by Σ the Popov triple (A,B,Π). Finally, let $H = \begin{bmatrix} H_1 & H_2 \\ H_2^\top & H_3 \end{bmatrix} = H^\top \geq 0$ with $H_1, H_2, H_3\in\mathbb{R}^{n\times n}$ and $h_0, h_N\in\mathbb{R}^n$.

Problem 1 *Find $u(t)$, $t\in\{0,\ldots,N-1\}$ and $x(t)$, $t\in\{0,\ldots,N\}$, minimising*

$$\begin{aligned} J(x,u) \triangleq\; & \sum_{t=0}^{N-1} \big[\, x^\top(t) \;\;\; u^\top(t)\,\big]\,\Pi \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \\ & + \big[\, x^\top(0)-h_0^\top \;\;\; x^\top(N)-h_N^\top \,\big]\, H \begin{bmatrix} x(0)-h_0 \\ x(N)-h_N \end{bmatrix}, \end{aligned} \tag{3}$$

under the constraints (1-2).

As discussed in [4], the formulation of Problem 1 is very general, since the cost index in (3) involves the most general type of positive semidefinite quadratic penalisation on the extreme states, and (2) represents the most general affine constraint on these states. As particular cases of Problem 1 we have **1)** the standard case where $x(0)$ is assigned and $x(N)$ is weighted in (3); **2)** the fixed end-point case, where the states at the end-points are sharply assigned; **3)** the point-to-point case, where the extreme values of an output $y(t)=C\,x(t)$ are constrained to be equal to two assigned vectors y_0 and y_N, respectively. Further non-standard LQ problems that can be useful in practice are particular cases of Problem 1: consider for example an LQ problem in which the states at the end-points $x(0)$ and $x(N)$ are not assigned, but they are constrained to be equal, i.e., $x(0)=x(N)$. This case can be obtained by Problem 1 by setting $V_0=I_n$, $V_N=-I_n$ and $v=0$.

Lemma 1 **[4, Lemma 1]** *If $u(t)$ and $x(t)$ are optimal for Problem 1, then $\lambda(t)\in\mathbb{R}^n$, $t\in\{0,\ldots,N\}$ and $\eta\in\mathbb{R}^s$ exist such that $x(t)$, $\lambda(t)$, $u(t)$ and η satisfy the set of* *equations*

$$x(t+1) = Ax(t) + Bu(t) \qquad t \in \{0,\ldots,N-1\}, \tag{4}$$

$$V \begin{bmatrix} x(0) \\ x(N) \end{bmatrix} = v, \tag{5}$$

$$\lambda(t) = Qx(t) + A^\top \lambda(t+1) + Su(t) \qquad t \in \{0,\ldots,N-1\}, \tag{6}$$

$$\begin{bmatrix} -\lambda(0) \\ \lambda(N) \end{bmatrix} = H \begin{bmatrix} x(0)-h_0 \\ x(N)-h_N \end{bmatrix} + V^\top \eta, \tag{7}$$

$$0 = S^\top x(t) + B^\top \lambda(t+1) + Ru(t) \qquad t \in \{0,\ldots,N-1\}. \tag{8}$$

Conversely, if equations (4-8) admit solutions $x(t)$, $u(t)$, $\lambda(t)$, η, then $x(t)$, $u(t)$ minimise $J(x,u)$ subject to the constraints (1-2).

3 The generalised Riccati equation and the extended symplectic system

Since in the present setting we are not assuming that R is positive definite, (8) cannot be solved in $u(t)$ to obtain a set of $2n$ equations in $x(t)$ and $\lambda(t)$. A convenient form in which (4), (6) and (8) can be written, that does not require inversion of R, is the descriptor form

$$F\,p(t+1) = G\,p(t) \qquad t \in \{0,\ldots,N-1\}, \tag{9}$$

where

$$F \triangleq \begin{bmatrix} I_n & O & O \\ O & -A^\top & O \\ O & -B^\top & O \end{bmatrix}, \quad G \triangleq \begin{bmatrix} A & O & B \\ Q & -I_n & S \\ S^\top & O & R \end{bmatrix}, \quad p(t) \triangleq \begin{bmatrix} x(t) \\ \lambda(t) \\ u(t) \end{bmatrix}.$$

Notice that there is a small issue in the equivalence between equations (4), (6) and (8) and equation (9). In fact, $u(N)$ does not appear in (4), (6) and (8). Notice, however, that when $u(N)$ appears in (9) it is multiplied by 0, hence its value is irrelevant. Therefore, we can say that equations (4), (6) and (8) and equation (9) are equivalent, modulo the (arbitrary) value of $u(N)$. The matrix pencil $G - zF$ is known as the *extended symplectic pencil*, [11, 9], herein denoted concisely by ESP(Σ). In this paper we do not make the assumption of regularity of this pencil.

