On Linear Quadratic Optimal Control and Algebraic Riccati Equations for Infinite-Dimensional Differential-Algebraic Equations

Mark R. Opmeer and Olof J. Staffans

June 21, 2024

1 Introduction

There has been significant recent interest in infinite-dimensional differential algebraic equations (DAEs) [6, 10, 8, 12] and particularly in linear quadratic optimal control and (differential or algebraic) Riccati equations for infinite-dimensional DAEs [9, 1]. In this article we show how results and methods from [2, 13, 15] can be utilized to obtain an algebraic Riccati equation for an infinite-horizon linear quadratic optimal control problem for a very general class of infinite-dimensional DAEs.

This article is structured as follows. In Section 2 we review relevant notions from [2]. In Section 3 we precisely formulate the linear quadratic optimal control problem for the class of DAEs considered. We will utilize the (internal) Cayley transform to solve this problem. Therefore, in Section 4 we study the Cayley transform, in Section 5 we recall the solution of the discrete-time linear quadratic optimal control problem and reformulate this in a suitable form, and in Section 6 we relate the Cayley transform and linear quadratic optimal control as in [13, 15]. In Section 7 we present our main result: the solution to the linear quadratic optimal control problem for the considered class of DAEs through an algebraic Riccati equation. In Section 8 we illustrate this with a simple, but interesting, finite-dimensional example. The result in Section 7 assumes that every initial state has finite cost. In [15] we actually considered a more general case and as already noted in [17] for DAEs the case where not every initial state has finite cost is especially relevant. Therefore in Section 9 we consider this more general case; for that we introduce some further concepts from [2], we formulate a more general version of the result from Section 7, we compare that results to [17] and we consider a finite-dimensional example for which not every initial state has finite cost.

2 Preliminaries

In this section we discuss some general notions from [2] on the input/state/output (i/s/o) node approach to DAEs which are relevant for our later results on infinite-horizon linear quadratic optimal control and algebraic Riccati equations.

Definition 2.1. *Let $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{Y}$ be Hilbert spaces. An* i/s/o node *is a multi-valued operator from* $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$ *to* $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{Y}\end{smallmatrix}\right]$. *The graph of the i/s/o node S is (note that the components are the different way around than usual, this is to conform to the convention used in [2]):*

$$\operatorname{gph}(S) = \left\{ \begin{bmatrix} Sq \\ q \end{bmatrix} : q \in \operatorname{dom}(S) \right\}.$$

The i/s/o node is called closed *if S is a closed multi-valued operator (i.e. when* $\operatorname{gph}(S)$ *is a closed subspace) and* bounded *if S is a bounded single-valued operator with domain* $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$.

The definition of i/s/o node is adapted from [2, Definition 4.1.5]. Time-domain trajectories of varies kinds (classical, generalized and mild) are defined in [2, Definitions 4.1.5 and 4.1.7]. These notions capture that in some suitable sense time-domain trajectories should satisfy

$$\begin{bmatrix} \dot{x} \\ y \end{bmatrix} \in S \begin{bmatrix} x \\ u \end{bmatrix}, \qquad \text{i.e.} \quad \begin{bmatrix} \dot{x} \\ y \\ x \\ u \end{bmatrix} \in \operatorname{gph}(S). \tag{1}$$

Since we won't need time-domain trajectories in the sense of [2], we will not elaborate further.

Remark 2.2. *The connection between i/s/o nodes and the "conventional" approach to DAEs becomes most clear from the notion of a* kernel representation *from [2, Definition 4.1.16] (see also [2, Lemma 4.1.15]): for a closed i/s/o node there exist a Hilbert space $\mathscr{Z}$ and bounded single-valued everywhere-defined operators $E : \mathscr{X} \to \mathscr{Z}$, $M : \mathscr{X} \to \mathscr{Z}$, $N_{\text{in}} : \mathscr{U} \to \mathscr{Z}$ and $N_{\text{out}} : \mathscr{Y} \to \mathscr{Z}$ such that*

$$\operatorname{gph}(S) = \left\{ \begin{bmatrix} z \\ y \\ x \\ u \end{bmatrix} \in \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \\ \mathscr{X} \\ \mathscr{U} \end{bmatrix} : Ez + N_{\text{out}}y = Mx + N_{\text{in}}u \right\},$$

and conversely, the above defines the graph of a closed i/s/o node. The notion of time-domain trajectory then means that in some suitable sense it should satisfy

$$E\dot{x} + N_{\text{out}}y = Mx + N_{\text{in}}u.$$

Note that compared to the usual form of a DAE, there are no separate equations for $E\dot{x}$ and y, but these instead are generally coupled. This makes i/s/o nodes more general than "conventional" DAEs (an example of an i/s/o node which is not a conventional DAE is given in Example 8.4).

Example 2.3. *Let $\mathscr{U}$, $\mathscr{X}$, $\mathscr{Z}$ and $\mathscr{Y}$ be Hilbert spaces and let $\mathbf{A} : \mathscr{X} \to \mathscr{Z}$, $\mathbf{B} : \mathscr{U} \to \mathscr{Z}$, $\mathbf{C} : \mathscr{X} \to \mathscr{Y}$, $\mathbf{D} : \mathscr{U} \to \mathscr{Y}$ and $\mathbf{E} : \mathscr{X} \to \mathscr{Z}$ be bounded single-valued everywhere-defined operators. The conventional DAE*

$$\mathbf{E}\dot{x} = \mathbf{A}x + \mathbf{B}u, \qquad y = \mathbf{C}x + \mathbf{D}u,$$

is described by the closed i/s/o node

$$\mathrm{gph}(S) = \left\{ \begin{bmatrix} z \\ y \\ x \\ u \end{bmatrix} \in \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \\ \mathscr{X} \\ \mathscr{U} \end{bmatrix} : \mathbf{E}z = \mathbf{A}x + \mathbf{B}u,\ y = \mathbf{C}x + \mathbf{D}u \right\}.$$

Definition 2.4. *For $\lambda \in \mathbb{C}$, the* formal i/s/o resolvent *of the i/s/o node S is the multi-valued operator $\widehat{\mathfrak{G}}(\lambda)$ from $\left[\begin{smallmatrix} \mathscr{X} \\ \mathscr{U} \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} \mathscr{X} \\ \mathscr{Y} \end{smallmatrix}\right]$ whose graph is given by*

$$\mathrm{gph}(\widehat{\mathfrak{G}}(\lambda)) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathrm{gph}(S).$$

The i/s/o resolvent set $\rho(S)$ *of S consists of those $\lambda \in \mathbb{C}$ for which $\widehat{\mathfrak{G}}(\lambda)$ is a bounded single-valued operator with domain $\left[\begin{smallmatrix} \mathscr{X} \\ \mathscr{U} \end{smallmatrix}\right]$. The i/s/o node is called* resolvable *if $\rho(S)$ is non-empty and* future-resolvable *if $\rho(S) \cap \mathbb{C}_+ \neq \emptyset$. For $\lambda \in \rho(S)$ we have*

$$\widehat{\mathfrak{G}}(\lambda) = \begin{bmatrix} \widehat{\mathfrak{A}}(\lambda) & \widehat{\mathfrak{B}}(\lambda) \\ \widehat{\mathfrak{C}}(\lambda) & \widehat{\mathfrak{D}}(\lambda) \end{bmatrix},$$

where $\widehat{\mathfrak{A}}$ is called the state/state resolvent*, $\widehat{\mathfrak{B}}$ is called the* input/state resolvent*, $\widehat{\mathfrak{C}}$ is called the* state/output resolvent *and $\widehat{\mathfrak{D}}$ is called the* input/output resolvent*.*

The notions in Definition 2.4 are taken from [2, Definition 5.5.8]. For the connection with linear quadratic optimal control, we need the notion of future-resolvable (i.e. $\rho(S)$ contains an element in the open right-half plane) rather than the weaker notion of resolvable.

We have

$$\mathrm{gph}(S) = \begin{bmatrix} \lambda & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathrm{gph}(\widehat{\mathfrak{G}}(\lambda)).$$

Remark 2.5. *Related to Remark 2.2, by [2, Lemma 5.5.5] we have for a closed i/s/o node*

$$\mathrm{gph}(\widehat{\mathfrak{G}}(\lambda)) = \left\{ \begin{bmatrix} x \\ y \\ x^0 \\ u \end{bmatrix} \in \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \\ \mathscr{X} \\ \mathscr{U} \end{bmatrix} : (\lambda E - M)x + N_{\mathrm{out}}y = Ex^0 + N_{\mathrm{in}}u \right\}.$$

Example 2.6. *The i/s/o resolvent set of the conventional DAE from Example 2.3 consists of those $\lambda \in \mathbb{C}$ for which $\lambda\mathbf{E} - \mathbf{A}$ has a (bounded single-valued everywhere-defined) inverse and the various resolvent operators are given by*

$$\widehat{\mathfrak{A}}(\lambda) = (\lambda\mathbf{E} - \mathbf{A})^{-1}\mathbf{E}, \qquad \widehat{\mathfrak{B}}(\lambda) = (\lambda\mathbf{E} - \mathbf{A})^{-1}\mathbf{B},$$
$$\widehat{\mathfrak{C}}(\lambda) = \mathbf{C}(\lambda\mathbf{E} - \mathbf{A})^{-1}\mathbf{E}, \qquad \widehat{\mathfrak{D}}(\lambda) = \mathbf{C}(\lambda\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}.$$

Remark 2.7. *By [2, Theorems 10.2.9 and 10.2.14], the (formal) i/s/o resolvent of a resolvable i/s/o node is an i/s/o pseudoresolvent and conversely. Under the name* resolvent linear system*, i/s/o pseudoresolvents were studied in [13] in connection with linear quadratic optimal control. The representation results of [2] now allow us to connect the results from [13] more clearly to DAEs.*

Definition 2.8. *Let Ω be a non-empty open subset of $\mathbb{C}$. A* frequency domain Ω trajectory *of an i/s/o node is a quadruple $(\hat{x}, \hat{y}, x^0, \hat{u})$ where $\hat{x}$, $\hat{y}$ and $\hat{u}$ are holomorphic functions defined on Ω with values in $\mathscr{X}$, $\mathscr{Y}$ and $\mathscr{U}$ respectively and $x^0 \in \mathscr{X}$ such that for all $\lambda \in \Omega$*

$$\begin{bmatrix} \hat{x}(\lambda) \\ \hat{y}(\lambda) \\ x^0 \\ \hat{u}(\lambda) \end{bmatrix} \in \operatorname{gph}(\widehat{\mathfrak{G}}(\lambda)).$$

Remark 2.9. *The above is [2, Definition 11.1.1]. By [2, Lemma 11.1.6], for a resolvable i/s/o node with $\Omega \subset \rho(S)$, for every $x^0 \in \mathscr{X}$ and every holomorphic $\mathscr{U}$-valued $\hat{u}$, there exist unique $\hat{x}$ and $\hat{y}$ such that the quadruple forms an Ω trajectory; namely*

$$\begin{bmatrix} \hat{x}(\lambda) \\ \hat{y}(\lambda) \end{bmatrix} = \widehat{\mathfrak{G}}(\lambda) \begin{bmatrix} x^0 \\ \hat{u}(\lambda) \end{bmatrix}.$$

The following definition allows us to add an output to an i/s/o node. This is relevant in linear quadratic optimal control since the optimal control can be characterized by adding a certain output and subsequently putting that additional output equal to zero.

Definition 2.10. *Let S be an i/s/o node, let $\mathscr{Y}_0$ be a Hilbert space and let $C = \begin{bmatrix} C_1 & C_0 \end{bmatrix} : \begin{bmatrix} \mathscr{X} \\ \mathscr{X} \end{bmatrix} \to \mathscr{Y}_0$ and $D = \begin{bmatrix} D_1 & D_0 \end{bmatrix} : \begin{bmatrix} \mathscr{Y} \\ \mathscr{U} \end{bmatrix} \to \mathscr{Y}_0$ be (bounded single-valued everywhere-defined) operators. The* nonstandard output extension S^{ext} *of S with observation extension C and feedthrough extension D is defined by*

$$\operatorname{gph}(S^{\text{ext}}) = \left\{ \begin{bmatrix} z \\ \begin{bmatrix} C_1 z + C_0 x + D_1 y + D_0 u \\ y \end{bmatrix} \\ x \\ u \end{bmatrix} \in \begin{bmatrix} \mathscr{X} \\ \begin{bmatrix} \mathscr{Y}_0 \\ \mathscr{Y} \end{bmatrix} \\ \mathscr{X} \\ \mathscr{U} \end{bmatrix} : \begin{bmatrix} z \\ y \\ x \\ u \end{bmatrix} \in \operatorname{gph}(S) \right\}.$$

A standard output extension *is a nonstandard output extension where $C_1 = 0$ and $D_1 = 0$.*

We equivalently have by [2, (5.1.12b)]

$$\operatorname{gph}(S^{\text{ext}}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \begin{bmatrix} C_1 \\ 0 \end{bmatrix} & \begin{bmatrix} D_1 \\ 1 \end{bmatrix} & \begin{bmatrix} C_0 \\ 0 \end{bmatrix} & \begin{bmatrix} D_0 \\ 0 \end{bmatrix} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(S),$$

and we can recover S from S^{ext} by [2, (5.1.13b)] through

$$\operatorname{gph}(S) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \begin{bmatrix} 0 & 1 \end{bmatrix} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(S^{\text{ext}}).$$

Definition 2.10 is from [2, Definition 5.1.23 (ii)] and [2, Definition 5.1.33 (ii)]. If S is bounded, then a nonstandard output extension is equivalent to a standard output extension [2, Lemma 6.2.1 (vii)].

Remark 2.11. *From [2, Lemma 5.5.15] we infer that $\rho(S^{\text{ext}}) = \rho(S)$ for any nonstandard output extension S^{ext} of a resolvable i/s/o node S.*

3 Linear quadratic optimal control

For the purposes of linear quadratic optimal control, we restrict the set Ω in Definition 2.8 to be a subset of $\rho(S) \cap \mathbb{C}_+$ (as was done in [15]). In that case, for certain Ω trajectories we can give a time-domain interpretation.