We will show how to obtain a decomposition of ESP(Σ) that can be used to solve Problem 1 by exploiting the solutions of the following constrained matrix equation

$$X = A^\top XA - (A^\top XB + S)(R + B^\top XB)^\dagger (B^\top XA + S^\top) + Q, \tag{10}$$

$$\ker(R + B^\top XB) \subseteq \ker(A^\top XB + S), \tag{11}$$

where (10) has been obtained from the standard discrete algebraic Riccati equation (DARE) by replacing the inverse with the Moore-Penrose pseudo-inverse. Eq. (10) is known in the literature as the *generalised discrete-time algebraic Riccati equation* GDARE(Σ), [15, 8]. GDARE(Σ) with the additional constraint given by (11) is sometimes referred to as *constrained generalised discrete-time algebraic Riccati equation* CGDARE(Σ). Clearly (10) constitutes a generalisation of the classic DARE(Σ), in the sense that any solution of DARE(Σ) is also a solution of GDARE(Σ) – and therefore also of CGDARE(Σ) – but the *vice-versa* is not true in general. Results on the existence of solutions of GDARE(Σ) in terms of deflating subspaces of the extended symplectic pencil are given in [8] and [9]. We now introduce a standing assumption.

Assumption 3.1 *Assume that CGDARE(Σ) has solutions.*

Notice that Assumption 3.1 is generically satisfied. The situations in which CGDARE(Σ) does not admit solutions happen to be extremely pathological. Indeed, to the best of the authors' knowledge, no necessary and sufficient existence conditions expressed in terms of the problem data are available for CGDARE(Σ). There are, however, very weak sufficient conditions (see e.g. modulus controllability, [4]) that guarantee existence of solutions of DARE(Σ) – and therefore also of CGDARE(Σ). On the other hand, CGDARE(Σ) generalises DARE(Σ), and may admit solutions even when DARE(Σ) does not. Thus, even in cases in which the aforementioned weak system-theoretic conditions are not satisfied, CGDARE(Σ) may still have solutions. Such solutions can be computed via a reduction to a reduced-order DARE(Σ), see the MATLAB® routine `rdare.m` in [2], see also [7]. We now introduce some notation that will be used throughout the paper. First, to any matrix $X = X^\top \in \mathbb{R}^{n\times n}$ we associate the following matrices:

$$S_X \triangleq A^\top X B + S, \quad R_X \triangleq R + B^\top X B, \quad G_X \triangleq I_m - R_X^\dagger R_X, \tag{12}$$

$$K_X \triangleq R_X^\dagger S_X^\top, \quad A_X \triangleq A - B K_X. \tag{13}$$

The term $R_X^\dagger R_X$ is the orthogonal projector that projects onto $\operatorname{range} R_X^\dagger = \operatorname{range} R_X$ so that G_X is the orthogonal projector that projects onto $\ker R_X$. Hence, $\ker R_X = \operatorname{range} G_X$.

Example 3.1 The following Popov triple is used as a running example throughout the paper:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad S = R = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The extended symplectic pencil in this case is not regular. As such, DARE(Σ) does not admit solutions. On the other hand, in this case CGDARE(Σ) admits the solution $X = \mathrm{diag}\{0,1\}$, that can be computed by resorting to the algorithm proposed in [2]. In this case, $R_X = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, and $A_X = \mathrm{diag}\{1,0\}$. Observe that the spectrum of A_X is not unmixed, see e.g. [4]. $\square$

The following result adapts [6, Lemma 2.5] to the case when $G - zF$ may be singular.

Lemma 2 *Let $X = X^\top$ be a solution of CGDARE(Σ). Then, two invertible matrices $U_X, V_X \in \mathbb{R}^{2n+m}$ exist such that*

$$U_X\,(G - zF)\,V_X = \begin{bmatrix} A_X - zI_n & O & B \\ O & I_n - zA_X^\top & O \\ O & -zB^\top & R_X \end{bmatrix}. \tag{14}$$

Proof: By direct computation we find

$$U_X(G-zF)V_X = \begin{bmatrix} A_X - zI_n & O & B \\ \Xi_{21} & I_n - zA_X^\top & \Xi_{23} \\ \Xi_{23}^\top & -zB^\top & R_X \end{bmatrix}$$

with

$$U_X \triangleq \begin{bmatrix} I_n & O & O \\ A_X^\top X & I_n & -K_X^\top \\ B^\top X & O & I_m \end{bmatrix} \quad \text{and} \quad V_X \triangleq \begin{bmatrix} I_n & O & O \\ X & -I_n & O \\ -K_X & O & I_m \end{bmatrix}.$$

The term Ξ_{21} is given by

$$\begin{aligned}\Xi_{21} &= A_X^\top X A - A_X^\top X B K_X + Q - X - S K_X - K_X^\top S^\top + K_X^\top R K_X \\ &\quad - z\,(A^\top X - A_X^\top X + K_X^\top B^\top X).\end{aligned}$$

The term multiplying z is zero since $A_X = A - BK_X$. Moreover, since GDARE(Σ) can be written as $X = A^\top X A - S_X K_X + Q$ we find $\Xi_{21} = K_X^\top (R_X K_X - S_X^\top) = S_X R_X^\dagger R_X R_X^\dagger S_X^\top - S_X R_X^\dagger S_X^\top = 0$. Finally, $\Xi_{23} = A^\top X B - zXB - K_X^\top B^\top X B + S + zXB - K_X^\top R = S_X G_X$. In view of (11), we have $S_X G_X = 0$, so that (14) holds. $\blacksquare$

Remark 1 It is known that the dynamics associated with a matrix pencil is governed by its generalised eigenvalues. [1] If X is a solution of CGDARE(Σ), from (14) we have

$$\det(G - zF) = (-1)^n \cdot \det(A_X - zI_n) \cdot \det(I_n - zA_X^\top) \cdot \det R_X. \tag{15}$$

[1] Recall that a generalised eigenvalue of a matrix pencil $G - zF$ is a value of $z \in \mathbb{C}$ for which the rank of the matrix pencil $G - zF$ is lower than its normal rank.

When R_X is non-singular (i.e. X is also a solution of DARE(Σ)), the generalised eigenvalues of $G-zF$ are immediately seen to be given by the eigenvalues of A_X, the reciprocal of the non-zero eigenvalues of A_X, and a generalised eigenvalue at infinity whose algebraic multiplicity is equal to m plus the algebraic multiplicity of the eigenvalue of A_X at the origin. When the matrix R_X is singular, the computation of the generalised eigenvalues of $G-zF$ is much more complex. Indeed, in such case (15) still holds but provides no information since $\det R_X = 0$. We show this fact with a simple example.

Example 3.2 Consider Example 3.1. Matrix $X = \mathrm{diag}\{0,1\}$ is a solution of CGDARE(Σ), and the corresponding closed-loop matrix is $A_X = \mathrm{diag}\{1,0\}$. From Lemma 2 we find

$$U_X\,(G-zF)\,V_X = \left[\begin{array}{cc|cc|cc} 1-z & 0 & 0 & 0 & 2 & 0 \\ 0 & -z & 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 1-z & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -2z & -z & 1 & 1 \\ 0 & 0 & 0 & -z & 1 & 1 \end{array}\right],$$

whose normal rank (which coincides with that of $G-zF$) is easily seen to be equal to 5. The eigenvalues of A_X are 0 and 1. However, it is not true that $z=1$ is a generalised eigenvalue of $G-zF$. In fact, a direct check shows that the rank of $G-F$ is equal to 5.[2] $\square$

From these considerations, it turns out that when R_X is singular, the computation of the eigenstructure of the pencil $G-zF$ is more difficult, and requires a different machinery. This machinery hinges on a decomposition of the matrix pencil $G-zF$ for which we need to introduce the following notation. Consider a change of coordinates in the input space $\mathbb{R}^m$ induced by the $m \times m$ orthogonal matrix $T = [\,T_1\;\; T_2\,]$ where $\mathrm{range}\,T_1 = \mathrm{range}\,R_X$ and $\mathrm{range}\,T_2 = \mathrm{range}\,G_X = \ker R_X$. From [5, Theorems 4.3-4.4], T is independent of the solution X of CGDARE(Σ). Thus $T^\top R_X\,T = \mathrm{diag}\{R_{X,0}, O\}$, where $R_{X,0}$ is invertible. Its dimension is denoted by m_1. Consider the block matrix $\hat{T} \triangleq \mathrm{diag}(I_n, I_n, T)$. Defining the matrices $B_1 \triangleq B\,T_1$ and $B_2 \triangleq B\,T_2$ we get

$$\hat{T}^\top (U_X(G-zF)V_X)\,\hat{T} = \begin{bmatrix} A_X - zI_n & O & B_1 & B_2 \\ O & I_n - zA_X^\top & O & O \\ O & -zB_1^\top & R_{X,0} & O \\ O & -zB_2^\top & O & O \end{bmatrix}. \tag{16}$$

[2] We warn that the routine `eig.m` of the software MATLAB® (version 7.11.0.584(R2010b)) in this case fails to provide the right answer. It indeed returns 1 as a generalised eigenvalue of the pencil $G-zF$.

From $\ker R_X = \operatorname{range} G_X$, we obtain $\operatorname{range} B_2 = \operatorname{range}(B G_X)$. Matrix B_1 has m_1 columns. Let $m_2 \triangleq m - m_1$ be the number of columns of B_2. Let us take $U = [U_1\ U_2]$ such that U_1 spans the reachable subspace associated with the pair (A_X, B_2), denoted by $\mathscr{R}_X$, and U_2 is such that U is invertible. We have

$$\begin{aligned} U^{-1} A_X U &= \begin{bmatrix} A_{X,11} & A_{X,12} \\ O & A_{X,22} \end{bmatrix}, \quad U^{-1} B_2 = \begin{bmatrix} B_{21} \\ O \end{bmatrix}, \\ U^{-1} B_1 &= \begin{bmatrix} B_{11} \\ B_{12} \end{bmatrix}. \end{aligned} \tag{17}$$

Now, we are ready to state the main result of the paper.

Theorem 1 *Let Assumption 3.1 hold. Two invertible matrices $\hat{U}_X$ and $\hat{V}_X$ exist such that*

$$\hat{U}_X (G - zF) \hat{V}_X = \left[\begin{array}{cc|c|ccc} A_{X,11} - zI_r & B_{21} & O & A_{X,12} & O & B_{11} \\ \hline O & O & I_r - zA_{X,11}^\top & O & O & O \\ O & O & -zB_{21}^\top & O & O & O \\ \hline O & O & O & A_{X,22} - zI_{n-r} & O & B_{12} \\ O & O & -zA_{X,12}^\top & O & I_{n-r} - zA_{X,22}^\top & O \\ O & O & -zB_{11}^\top & O & -z\,B_{12}^\top & R_{X,0} \end{array}\right], \tag{18}$$

where the pair $(A_{X,11}, B_{21})$ is reachable and $R_{X,0}$ is invertible. Moreover, the matrix pencil $P_1(z) \triangleq \begin{bmatrix} A_{X,22} - zI_{n-r} & O & B_{12} \\ O & I_{n-r} - zA_{X,22}^\top & O \\ O & -zB_{12}^\top & R_{X,0} \end{bmatrix}$ in (18) is regular, and the generalised eigenvalues of the pencil $G - zF$ are the generalised eigenvalues of $P_1(z)$.