Definition 3.1. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty open subset of $\rho(S) \cap \mathbb{C}_+$. For $x^0 \in \mathscr{X}$ the set of* i/o stable Ω trajectories *is defined as follows. Let $u \in L^2(\mathbb{R}^+; \mathscr{U})$ and let $\hat{u}$ be the restriction to Ω of the Laplace transform of u. Let $\hat{y}$ be the output component of the corresponding Ω trajectory. If there exists a (necessarily unique) $y \in L^2(\mathbb{R}^+; \mathscr{Y})$ whose Laplace transform restricted to Ω equals $\hat{y}$, then we call (y, u) an i/o stable Ω trajectory with initial condition x^0.*

We say that S satisfies the Ω finite cost condition *if for all $x^0 \in \mathscr{X}$ the corresponding set of i/o stable Ω trajectories is non-empty.*

Remark 3.2. *The above essentially coincides with the notion of stable input/output pairs from [13]. The difference is that in [13] an additional assumption is made on the resolvent linear system (i.e. resolvable i/s/o node) which allows for frequency domain trajectories for Ω an exponential region to always be interpreted as Laplace transforms of distributions. By using the ideas in [15], we can circumvent this additional assumption (and can allow for more general Ω).*

Remark 3.3. *The definition of trajectories and therefore of the optimal control problem considered depends on the choice of Ω. In most applications, $\rho(S) \cap \mathbb{C}_+$ is connected and then the choice of Ω is immaterial (see [15]). More generally, $\rho(S) \cap \mathbb{C}_+$ usually contains a subset of the form $[r, \infty)$ for some $r > 0$ and the natural choice of Ω is then as (a subset of) this connected component of $\rho(S) \cap \mathbb{C}_+$ (this is the choice which is made in [13]).*

Definition 3.4. *The linear quadratic optimal control problem for a future-resolvable i/s/o node is: for given $x^0 \in \mathscr{X}$ find the i/o stable Ω trajectory with initial condition x^0 of minimal norm, i.e. minimize $\|u\|^2_{L^2(\mathbb{R}^+;\mathscr{U})} + \|y\|^2_{L^2(\mathbb{R}^+;\mathscr{Y})}$.*

4 The internal Cayley transform

As in [15] and [13], the easiest way to approach the linear quadratic optimal control problem at this high level of generality is through utilizing the internal Cayley transform to translate the problem to a discrete-time linear quadratic optimal control problem.

Definition 4.1. *For $\alpha \in \mathbb{C}$ with* $\mathrm{Re}(\alpha) > 0$*, the* Cayley transform *of the i/s/o node S is the multi-valued operator S_d from $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$ to $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{Y}\end{smallmatrix}\right]$ whose graph is given by*

$$\mathrm{gph}(S_d) = \begin{bmatrix} \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{\overline{\alpha}}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{-1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{\alpha}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathrm{gph}(S).$$

Definition 4.1 is from [2, Definition 14.9.7].

We have [2, (14.9.6(b)]

$$\mathrm{gph}(S) = \begin{bmatrix} \frac{\alpha}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{-\overline{\alpha}}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathrm{gph}(S_d),$$

and

$$\mathrm{gph}(\widehat{\mathfrak{G}}(\alpha)) = \begin{bmatrix} \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \sqrt{2\mathrm{Re}(\alpha)} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathrm{gph}(S_d),$$

and (see [2, proof of Lemma 14.9.8])

$$\mathrm{gph}(S_d) = \begin{bmatrix} \sqrt{2\mathrm{Re}(\alpha)} & 0 & \frac{-1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathrm{gph}(\widehat{\mathfrak{G}}(\alpha)).$$

If S is future-resolvable and $\alpha \in \rho(S) \cap \mathbb{C}_+$, then the Cayley transform with parameter α is a single-valued bounded operator with domain $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$ and in particular it therefore can be written as

$$S_d = \begin{bmatrix} A_d & B_d \\ C_d & D_d \end{bmatrix},$$

for (single-valued bounded everywhere-defined) operators $A_d : \mathscr{X} \to \mathscr{X}$, $B_d : \mathscr{U} \to \mathscr{X}$, $C_d : \mathscr{X} \to \mathscr{Y}$ and $D_d : \mathscr{U} \to \mathscr{Y}$. Explicitly we have

$$A_d = -I + 2\,\mathrm{Re}(\alpha)\widehat{\mathfrak{A}}(\alpha), \quad B_d = \sqrt{2\mathrm{Re}(\alpha)}\,\widehat{\mathfrak{B}}(\alpha), \quad C_d = \sqrt{2\mathrm{Re}(\alpha)}\,\widehat{\mathfrak{C}}(\alpha), \quad D_d = \widehat{\mathfrak{D}}(\alpha). \tag{2}$$

The formal i/s/o resolvent of S_d (in accordance with Definition 2.4, but using w for the resolvent variable) is given by

$$\operatorname{gph}(\widehat{\mathfrak{G}}_d(w)) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & w & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(S_d) = \begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ -\sqrt{2\mathrm{Re}(\alpha)} & 0 & \frac{1+w}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(\widehat{\mathfrak{G}}(\alpha)).$$

Using [2, (10.2.1c)] (which is basically the resolvent identity)

$$\operatorname{gph}(\widehat{\mathfrak{G}}(\alpha)) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \alpha-\lambda & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(\widehat{\mathfrak{G}}(\lambda)), \tag{3}$$

we obtain the following relation between the formal i/s/o resolvents

$$\operatorname{gph}(\widehat{\mathfrak{G}}_d(w)) = \begin{bmatrix} \frac{\alpha-\lambda}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ -\sqrt{2\mathrm{Re}(\alpha)} + \frac{1+w}{\sqrt{2\mathrm{Re}(\alpha)}}(\alpha-\lambda) & 0 & \frac{1+w}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(\widehat{\mathfrak{G}}(\lambda)).$$

With the following correspondence between the resolvent variables

$$\lambda = \frac{\alpha w - \overline{\alpha}}{w+1}, \quad w = \frac{\overline{\alpha}+\lambda}{\alpha-\lambda},$$

the above relation becomes

$$\operatorname{gph}(\widehat{\mathfrak{G}}_d(w)) = \begin{bmatrix} \frac{\alpha-\lambda}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{\sqrt{2\mathrm{Re}(\alpha)}}{\alpha-\lambda} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(\widehat{\mathfrak{G}}(\lambda)).$$

From this we can deduce the following relation between the various resolvent operators

$$\widehat{\mathfrak{A}}_d(w) = \frac{(\alpha-\lambda)^2}{\overline{\alpha}+\alpha}\widehat{\mathfrak{A}}(\lambda) + \frac{\alpha-\lambda}{\overline{\alpha}+\alpha}I, \qquad \widehat{\mathfrak{B}}_d(w) = \frac{\alpha-\lambda}{\sqrt{2\mathrm{Re}(\alpha)}}\widehat{\mathfrak{B}}(\lambda),$$
$$\widehat{\mathfrak{C}}_d(w) = \frac{\alpha-\lambda}{\sqrt{2\mathrm{Re}(\alpha)}}\widehat{\mathfrak{C}}(\lambda), \qquad \widehat{\mathfrak{D}}_d(w) = \widehat{\mathfrak{D}}(\lambda),$$

and we obtain that $\lambda \in \rho(S)$ if and only if $w \in \rho(S_d)$.

Remark 4.2. *It will later be important how the Cayley transform interacts with nonstandard output extensions. Since*

$$\begin{bmatrix} \frac{\alpha}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{-\overline{\alpha}}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ \begin{bmatrix} C_{d,1} \\ 0 \end{bmatrix} & \begin{bmatrix} D_{d,1} \\ 1 \end{bmatrix} & \begin{bmatrix} C_{d,0} \\ 0 \end{bmatrix} & \begin{bmatrix} D_{d,0} \\ 0 \end{bmatrix} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{\overline{\alpha}}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{-1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{\alpha}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ \begin{bmatrix} C_1 \\ 0 \end{bmatrix} & \begin{bmatrix} D_1 \\ 1 \end{bmatrix} & \begin{bmatrix} C_0 \\ 0 \end{bmatrix} & \begin{bmatrix} D_0 \\ 0 \end{bmatrix} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

with

$$C_1 = \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}}\left(C_{d,1} - C_{d,0}\right), \qquad D_1 = D_{d,1}, \qquad C_0 = \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}}\left(\overline{\alpha}C_{d,1} + \alpha C_{d,0}\right), \qquad D_0 = D_{d,0},$$

we see that the inverse Cayley transform of a nonstandard output extension of the Cayley transform of S is a nonstandard output extension of S. If the output extension of the Cayley transform is standard (i.e. $C_{d,1} = 0$ and $D_{d,1} = 0$), then the output extension of S need not be standard since then $C_1 = \frac{-1}{\sqrt{2\mathrm{Re}(\alpha)}} C_{d,0}$, which is generally nonzero.

5 Discrete-time linear quadratic optimal control

For a bounded i/s/o node

$$S_d = \begin{bmatrix} A_d & B_d \\ C_d & D_d \end{bmatrix},$$

(which in our application will come from the Cayley transform) we consider the discrete-time dynamics

$$(x_d)_{n+1} = A_d(x_d)_n + B_d(x_d)_n, \qquad (y_d)_n = C_d(x_d)_n + D_d(u_d)_n, \tag{4}$$

i.e.

$$\begin{bmatrix} (x_d)_{n+1} \\ (y_d)_n \\ (x_d)_n \\ (u_d)_n \end{bmatrix} \in \mathrm{gph}(S_d), \tag{5}$$

(this should be compared to the continuous-time case (1), noting that in discrete-time the sense in which the equation should be understood is completely obvious). We call (x,y,u) a *discrete-time trajectory* if (5) is satisfied for all $n \in \mathbb{N}_0$.

For a discrete-time system we define for a given initial condition $x^0 \in \mathscr{X}$ the set of i/o stable discrete-time trajectories as consisting of those $u_d \in \ell^2(\mathbb{N}_0;\mathscr{U})$ and $y_d \in \ell^2(\mathbb{N}_0;\mathscr{Y})$ for which there exists a $x_d : \mathbb{N}_0 \to \mathscr{X}$ such that $(x_d)_0 = x^0$ and (4) (or equivalently (5)) is satisfied. If for all $x^0 \in \mathscr{X}$ this set is non-empty, then it is said that the *discrete-time finite cost condition* holds. The discrete-time linear quadratic optimal control problem is: for a given $x^0 \in \mathscr{X}$ find the i/o stable discrete-time trajectory with initial condition x^0 of minimal norm, i.e. minimize $\|u_d\|^2_{\ell^2(\mathbb{N}_0;\mathscr{U})} + \|y_d\|^2_{\ell^2(\mathbb{N}_0;\mathscr{Y})}$.

By standard discrete-time theory (see e.g. [14]), if the discrete-time finite cost condition is satisfied, then there exist K_d, L_d and X which satisfy the (Lur'e form of the) discrete-time Riccati equation:

$$\begin{aligned} A_d^* X A_d - X + C_d^* C_d &= K_d^* K_d, \\ B_d^* X B_d + D_d^* D_d + I &= L_d^* L_d, \\ B_d^* X A_d + D_d^* C_d &= L_d^* K_d, \end{aligned}$$

the optimal cost is given by $\langle X x^0, x^0 \rangle$ and the optimal control is given by

$$0 = K_d (x_d)_n + L_d (u_d)_n,$$

which noting that by the middle Lur'e equation, L_d has the left-inverse $(B_d^* X B_d + D_d^* D_d + I)^{-1} L_d^*$, can be explicitly written as

$$(u_d)_n = -(B_d^* X B_d + D_d^* D_d + I)^{-1} L_d^* K_d (x_d)_n = -(B_d^* X B_d + D_d^* D_d + I)^{-1} \left(B_d^* X A_d + D_d^* C_d\right)(x_d)_n.$$

Similarly, L_d and K_d can be eliminated from the Lur'e equations to obtain the standard form of the Riccati equation

$$A_d^* X A_d - X + C_d^* C_d - (C_d^* D_d + A_d^* X B_d)(B_d^* X B_d + D_d^* D_d + I)^{-1}(B_d^* X A_d + D_d^* C_d) = 0.$$

For our purposes it will be convenient to write the Lur'e form of the Riccati equation as

$$\langle z_d, X z_d \rangle - \langle x_d, X x_d \rangle + \|y_d\|^2 + \|u_d\|^2 = \|w_d\|^2, \quad \text{for all} \quad \begin{bmatrix} z_d \\ \begin{bmatrix} w_d \\ y_d \end{bmatrix} \\ x_d \\ u_d \end{bmatrix} \in \operatorname{gph}(S_d^{\text{ext}}), \tag{6}$$

where

$$\operatorname{gph}(S_d^{\text{ext}}) = \left\{ \begin{bmatrix} z_d \\ \begin{bmatrix} w_d \\ y_d \end{bmatrix} \\ x_d \\ u_d \end{bmatrix} : \begin{bmatrix} z_d \\ y_d \\ x_d \\ u_d \end{bmatrix} \in \operatorname{gph}(S_d),\ w_d = K_d x_d + L_d u_d \right\}.$$

Using Definition 2.10, S_d^{ext} is the standard output extension of S_d with observation extension $\begin{bmatrix} 0 & K_d \end{bmatrix}$ and feedthrough extension $\begin{bmatrix} 0 & L_d \end{bmatrix}$ (note that by [2, Lemma 6.2.1 (vii)] since S_d is bounded, any non-standard output extension is equivalent to a standard output extension).

Remark 5.1. *By the above, we can therefore formulate the solution of the discrete-time optimal control problem as follows: if S_d satisfies the discrete-time finite cost condition, then for all $x^0 \in \mathscr{X}$ a unique optimal control exists, the optimal cost is given by $\langle X x^0, x^0 \rangle$, S has a standard output extension S_d^{ext} with a feedthrough extension which has left-invertible standard part such that* (6) *holds and the optimal control is characterized by putting the additional output in S_d^{ext} equal to zero.*

6 The internal Cayley transform in linear quadratic optimal control

The crucial observation (utilized in [13] and in [15]) is that stable i/o trajectories in continuous- and discrete-time correspond to each other. Let $\mathscr{L}$ denote the Laplace transform and note that by the Paley–Wiener theorem for a Hilbert space $\mathscr{K}$ this is an isometric isomorphism between $L^2(\mathbb{R}_+;\mathscr{K})$ and the Hardy space $H^2(\mathbb{C}_+;\mathscr{K})$. The Z-transform $\mathscr{Z}$ maps a sequence $(h_n)_{n\in\mathbb{N}_0}$ to the corresponding formal power series $\sum_{n=0}^{\infty} h_n z^n$ and gives an isometric isomorphism between $\ell^2(\mathbb{N}_0;\mathscr{K})$ and the Hardy space of the disc $H^2(\mathbb{D};\mathscr{K})$. Finally, for $\alpha \in \mathbb{C}$ with $\mathrm{Re}(\alpha) > 0$, the linear fractional transformation

$$(F_\alpha g)(z) = \frac{\sqrt{\mathrm{Re}(2\alpha)}}{1+z}\, g\left(\frac{\alpha - \overline{\alpha} z}{1+z}\right), \qquad (F_\alpha^{-1} f)(\lambda) = \frac{\sqrt{\mathrm{Re}(2\alpha)}}{\overline{\alpha}+\lambda}\, f\left(\frac{\alpha-\lambda}{\overline{\alpha}+\lambda}\right),$$

gives an isometric isomorphism between the Hardy spaces $H^2(\mathbb{C}_+;\mathscr{K})$ and $H^2(\mathbb{D};\mathscr{K})$.