The proof of Theorem 1 can be found in the Appendix. The decomposition introduced in Theorem 1 essentially isolates the regular part $P_1(z)$ of the pencil $G - zF$. A consequence of this fact is that, unlike the regular case, not all the eigenvalues of A_X appear as generalised eigenvalues of ESP(Σ). Indeed, from (18) we have the following

Corollary 1 *The finite generalised eigenvalues of $G - zF$ are the uncontrollable eigenvalues of the pair (A_X, B_2) plus the reciprocals of those eigenvalues that are not zero.*

Example 3.3 Consider Example 3.1. Using the solution $X = \operatorname{diag}\{0, 1\}$ of CGDARE(Σ), the null-space and image of R_X are respectively spanned by the vectors $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. By taking $T = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ we obtain $T^\top R_X T = \operatorname{diag}\{4, 0\}$. Hence, in this case $m_1 = m_2 = 1$. We partition BT as $BT = \begin{bmatrix} 2 & -2 \\ 2 & 0 \end{bmatrix}$, so that $B_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $B_2 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$. The normal rank of ESP(Σ) is equal to $2n + m_1 = 5$. The generalised eigenvalues of $G - zF$ are given by the uncontrollable eigenvalues of the pair $(A_X, B_2) = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix} \right)$ plus their reciprocals. Therefore, ESP(Σ) has a generalised eigenvalue at the origin. Since $A_{X,22} = 0$ and $B_{12} = 2$, it also has an eigenvalue at infinity with multiplicities equal to the multiplicities of the zero eigenvalue of $\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$. By writing this pencil in the form given by (18), we get

$$P(z) = \left[\begin{array}{cc|c|ccc} 1-z & -2 & 0 & 0 & 0 & 2 \\ \hline 0 & 0 & 1-z & 0 & 0 & 0 \\ 0 & 0 & 2z & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -z & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2z & 0 & -2z & 4 \end{array}\right],$$

from which we see that zero is indeed the only finite generalised eigenvalue of ESP(Σ). □

4 Solution of the LQ problem

We now consider the problem in the basis constructed in the previous section. Let $\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = U^{-1} x(t)$ be the coordinates of the state in this basis, partitioned conformably with U. Similarly, let $\begin{bmatrix} \lambda_1(t) \\ \lambda_2(t) \end{bmatrix} = U^\top \lambda(t)$ and $T^\top u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$. In this section, we show that in this basis the problem can be easily solved in closed form. More precisely, we first parameterise the solutions of (9) in terms of $x_1(0)$, $x_2(0)$, $x_2(N)$ and $\lambda_2(N)$. Then we parameterise the optimal values of $x_1(0)$, $x_2(0)$, $x_2(N)$ and $\lambda_2(N)$ by imposing the boundary conditions.

In the new bases, equations (9) can be written for $t \in \{0, \ldots, N-1\}$ as

$$\begin{aligned} x_1(t+1) &= A_{X,11} x_1(t) + B_{21} u_1(t) + A_{X,12} x_2(t) + B_{11} u_2(t), && (19) \\ \lambda_1(t) &= A_{X,11}^\top \lambda_1(t+1), && (20) \\ 0 &= -B_{21}^\top \lambda_1(t+1), && (21) \\ x_2(t+1) &= A_{X,22} x_2(t) + B_{12} u_2(t), && (22) \\ \lambda_2(t) &= A_{X,22}^\top \lambda_2(t+1) + A_{X,12}^\top \lambda_1(t+1), && (23) \\ u_2(t) &= R_{X,0}^{-1} B_{12}^\top \lambda_2(t+1) + R_{X,0}^{-1} B_{11}^\top \lambda_1(t+1). && (24) \end{aligned}$$

Since by construction the pair $(A_{X,11}, B_{21})$ is reachable, $\ker \begin{bmatrix} A_{X,11}^\top \\ B_{21}^\top \end{bmatrix} = \{0\}$, which means (20-21) yield $\lambda_1(t) = 0$ for all $t \in \{0, \dots, N-1\}$. Thus, (23-24) can be simplified as

$$\lambda_2(t) = A_{X,22}^\top \lambda_2(t+1), \quad \text{and} \quad u_2(t) = R_{X,0}^{-1} B_{12}^\top \lambda_2(t+1). \tag{25}$$

It is clear at this point that we can parameterise all the trajectories generated by the difference equations (22) and (25) in terms of $x_2(0)$ and $\lambda_2(N)$. Indeed, the first of (25) leads to

$$\lambda_2(t) = (A_{X,22}^\top)^{N-t} \lambda_2(N) \qquad \forall t \in \{0, \dots, N\}. \tag{26}$$

This expression can be plugged into $u_2(t)$, and gives

$$u_2(t) = R_{X,0}^{-1} B_{12}^\top (A_{X,22}^\top)^{N-t-1} \lambda_2(N). \tag{27}$$

Plugging (26) and (27) into (22) gives

$$\begin{aligned} x_2(t) &= A_{X,22}^t x_2(0) \\ &\quad + \sum_{j=0}^{t-1} A_{X,22}^{t-j-1} B_{12} R_{X,0}^{-1} B_{12}^\top (A_{X,22}^\top)^{N-j-1} \lambda_2(N). \end{aligned} \tag{28}$$