Remark 6.1. *In the above we use the discrete-time frequency domain variable z which in the stable case belongs to the unit disc. The discrete-time resolvent parameter w relates to this z though $w = \frac{1}{z}$. We could have written the above in terms of w by utilizing the Hardy space of the exterior of the unit disc.*

Lemma 6.2. *Let S be a future-resolvable i/s/o node. Let Ω be a non-empty connected open subset of $\rho(S) \cap \mathbb{C}_+$ and let $\alpha \in \Omega$. Let S_d be the Cayley transform of S with parameter α. Let $x^0 \in \mathscr{X}$. The set of i/o stable Ω trajectories of S and the set of i/o stable discrete-time trajectories of S_d, both with initial condition x^0, are isometrically isomorphic through the map $\mathscr{Z}^{-1} F_\alpha \mathscr{L}$.*

Proof. This is essentially contained in [15, Section 4.1] and also in [13, Theorem 6.5]. □

Proposition 6.3. *Let $X : \mathscr{X} \to \mathscr{X}$ be a bounded single-valued everywhere-defined self-adjoint operator. Let $\widetilde{S}$ be a closed i/s/o node with state space $\mathscr{X}$, input space $\mathscr{U}$ and output space $\left[\begin{smallmatrix} \mathscr{Y} \\ \mathscr{U} \end{smallmatrix}\right]$ and let $\widetilde{S}_d$ be its Cayley transform with parameter $\alpha \in \mathbb{C}$ with* $\mathrm{Re}(\alpha) > 0$. *The following are equivalent: (here $\widehat{\mathfrak{G}}$ is the formal i/s/o resolvent of $\widetilde{S}$ and $\lambda \in \mathbb{C}$)*

(i)

$$\langle z, Xx\rangle + \langle Xx, z\rangle + \|y\|^2 + \|u\|^2 = \|w\|^2, \quad \textit{for all} \quad \begin{bmatrix} z \\ \begin{bmatrix} w \\ y \end{bmatrix} \\ x \\ u \end{bmatrix} \in \operatorname{gph}(\widetilde{S});$$

(ii)

$$\langle \lambda\hat{z} - \hat{x}, X\hat{x}\rangle + \langle X(\lambda\hat{z} - \hat{x}), \hat{x}\rangle + \|\hat{y}\|^2 + \|\hat{u}\|^2 = \|\hat{w}\|^2, \quad \textit{for all} \quad \begin{bmatrix} \hat{z} \\ \begin{bmatrix} \hat{w} \\ \hat{y} \end{bmatrix} \\ \hat{x} \\ \hat{u} \end{bmatrix} \in \operatorname{gph}(\widehat{\mathfrak{G}}(\lambda));$$

(iii)

$$\langle z_d, Xz_d\rangle - \langle x_d, Xx_d\rangle + \|y_d\|^2 + \|u_d\|^2 = \|w_d\|^2, \quad \textit{for all} \quad \begin{bmatrix} z_d \\ \begin{bmatrix} w_d \\ y_d \end{bmatrix} \\ x_d \\ u_d \end{bmatrix} \in \operatorname{gph}(\widetilde{S}_d).$$

Proof. This follows easily from the relations between the graphs. As an example we show how the first equation implies the third in detail. Let

$$\begin{bmatrix} z_d \\ \begin{bmatrix} w_d \\ y_d \end{bmatrix} \\ x_d \\ u_d \end{bmatrix} \in \operatorname{gph}(\widetilde{S}_d).$$

Then there exists

$$\begin{bmatrix} z \\ \begin{bmatrix} w \\ y \end{bmatrix} \\ x \\ u \end{bmatrix} \in \operatorname{gph}(\widetilde{S}), \quad \text{such that} \quad \begin{bmatrix} z_d \\ \begin{bmatrix} w_d \\ y_d \end{bmatrix} \\ x_d \\ u_d \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2\operatorname{Re}(\alpha)}} & 0 & \frac{\overline{\alpha}}{\sqrt{2\operatorname{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{-1}{\sqrt{2\operatorname{Re}(\alpha)}} & 0 & \frac{\alpha}{\sqrt{2\operatorname{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z \\ \begin{bmatrix} w \\ y \end{bmatrix} \\ x \\ u \end{bmatrix}.$$

With this the third equation then is

$$\begin{aligned} &\left\langle \frac{1}{\sqrt{2\operatorname{Re}(\alpha)}} z + \frac{\overline{\alpha}}{\sqrt{2\operatorname{Re}(\alpha)}} x, X\left(\frac{1}{\sqrt{2\operatorname{Re}(\alpha)}} z + \frac{\overline{\alpha}}{\sqrt{2\operatorname{Re}(\alpha)}} x\right)\right\rangle \\ &\qquad - \left\langle \frac{-1}{\sqrt{2\operatorname{Re}(\alpha)}} z + \frac{\alpha}{\sqrt{2\operatorname{Re}(\alpha)}} x, X\left(\frac{-1}{\sqrt{2\operatorname{Re}(\alpha)}} z + \frac{\alpha}{\sqrt{2\operatorname{Re}(\alpha)}} x\right)\right\rangle \\ &\qquad\qquad + \|y\|^2 + \|u\|^2 = \|w\|^2. \end{aligned}$$

Simplifying this gives the first equation, which by assumption is satisfied, so that the third equation is satisfied. □

7 The algebraic Riccati equation for future-resolvable i/s/o nodes

Theorem 7.1. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty connected open subset of $\rho(S)\cap\mathbb{C}_+$. If the Ω finite cost condition is satisfied, then for every $x^0\in\mathscr{X}$ there exists a unique optimal control, there exists a bounded single-valued everywhere-defined self-adjoint operator $X:\mathscr{X}\to\mathscr{X}$ such that the optimal cost is given by $\langle Xx^0,x^0\rangle$, S has a non-standard output extension S^{ext} with a feedthrough extension which is standard and has left-invertible standard part and is such that*

$$\langle z,Xx\rangle+\langle Xx,z\rangle+\|y\|^2+\|u\|^2=\|w\|^2,\quad \textit{for all}\quad \begin{bmatrix} z\\ \begin{bmatrix} w\\ y\end{bmatrix}\\ x\\ u\end{bmatrix}\in\operatorname{gph}(S^{\text{ext}}); \tag{7}$$

holds and the optimal control is characterized by putting the additional output in S^{ext} equal to zero.

Proof. Let $\alpha\in\Omega$ and let S_d be the Cayley transform of S with parameter α. Since the Ω finite cost condition is satisfied for S, it follows from Lemma 6.2 that the discrete-time finite cost condition is satisfied for S_d. By Remark 5.1, a unique discrete-time optimal control exists, which by Lemma 6.2 transforms to a unique continuous-time optimal control. By Remark 5.1 the discrete-time optimal cost is given by $\langle Xx^0,x^0\rangle$, which by Lemma 6.2 is also the continuous-time optimal cost. By Remark 5.1, S_d has a standard output extension S_d^{ext} whose feedthrough extension has a left-invertible standard part such that the discrete-time Riccati equation (6) holds. By Remark 2.11 the resolvent set of S_d^{ext} is the same as that of S_d. Define S^{ext} as the inverse Cayley transform (with the same parameter α) of S_d^{ext}. By Remark 4.2, this is a non-standard output extension of S, the feedthrough extension is standard (i.e. $D_1=0$ in Remark 4.2 since $D_{d,1}=0$ as the discrete-time feedthrough extension is standard) and its standard part is left-invertible (since $D_0=D_{d,0}$ and the discrete-time feedthrough $D_{d,0}$ is left-invertible). By Remark 2.11 the resolvent set of S^{ext} is the same as that of S. By Proposition 6.3 (applied with $\widetilde{S}=S^{\text{ext}}$ and therefore $\widetilde{S}_d=S_d^{\text{ext}}$) we have that (7) holds (since (6) holds as we saw above). Since in discrete-time the optimal control is characterized by putting the additional output in S_d^{ext} equal to zero, the equivalent is true in continuous-time because these output are Cayley transforms of each other. □

The Riccati equation (7) is representation-independent in that we can substitute any representation of S into it: for example a kernel representation from Remark 2.2 or (if such a representation exists) the conventional DAE form from Example 2.3. A particularly attractive representation in this respect is an image representation.

By [2, Lemma 4.1.15 and Definition 4.1.16], every closed i/s/o node $\widetilde{S}$ has an *image representation*, i.e. there exist a Hilbert space $\mathscr{V}$ and bounded single-valued everywhere-defined operators $F:\mathscr{V}\to\mathscr{X}$, $L_{\text{out}}:\mathscr{V}\to\mathscr{Y}$, $K:\mathscr{V}\to\mathscr{X}$, $L_{\text{in}}:\mathscr{V}\to\mathscr{U}$

such that

$$\mathrm{gph}(\widetilde{S}) = \left\{ \begin{bmatrix} z \\ y \\ x \\ u \end{bmatrix} \in \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \\ \mathscr{X} \\ \mathscr{U} \end{bmatrix} : z = Fv,\ y = L_{\mathrm{out}}v,\ x = Kv,\ u = L_{\mathrm{in}}v \right\}.$$

Applying this result with $\widetilde{S} = S^{\mathrm{ext}}$ from Theorem 7.1, the Riccati equation (7) can be written as

$$F^*XK + K^*XF + L^*_{\mathrm{out,y}}L_{\mathrm{out,y}} + L^*_{\mathrm{in}}L_{\mathrm{in}} = L^*_{\mathrm{out,w}}L_{\mathrm{out,w}}, \tag{8}$$

and the optimal control is obtained from

$$L_{\mathrm{out,w}}v = 0.$$

Of course, it is not always easy to explictly find an image representation of a given DAE (e.g. in terms of the coefficients **A**, **B**, **C**, **D**, **E** of the conventional DAE form considered in Example 2.3), but once such a represenation is obtained, (8) immediately gives the appropriate Riccati equation very explicitly in terms of the operators appearing in the image representation.

8 An example

Example 8.1. *We return to the example of a conventional DAE from Example 2.3. In that case we have for a nonstandard output extension as in Theorem 7.1 with a feedthrough extension which is standard*

$$\mathrm{gph}(S^{\mathrm{ext}}) = \left\{ \begin{bmatrix} z \\ \begin{bmatrix} w \\ y \end{bmatrix} \\ x \\ u \end{bmatrix} : \mathbf{E}z = \mathbf{A}x + \mathbf{B}u, \quad y = \mathbf{C}x + \mathbf{D}u, \quad w = \mathbf{K}_1 z + \mathbf{K}_0 x + \mathbf{L}_0 u \right\}.$$

Note that setting $w = 0$ gives

$$\mathbf{K}_1 z + \mathbf{K}_0 x + \mathbf{L}_0 u = 0,$$

which since $\mathbf{L}_0$ is left-invertible in the context of Theorem 7.1 gives (here $\mathbf{L}_0^{-1}$ is a left-inverse of $\mathbf{L}_0$):

$$u = -\mathbf{L}_0^{-1}\mathbf{K}_1 z - \mathbf{L}_0^{-1}\mathbf{K}_0 x.$$

We note that this feedback depends both on x and z; for trajectories this means that the optimal control will be a feedback of both the state and its derivative (rather than only the state as is known to be the case in the purely differential equation situation).

Example 8.2. *We again consider the situation in Example 8.1. We show that in the purely differential equation situation, the conclusions of Theorem 7.1 reduce to the usual ones.*

Assume that $\mathbf{E}$ *is invertible (i.e. has a bounded single-valued everywhere-defined inverse). Then we can easily eliminate* z *and re-write the Riccati equation* (7) *as (here* $\mathbf{K} := \mathbf{K}_1\mathbf{E}^{-1}\mathbf{A} + \mathbf{K}_0$ *and* $\mathbf{L} := \mathbf{L}_0 + \mathbf{K}_1\mathbf{E}^{-1}\mathbf{B}$*)*

$$\langle \mathbf{E}^{-1}\mathbf{A}x + \mathbf{E}^{-1}\mathbf{B}u, Xx\rangle + \langle Xx, \mathbf{E}^{-1}\mathbf{A}x + \mathbf{E}^{-1}\mathbf{B}u\rangle + \|\mathbf{C}x + \mathbf{D}u\|^2 + \|u\|^2 = \|\mathbf{K}x + \mathbf{L}_0 u\|^2.$$

Since $u \in \mathscr{U}$ *and* $x \in \mathscr{X}$ *are arbitrary, this can be written as the Lur'e equations*

$$\begin{aligned} \mathbf{A}^*\mathbf{E}^{-*}X + X\mathbf{E}^{-1}\mathbf{A} + \mathbf{C}^*\mathbf{C} &= \mathbf{K}^*\mathbf{K}, \\ \mathbf{D}^*\mathbf{D} + I &= \mathbf{L}^*\mathbf{L}, \\ \mathbf{B}^*\mathbf{E}^{-1}X + \mathbf{D}^*\mathbf{C} &= \mathbf{L}^*\mathbf{K}, \end{aligned}$$

which can in turn be written as the standard algebraic Riccati equation

$$\mathbf{A}^*\mathbf{E}^{-*}X + X\mathbf{E}^{-1}\mathbf{A} + \mathbf{C}^*\mathbf{C} - (X\mathbf{E}^{-*}\mathbf{B} + \mathbf{C}^*\mathbf{D})(\mathbf{D}^*\mathbf{D} + I)^{-1}(\mathbf{B}^*\mathbf{E}^{-1}X + \mathbf{D}^*\mathbf{C}) = 0.$$

The optimal control is characterized by $0 = \mathbf{K}_1 z + \mathbf{K}_0 x + \mathbf{L}_0 u$*, which is*

$$0 = \mathbf{K}x + \mathbf{L}u.$$

From this we obtain

$$u = -(\mathbf{D}^*\mathbf{D} + I)^{-1}\left(\mathbf{B}^*\mathbf{E}^{-1}X + \mathbf{D}^*\mathbf{C}\right)x.$$

Hence we have that the conclusions of Theorem 7.1 become the usual ones.

We note that [15] considers a more general version of this "usual" Riccati equation which allows for example for the solution X *to be unbounded (this relates to not every initial condition having a finite cost).*

Example 8.3. *We consider an example from [5] (see also [3, 11, 7]). Although this is a seemingly simple finite-dimensional example, it is interesting since the "naive" algebraic Riccati equation*

$$\mathbf{A}^*Z\mathbf{E} + \mathbf{E}^*Z\mathbf{A} + \mathbf{C}^*\mathbf{C} - (\mathbf{E}^*Z\mathbf{B} + \mathbf{C}^*\mathbf{D})(I + \mathbf{D}^*\mathbf{D})^{-1}(\mathbf{B}^*Z\mathbf{E} + \mathbf{D}^*\mathbf{C}) = 0,$$

does not have a solution (see [3]). The example is of the conventional DAE form from Example 2.3 with

$$\mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Here

$$\lambda\mathbf{E} - \mathbf{A} = \begin{bmatrix} \lambda & -1 \\ -1 & 0 \end{bmatrix}, \qquad (\lambda\mathbf{E} - \mathbf{A})^{-1} = \begin{bmatrix} 0 & -1 \\ -1 & -\lambda \end{bmatrix},$$

so that the i/s/o resolvent set equals $\mathbb{C}$ *(which implies that the choice of* Ω *is immaterial; we choose* Ω *equal to the open right half-plane* $\mathbb{C}_+$ *since this works nicely with Laplace transforms). We further have*

$$\operatorname{gph}(S) = \left\{ \begin{bmatrix} z \\ y \\ x \\ u \end{bmatrix} : z_1 = x_2,\ 0 = x_1 + u,\ y = x \right\},$$

and

$$\widehat{\mathfrak{A}}(\lambda) = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, \qquad \widehat{\mathfrak{B}}(\lambda) = \begin{bmatrix} -1 \\ -\lambda \end{bmatrix}, \qquad \widehat{\mathfrak{C}}(\lambda) = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, \qquad \widehat{\mathfrak{D}}(\lambda) = \begin{bmatrix} -1 \\ -\lambda \end{bmatrix}.$$