It is worth observing that

$$x_2(N) = A_{X,22}^N x_2(0) + P \lambda_2(N), \tag{29}$$

where

$$P \triangleq \sum_{j=0}^{N-1} A_{X,22}^{N-j-1} B_{12} R_{X,0}^{-1} B_{12}^\top (A_{X,22}^\top)^{N-j-1}.$$

It is easy to see that matrix P can be re-written as $P = \sum_{j=0}^{N-1} A_{X,22}^j B_{12} R_{X,0}^{-1} B_{12}^\top (A_{X,22}^\top)^j$. Therefore, P satisfies the discrete Lyapunov equation

$$P = A_{X,22} P A_{X,22}^\top - A_{X,22}^N B_{12} R_{X,0}^{-1} B_{12}^\top (A_{X,22}^\top)^N + B_{12} R_{X,0}^{-1} B_{12}^\top.$$

If $A_{X,22}$ has unmixed spectrum, this equation can be used to determine P instead of computing the sum in (29). At this point we can solve (19), which can be written as

$$x_1(t+1) \;=\; A_{X,11} x_1(t) + B_{21} u_1(t) + \xi(t), \tag{30}$$

where $\xi(t) = A_{X,12}\,x_2(t) + B_{11}\,u_2(t)$. Using (28) and (27) we find

$$\xi(t) = A_{X,12}A_{X,22}^{t}x_2(0) + \Big(B_{11}R_{X,0}^{-1}B_{12}^{\top}(A_{X,22}^{\top})^{N-t-1} + A_{X,12}\sum_{j=0}^{t-1}A_{X,22}^{t-j-1}B_{12}R_{X,0}^{-1}B_{12}^{\top}(A_{X,22}^{\top})^{N-j-1}\Big)\lambda_2(N).$$

Let $R_1 = [B_{21} \mid A_{X,11}\,B_{21} \mid A_{X,11}^{2}\,B_{21} \mid \cdots \mid A_{X,11}^{N-1}B_{21}]$ and $R_2 = [I \mid A_{X,11} \mid A_{X,11}^{2} \mid \cdots \mid A_{X,11}^{N-1}]$. Then, we can write $x_1(N) = A_{X,11}^{N}x_1(0) + R_2\,\Xi + R_1\,U_1$ where $\Xi \triangleq \begin{bmatrix} \xi(N-1) \\ \vdots \\ \xi(0) \end{bmatrix}$ and $U_1 \triangleq \begin{bmatrix} u_1(N-1) \\ \vdots \\ u_1(0) \end{bmatrix}$. We assume that N is greater than the controllability index of the pair $(A_{X,11}, B_{21})$. All the solutions of this equation are parameterised by

$$U_1 = R_1^{\dagger}\left(x_1(N) - A_{X,11}^{N}x_1(0) - R_2\,\Xi\right) + (I - R_1^{\dagger}R_1)\,v_1, \tag{31}$$

where v_1 is arbitrary.

4.1 Boundary conditions

In the new basis, the state, co-state and transversality equations can be written again as in (4), (6) and (8), where A, B, Q, S, V, H, h_0 and h_N are replaced by $\tilde{A} = U^{-1}AU$, $\tilde{B} = U^{-1}B$, $\tilde{Q} = U^{\top}QU$, $\tilde{S} = U^{\top}S$, $\tilde{V} = V\begin{bmatrix} U & O \\ O & U \end{bmatrix}$, $\tilde{H} = \begin{bmatrix} U & O \\ O & U \end{bmatrix}^{\top} H \begin{bmatrix} U & O \\ O & U \end{bmatrix}$, $\tilde{h}_0 = U^{-1}h_0$ and $\tilde{h}_N = U^{-1}h_N$, respectively. Now, let us consider the boundary conditions. In this basis, if we partition $\tilde{V}_0$ and $\tilde{V}_N$ conformably with the state vector, i.e, $\tilde{V}_0 = [\tilde{V}_{0,1}\;\; \tilde{V}_{0,2}]$ and $\tilde{V}_N = [\tilde{V}_{N,1}\;\; \tilde{V}_{N,2}]$, (5) can be re-written in this basis as

$$\begin{bmatrix} \tilde{V}_{0,1} & \tilde{V}_{N,1} & \tilde{V}_{0,2} + \tilde{V}_{N,2}A_{X,22}^{N} & \tilde{V}_{N,2} \end{bmatrix} x = v, \tag{32}$$

where $x = [x_1^{\top}(0)\;\; x_1^{\top}(N)\;\; x_2^{\top}(0)\;\; \lambda_2^{\top}(N)]^{\top}$. Let us now consider (7), and let $\tilde{H} = \begin{bmatrix} \tilde{H}_1 & \tilde{H}_2 \\ \tilde{H}_3 & \tilde{H}_4 \end{bmatrix}$ be partitioned conformably with $\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$, where $\tilde{H}_3 = \tilde{H}_2^{\top}$. Finally, let K_0 and K_N be basis matrices for $\ker\tilde{V}_0$ and $\ker\tilde{V}_N$, respectively, to be used to eliminate the multiplier η. Thus, (7) can be re-written as

$$-K_0^{\top}\begin{bmatrix} \lambda_1(0) \\ \lambda_2(0) \end{bmatrix} = K_0^{\top}\tilde{H}_1\left(\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} - \tilde{h}_0\right) + K_0^{\top}\tilde{H}_2\left(\begin{bmatrix} x_1(N) \\ x_2(N) \end{bmatrix} - \tilde{h}_N\right), \tag{33}$$

$$K_N^{\top}\begin{bmatrix} \lambda_1(N) \\ \lambda_2(N) \end{bmatrix} = K_N^{\top}\tilde{H}_3\left(\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} - \tilde{h}_0\right) + K_N^{\top}\tilde{H}_4\left(\begin{bmatrix} x_1(N) \\ x_2(N) \end{bmatrix} - \tilde{h}_N\right). \tag{34}$$