Therefore the equations for Ω trajectories from Remark 2.9 are

$$\hat{y}(\lambda) = \hat{x}(\lambda) = \begin{bmatrix} -\hat{u}(\lambda) \\ -x_1^0 - \lambda \hat{u}(\lambda) \end{bmatrix}.$$

From this we see that the Ω finite-cost condition is satisfied. We can choose $u(t) = -x_1^0 \mathrm{e}^{-t}$ (which is in $L^2(\mathbb{R}^+)$) so that

$$\hat{u} = \frac{-x_1^0}{\lambda + 1}, \qquad \hat{y} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \frac{x_1^0}{\lambda + 1},$$

and we see that $\hat{y}$ is the Laplace transform of

$$y(t) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} x_1^0 \mathrm{e}^{-t},$$

which is in $L^2(\mathbb{R}^+)$.

In this example it is easy to obtain an image representation (the idea is that z_2, x_2 and u can be chosen as "free" variables which determine the others):

$$F = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad L_{\text{out,y}} = K = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad L_{\text{in}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$$

Writing (using that X is symmetric):

$$X = \begin{bmatrix} X_1 & X_0 \\ X_0 & X_2 \end{bmatrix}, \qquad L_{\text{out,w}} = \begin{bmatrix} L_1 & L_2 & L_3 \end{bmatrix},$$

the image representation form (8) *of the Riccati equation is*

$$\begin{bmatrix} 0 & X_2 & -X_0 \\ X_2 & 2X_0 + 1 & -X_1 \\ -X_0 & -X_1 & 2 \end{bmatrix} = \begin{bmatrix} L_1^2 & L_1 L_2 & L_1 L_3 \\ L_1 L_2 & L_2^2 & L_2 L_3 \\ L_1 L_3 & L_2 L_3 & L_3^2 \end{bmatrix}.$$

From the top-left entry we obtain $L_1 = 0$, from which we deduce using the first row that $X_0 = X_2 = 0$. The bottom right 2-by-2 matrix then is

$$\begin{bmatrix} 1 & -X_1 \\ -X_1 & 2 \end{bmatrix} = \begin{bmatrix} L_2^2 & L_2 L_3 \\ L_2 L_3 & L_3^2 \end{bmatrix}.$$

From the diagonal entries we then obtain $L_2 = \pm 1$ and $L_3 = \pm\sqrt{2}$. Since $X_1 \geq 0$ (because X is positive semi-definite) L_2 and L_3 must have opposite signs. We choose $L_2 = 1$ and $L_3 = -\sqrt{2}$ (this sign choice is immaterial) and obtain

$$X = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}, \qquad L = \begin{bmatrix} 0 & 1 & -\sqrt{2} \end{bmatrix}.$$

Hence the optimal cost is $\langle Xx^0, x^0\rangle = \sqrt{2}(x_1^0)^2$ *and the optimal control is determined by*

$$\begin{bmatrix}0 & 1 & -\sqrt{2}\end{bmatrix}\begin{bmatrix}z_2\\ x_2\\ u\end{bmatrix} = 0,$$

i.e.

$$u = \frac{1}{\sqrt{2}}x_2,$$

as is also obtained in the above references using various different methods. Note that we can equivalently write this as the output feedback

$$u = \frac{1}{\sqrt{2}}y_2.$$

Example 8.4. *The computations in Example 8.3 can be simplified by choosing a smaller state space. To write the relation between the input, the (relevant part of the) initial state and the output in conventional DAE form, a two-dimensional state is needed. However, because i/s/o nodes are more general than conventional DAEs, it is possible to described the same relation using a one-dimensional state space using i/s/o nodes. Consider the i/s/o node with* $\mathscr{X}$ *and* $\mathscr{U}$ *one-dimensional and* $\mathscr{Y}$ *two-dimensional given by*

$$\operatorname{gph}(S) = \left\{\begin{bmatrix}z\\ y\\ x\\ u\end{bmatrix} : z = y_2,\ y_1 = -u,\ x = -u\right\}.$$

We have that the i/s/o resolvent set equals $\mathbb{C}$ *and that*

$$\widehat{\mathfrak{A}}(\lambda) = 0, \qquad \widehat{\mathfrak{B}}(\lambda) = -1, \qquad \widehat{\mathfrak{C}}(\lambda) = \begin{bmatrix}0\\ -1\end{bmatrix}, \qquad \widehat{\mathfrak{D}}(\lambda) = \begin{bmatrix}-1\\ -\lambda\end{bmatrix}.$$

We in particular see that $\widehat{\mathfrak{D}}$ *is the same as in Example 8.3 and that* $\widehat{\mathfrak{C}}x_1 = \widehat{\underline{\mathfrak{C}}}\,\mathbf{E}\left[\begin{smallmatrix}x_1\\ x_2\end{smallmatrix}\right]$*, where* $\widehat{\underline{\mathfrak{C}}}$ *is the state/output resolvent from Example 8.3. Since* $\mathbf{E}\left[\begin{smallmatrix}x_1\\ x_2\end{smallmatrix}\right]$ *is the "relevant part" of the initial state* $\left[\begin{smallmatrix}x_1\\ x_2\end{smallmatrix}\right]$*, we see that the state/output resolvents therefore also essentially coincide. Hence the i/s/o node from this example and the one from Example 8.3 are from the linear quadratic optimal control perspective equivalent.*

A kernel representation (as in Remark 2.2) is (here the space $\mathscr{X}$ *is three-dimensional)*

$$E = \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}, \qquad N_{\text{out}} = \begin{bmatrix}0 & -1\\ 1 & 0\\ 0 & 0\end{bmatrix}, \qquad M = \begin{bmatrix}0\\ 0\\ 1\end{bmatrix}, \qquad N_{\text{in}} = \begin{bmatrix}0\\ -1\\ 1\end{bmatrix}.$$

An image representation is (the idea is that y_2 *and* u *are "free")*

$$F = \begin{bmatrix}1 & 0\end{bmatrix}, \qquad L_{\text{out,y}} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}, \qquad K = \begin{bmatrix}0 & -1\end{bmatrix}, \qquad L_{\text{in}} = \begin{bmatrix}0 & 1\end{bmatrix}.$$

The image representation form (8) *of the Riccati equation is*

$$\begin{bmatrix} 1 & -X \\ -X & 2 \end{bmatrix} = \begin{bmatrix} L_1^2 & L_1 L_2 \\ L_1 L_2 & L_2^2 \end{bmatrix}.$$

From this we obtain that L_1 and L_2 must have opposite signs (since their product equals $-X$) and picking an arbitrary sign convention gives $L_1 = -1$, $L_2 = \sqrt{2}$, $X = \sqrt{2}$. Hence the optimal cost is $\sqrt{2}(x^0)^2$ and the optimal control is determined by

$$-y_2 + \sqrt{2}u = 0,$$

i.e.

$$u = \frac{1}{\sqrt{2}} y_2.$$

This solution is consistent with what we obtained in Example 8.3 noting that x^0 here corresponds to x_1^0 there.