Now, defining $\tilde{H}_i = \begin{bmatrix} \tilde{H}_{i,11} & \tilde{H}_{i,12} \\ \tilde{H}_{i,21} & \tilde{H}_{i,22} \end{bmatrix}$ and also $\tilde{H}_i^1 = \begin{bmatrix} \tilde{H}_{i,11} \\ \tilde{H}_{i,21} \end{bmatrix}$ and $\tilde{H}_i^2 = \begin{bmatrix} \tilde{H}_{i,12} \\ \tilde{H}_{i,22} \end{bmatrix}$ for $i \in \{1,2,3,4\}$, in this basis (5) and (7) can be expressed as the single linear equation

$$\left[\begin{array}{c} F_1 \\ \hline F_2 \end{array}\right] x = g, \tag{35}$$

where

$$\begin{aligned} F_1 &= \left[\begin{array}{c|c|c|c} \tilde{V}_{0,1} & \tilde{V}_{N,1} & \tilde{V}_{0,2} + \tilde{V}_{N,2} A_{X,22}^N & \tilde{V}_{N,2} \end{array}\right], \\ F_2 &= \operatorname{diag}\{K_0^\top, K_N^\top\} \left[\begin{array}{c|c|c|c} \tilde{H}_1^1 & \tilde{H}_1^2 + \tilde{H}_2^2 A_{X,22}^N & \tilde{H}_2^1 & \begin{bmatrix} -\tilde{H}_{2,12} P \\ (A_{X,22}^\top)^N - \tilde{H}_{2,22} P \end{bmatrix} \\ \tilde{H}_3^1 & \tilde{H}_3^2 + \tilde{H}_4^2 A_{X,22}^N & \tilde{H}_4^1 & \begin{bmatrix} \tilde{H}_{4,12} P \\ \tilde{H}_{4,22} P - I \end{bmatrix} \end{array}\right], \\ g &= \left[\begin{array}{c} v \\ \hline \operatorname{diag}\{K_0^\top, K_N^\top\} \tilde{H} \begin{bmatrix} \tilde{h}_0 \\ \tilde{h}_N \end{bmatrix} \end{array}\right]. \end{aligned}$$

We have just proved the following result.

Theorem 2 *Under Assumption 3.1, Problem 1 admits solutions if and only if (35) does. For any solution $x = [x_1^\top(0) \;\; x_1^\top(N) \;\; x_2^\top(0) \;\; \lambda_2^\top(N)]^\top$ we get an optimal initial state $x(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}$ and a class of optimal controls parameterised by (27) and (31). The solutions obtained in this way are all the solutions of Problem 1.*

Example 4.1 Consider a finite-horizon LQ problem in the time interval $\{0,\dots,N\}$, involving the matrices given in Example 3.1. The initial and final states are constrained to be equal, i.e., $x(0) = x(N)$. Let $H = I_{2n}$, $h_0 = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$ and $h_N = 0$. As aforementioned, $X = \operatorname{diag}\{0,1\}$ is a solution of CGDARE(Σ), leading to $A_X = \operatorname{diag}\{1,0\}$. By taking $T = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, we obtained $T^\top R_X T = \operatorname{diag}\{4,0\}$, so that $R_{0,X} = 4$, $B_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $B_2 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$. Therefore, the reachable subspace of the pair (A_X, B_2) is spanned by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which means this system is already in the desired basis. Thus, $A_{X,11} = 1$, $A_{X,12} = A_{X,22} = 0$, $B_{11} = B_{12} = 2$ and $B_{21} = -2$. In this case, (22) and (25) yield $x_2(t+1) = B_{12} u_2(t)$, $\lambda_2(t) = 0 \cdot \lambda_2(t+1)$, and $u_2(t) = R_{X,0}^{-1} B_{12}^\top \lambda_2(t+1)$. This implies that $\lambda_2(t) = 0$ for all $t \in \{0,\dots,N-1\}$ and is equal to $\lambda_2(N)$ for $t = N$, so that $u_2(t) = 0$ for all $t \in \{0,\dots,N-2\}$ and $u_2(N-1) = R_{X,0}^{-1} B_{12}^\top \lambda_2(N)$. Thus $x_2(t)$ is equal to $x_2(0)$ at $t = 0$, is equal to zero for $t \in \{1,\dots,N-1\}$ and is equal to $B_{12}^\top R_{X,0}^{-1} B_{12}^\top \lambda_2(N) = \lambda_2(N)$ for $t = N$. In this basis, (5) gives rise to $x_1(0) = x_1(N)$ and $x_2(0) = x_2(N) = \lambda_2(N)$, which are linear in $x_1(N)$ and $\lambda_2(N)$, while (33-34) can be written as $x_1(0) + x_1(N) = h_1$ and

$x_2(0)+x_2(N)+\lambda_2(0)-\lambda_2(N)=h_2$. Since $\lambda_2(0)=0$ and $x_2(N)=\lambda_2(N)$, the latter can be written as $x_2(0)=h_2$. Therefore, the boundary conditions can be written in the form (35). This linear equation admits only the solution $x_1(0)=x_1(N)=h_1/2$ and $x_2(0)=\lambda_2(N)=h_2$. Now we can compute the optimal control law. First, $u_2(t)$ is zero for all $t\in\{0,\dots,N-2\}$ and $u_2(N-1)=R_{X,0}^{-1}B_{12}^{\top}\lambda_2(N)=h_2/2$. To compute u_1, we write (19) as $x_1(t+1)=1\cdot x_1(t)-2\,u_1(t)+\xi(t)$. The term $\xi(t)$ is zero for all $t\in\{0,\dots,N-2\}$ and $\xi(N-1)=B_{11}\,R_{0,X}^{-1}\,B_{12}^{\top}\,\lambda_2(N)=\lambda_2(N)=h_2$. We write (31) explicitly as

$$
\begin{aligned}
x_1(N) \;=\; & x_1(0)+\begin{bmatrix} I & A_{X,11} & A_{X,11}^2 & \dots & A_{X,11}^{N-1}\end{bmatrix}\begin{bmatrix} h_2\\ 0\\ \vdots\\ 0\end{bmatrix}\\
&+\underbrace{\begin{bmatrix} -2 & -2 & \dots & -2\end{bmatrix}}_{N}\begin{bmatrix} u_1(N-1)\\ u_1(N-2)\\ \vdots\\ u_1(0)\end{bmatrix}
\end{aligned}
$$

which gives

$$
\begin{bmatrix} u_1(N-1)\\ u_1(N-2)\\ \vdots\\ u_1(0)\end{bmatrix}=\frac{h_2}{2N}\begin{bmatrix}1\\1\\ \vdots\\ 1\end{bmatrix}+\begin{bmatrix} 1-N & 0 & \dots & 0\\ 1 & 2-N & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 1 & 1 & \dots & -1\\ 1 & 1 & \dots & 1\end{bmatrix}v,
$$

where v is arbitrary and represents the degree of freedom in the control u_1. $\square$

Remark 2 So far, we have not considered the problem of existence of solutions for Problem 1. In general, the existence of a state trajectory $x(t)$ satisfying the constraints (1-2) for some $u(t)$ is not ensured, since we have not assumed reachability on (1). A necessary and sufficient condition for the existence of optimal solutions is that there exist state and input trajectories satisfying (1-2) (feasible solutions). In fact, since the optimal control problem formulated in Section 2 involves a finite number of variables – precisely, $L=m\cdot N$ for the control plus n for the initial state – Problem 1 can be restated as a quadratic static optimisation problem in these $L+n$ variables with linear constraints. Thus, a solution to Problem 1 exists if and only if a feasible solution – i.e., a state and input functions satisfying both (1) and (2) – exists.

Proof of Theorem 1

Recall that we have defined the $m\times m$ orthogonal matrix $T=[T_1\;\;T_2]$ where $\mathrm{range}\,T_1=\mathrm{range}\,R_X$ and $\mathrm{range}\,T_2=\mathrm{range}\,G_X=\ker R_X$, so that $T^\top R_X\,T=\mathrm{diag}\{R_{X,0},O\}$, where $R_{X,0}$ is invertible. Its dimension is denoted by m_1. We also defined $B_1\triangleq BT_1$ and $B_2\triangleq BT_2$. Moreover, we considered $U=[U_1\;\,U_2]$ such that U_1 spans $\mathscr{R}_X$, and (17) holds. Let $\hat{U}=\mathrm{diag}\{U,U,I_{m_1},I_{m_2}\}$. Let r denote the size of $\mathscr{R}_X$. Using (17) into (16), and taking into consideration and defining the two unimodular matrices

$$\Omega_1\triangleq\begin{bmatrix} I_r & O & O & O & O & O\\ O & O & I_r & O & O & O\\ O & O & O & O & O & I_{m_2}\\ O & I_{n-r} & O & O & O & O\\ O & O & O & I_{n-r} & O & O\\ O & O & O & O & I_{m_1} & O\end{bmatrix}\quad\text{and}\quad\Omega_2\triangleq\begin{bmatrix} I_r & O & O & O & O & O\\ O & O & O & I_{n-r} & O & O\\ O & O & I_r & O & O & O\\ O & O & O & O & I_{n-r} & O\\ O & O & O & O & O & I_{m_1}\\ O & I_{m_2} & O & O & O & O\end{bmatrix},$$

along with $\hat{U}_X\triangleq\Omega_1\hat{U}^{-1}\hat{T}^\top U_X$ and $\hat{V}_X\triangleq V_X\,\hat{T}\hat{U}\Omega_2$, we get (18). Let $P(z)=\hat{U}_X\,(G-zF)\,\hat{V}_X$. Since in (18) the pair $(A_{X,11},B_{21})$ is reachable by construction, all the r rows of the submatrix $[A_{X,11}-zI_r\;\;B_{21}]$ are linearly independent for every $z\in\mathbb{C}\cup\{\infty\}$. This also means that of the $r+m_2$ columns of $[A_{X,11}-zI_r\;\;B_{21}]$, only r are linearly independent, and this gives rise to the presence of a null-space of $P(z)$ whose dimension m_2 is independent of $z\in\mathbb{C}\cup\{\infty\}$. We obtain[3]

$$\mathrm{rank}P(z)=r+\mathrm{rank}\left[\begin{array}{c|ccc} I_r-zA_{X,11}^\top & O & O & O\\ -zB_{21}^\top & O & O & O\\ \hline O & A_{X,22}-zI_{n-r} & O & B_{12}\\ -zA_{X,12}^\top & O & I_{n-r}-zA_{X,22}^\top & O\\ -zB_{11}^\top & O & -zB_{12}^\top & R_{X,0}\end{array}\right].$$

Now, consider the rank of $\left[\begin{smallmatrix} I_r-zA_{X,11}^\top\\ -zB_{21}^\top\end{smallmatrix}\right]$. Again, since the pair $(A_{X,11},B_{21})$ is reachable, this rank is constant and equal to r for every $z\in\mathbb{C}\cup\{\infty\}$. Thus, $\mathrm{rank}P(z)=2r+\mathrm{rank}P_1(z)$. Since $\det P_1(z)=\det(A_{X,22}-zI_{n-r})\cdot\det(I_{n-r}-zA_{X,22}^\top)\cdot\det R_{X,0}$, a value $z\in\mathbb{C}$ can be found for which $\det P_1(z)\neq 0$. Hence, the normal rank of $P_1(z)$ is equal to $2\,(n-r)+m_1$, and therefore the normal rank of $P(z)$ is $2r+2\,(n-r)+m_1=2n+m_1$. The generalised eigenvalues of the pencil $P(z)$ are the values $z\in\mathbb{C}\cup\{\infty\}$ for which the rank of $P_1(z)$ is smaller than its normal rank $2\,(n-r)+m_1$. ■

[3]Let $\Xi=\left[\begin{smallmatrix}\Xi_{11} & \Xi_{12}\\ O & \Xi_{22}\end{smallmatrix}\right]$. Observe that if either Ξ_{11} is full row-rank or Ξ_{22} is full column-rank, then $\mathrm{rank}\,\Xi=\mathrm{rank}\,\Xi_{11}+\mathrm{rank}\,\Xi_{22}$.

Acknowledgment

The research presented in this paper has been partially supported by the Italian Ministry for Education and Research (MIUR) under PRIN grant n. 20085FFJ2Z and by the Australian Research Council under the grant FT120100604.

References

[1] D.J. Clements and H. Wimmer. Existence and uniqueness of unmixed solutions of the discrete-time algebraic Riccati equation. *Syst. Control Lett.*, 50:343–346, 2003.

[2] A. Ferrante. On the structure of the solution of discrete-time algebraic Riccati equation with singular closed-loop matrix. *IEEE Trans. Aut. Control,* AC-49(11):2049–2054, 2004.

[3] A. Ferrante and L. Ntogramatzidis. "Employing the algebraic Riccati equation for a parametrization of the solutions of the finite-horizon LQ problem: the discrete-time case". *Syst. Control Lett.*, 54:693–703, 2005.

[4] A. Ferrante, and L. Ntogramatzidis, "A unified approach to finite-horizon generalized LQ optimal control problems for discrete-time systems". *Linear Algebra Appl.*, 425(2-3):242–260, 2007.

[5] A. Ferrante, and L. Ntogramatzidis, "The Generalised Discrete Algebraic Riccati Equation in LQ optimal control". *Automatica*. In press. Manuscript available at `http://arxiv.org/abs/1201.3704`.

[6] A. Ferrante, and H.K. Wimmer, "Order reduction of discrete-time algebraic Riccati equations with singular closed-loop matrix". *Oper. Matrices,* 1(1):61–70, 2007.

[7] A. Hansson, and P. Hagander, How to decompose semi-definite discrete-time algebraic Riccati equations *Europ. J. Control*, 5:245–258, 1999.

[8] V. Ionescu and C. Oară. Generalized discrete-time Riccati theory. *SIAM J. Control Optim.*, 34(2):601–619, 1996.

[9] V. Ionescu, C. Oară, and M. Weiss. *Generalized Riccati theory and robust control, a Popov function approach*. Wiley, 1999.

[10] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems*. John Wiley & Sons, New York, 1972.

[11] P. Lancaster and L. Rodman. *Algebraic Riccati equations*. Clarendon Press, Oxford, 1995.

[12] F.L. Lewis and V. Syrmos. *Optimal Control*. John Wiley & Sons, New York, 1995.

[13] J.H. Park, S. Han, and W.H. Kwon, LQ tracking controls with fixed terminal states and their application to receding horizon controls. *Syst. Control Lett.*, 57:772–777, 2008.

[14] A. Sideris, and L.A. Rodriguez, A Riccati approach for constrained linear quadratic optimal control. *Int. J. Control*, 84(2):370–380, 2011.

[15] A.A. Stoorvogel and A. Saberi. The discrete-time algebraic Riccati equation and linear matrix inequality. *Linear Algebra Appl.*, 274:317–365, 1998.

[16] H. Trentelman, A. Stoorvogel, and M. Hautus, *Control theory for linear systems*, ser. Communications and Control Engineering. Great Britain: Springer, 2001.

[17] P. Van Dooren, “The computation of Kroneckers canonical form of a singular pencil”. *Linear Algebra Appl.*, 27:103–140, 1979.