Note that the above cannot be written as a state feedback. Equivalently, the optimal output extension is necessarily nonstandard: we can re-write the equation determining the optimal control as

$$-z + \sqrt{2}u = 0.$$

Therefore the observation extension is $\begin{bmatrix} -1 & 0 \end{bmatrix}$ (i.e. $C_1 = -1$ and $C_0 = 0$, so this is nonstandard) and the feedthrough extension is $\begin{bmatrix} 0 & 0 & \sqrt{2} \end{bmatrix}$ (which is standard). Alternatively, from

$$-y_2 + \sqrt{2}u = 0,$$

we have the standard observation extension $\begin{bmatrix} 0 & 0 \end{bmatrix}$ and the nonstandard feedthrough extension $\begin{bmatrix} 0 & -1 & \sqrt{2} \end{bmatrix}$ (i.e. $D_1 = \begin{bmatrix} 0 & -1 \end{bmatrix}$ and $D_0 = \sqrt{2}$). We however cannot have both a standard observation extension and a standard feedthrough extension at the same time.

The construction of the i/s/o node in Example 8.4 from that in Example 8.3 can be done generally based on obtaining a minimal i/s/o node with the same input/output resolvent as a given i/s/o node, see [2].

Example 8.5. *It is possible to further simplify Example 8.4 by considering y_2 as the input and y_1 and u as the outputs (this is related to the notion of* canonical input space *from [2, Definition 2.1.23]). We therefore define*

$$\widetilde{u} = y_2, \qquad \widetilde{y} = \begin{bmatrix} y_1 \\ u \end{bmatrix}.$$

This change in perspective does not alter the cost function (or the state), but the dynamics instead become

$$\dot{x} = \widetilde{u}, \qquad \widetilde{y} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} x,$$

i.e. we have a standard state-space system with ($E = 1$ and)

$$A = 0, \qquad B = 1, \qquad C = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The usual Riccati equation is $2 - X^2 = 0$ and gives $X = \sqrt{2}$ and the state feedback $F = -\sqrt{2}$. The optimal control is therefore determined by

$$\widetilde{u} = -\sqrt{2}x,$$

which noting that $\widetilde{u} = y_2$ and $x = -\widetilde{y}_2 = -u$ gives

$$y_2 = \sqrt{2}u,$$

which is the same as what was obtained before.

9 Beyond the finite cost condition

To consider the case where not all initial conditions are required to have finite cost, we need some further notions from [2].

Definition 9.1. *Let S be an i/s/o node and let $\mathscr{X}_1 \subset \mathscr{X}$ be continuously embedded. Then the* part S_{part} *of S in $\mathscr{X}_1$ is defined by*

$$\mathrm{gph}(S_{\mathrm{part}}) = \mathrm{gph}(S) \cap \begin{bmatrix} \mathscr{X}_1 \\ \mathscr{Y} \\ \mathscr{X}_1 \\ \mathscr{U} \end{bmatrix}.$$

Definition 9.1 is adapted from [2, Definition 5.1.10 (i)] where it was assumed that $\mathscr{X}_1$ is a closed subspace of $\mathscr{X}$.

Remark 9.2. *It is easy to see that the Cayley transform of the part of S in $\mathscr{X}_1$ equals the part in $\mathscr{X}_1$ of the Cayley transform of S (i.e. taking the Cayley transform and taking the part commute).*

9.1 The discrete-time case

We reconsider the situation of a bounded i/s/o node S_d with discrete-time dynamics from Section 5.

Definition 9.3. *The initial state $x^0 \in \mathscr{X}$ is said to have* discrete-time finite cost *if the corresponding set of i/o stable trajectories is non-empty. We denote the subspace of discrete-time finite cost initial states by $\mathscr{X}_{d,\mathrm{finite}}$.*

By [14, Section 2], for every element x^0 of $\mathscr{X}_{d,\mathrm{finite}}$, a unique minimal norm i/o stable trajectory $(u_d^{\min}, y_d^{\min})$ exists. This defines a closed nonnegative sesquilinear symmetric form q on $\mathscr{X}$ with domain $\mathscr{X}_{d,\mathrm{finite}}$ given by

$$q[x_1^0, x_2^0] := \left\langle \begin{bmatrix} u_{d,1}^{\min} \\ y_{d,1}^{\min} \end{bmatrix}, \begin{bmatrix} u_{d,2}^{\min} \\ y_{d,2}^{\min} \end{bmatrix} \right\rangle,$$

which we call the *discrete-time optimal cost sequilinear form*. We will consider $\mathscr{X}_{d,\text{finite}}$ with the inner-product

$$\langle x_1^0, x_2^0 \rangle_{\mathscr{X}_{d,\text{finite}}} := \langle x_1^0, x_2^0 \rangle_{\mathscr{X}} + q[x_1^0, x_2^0],$$

(this is called the graph inner product in [14, Section 4]). With this, $\mathscr{X}_{d,\text{finite}}$ is a Hilbert space which is continuously embedded in $\mathscr{X}$.

Definition 9.4. *The bounded i/s/o node S_d satisfies the* discrete-time input finite future cost condition *if all initial states in* $\operatorname{im}(B_d)$ *have discrete-time finite cost, i.e. if* $\operatorname{im}(B_d) \subset \mathscr{X}_{d,\text{finite}}$.

Remark 9.5. *The above definition is adapted from [14, Definition 3.3] where the equivalent concept was called the* finite future incremental cost condition*. See [14, Lemma 3.4] for this equivalence and for the fact that $A_d \mathscr{X}_{d,\text{finite}} \subset \mathscr{X}_{d,\text{finite}}$.*

Definition 9.6. *Let S and S_1 be two bounded i/s/o nodes with the same input and output spaces and where $\mathscr{X}_1 \subset \mathscr{X}$. We call S_1 a* restriction *of S to $\mathscr{X}_1$ if the following two conditions hold*

(i) Every discrete-time trajectory of S_1 is also a discrete-time trajectory of S;

(ii) If (x, y, u) is a discrete-time trajectory of S with $x(0) \in \mathscr{X}_1$, then $x_n \in \mathscr{X}_1$ for all $n \in \mathbb{N}$ and (x, y, u) is a discrete-time trajectory of S_1.

Definition 9.6 is adapted from [2, Definition 5.4.37] (which is in continuous-time) as indicated in [2, Definition 6.5.7]. Moreover, we do not assume that $\mathscr{X}_1$ is a closed subspace of $\mathscr{X}$ as was done in [2, Definition 5.4.37].

Lemma 9.7. *Let S_d be a bounded i/s/o node and let $\mathscr{X}_1 \subset \mathscr{X}$ be continuously embedded. The following are equivalent:*

(i) S_d has a discrete-time restriction to $\mathscr{X}_1$;

(ii) S_d has a unique discrete-time restriction to $\mathscr{X}_1$;

(iii) $A_d \mathscr{X}_1 \subset \mathscr{X}_1$ and $\operatorname{im}(B_d) \subset \mathscr{X}_1$;

(iv) The part of S_d in $\mathscr{X}_1$ is bounded.

If these equivalent conditions hold, then the part of S_d in $\mathscr{X}_1$ is the unique discrete-time restriction to $\mathscr{X}_1$ from (ii).

Proof. This is part of [2, Theorem 6.3.19] (the discrete-time version of which holds by [2, Lemma 6.5.8]). There is was assumed that $\mathscr{X}_1 \subset \mathscr{X}$ is closed (rather than continuously embedded), but the result remains true (with essentially the same proof) under our weaker assumption. □

If the discrete-time input finite future cost condition holds for S_d, then with $\mathscr{X}_1 := \mathscr{X}_{d,\text{finite}}$ the condition (iii) in Lemma 9.7 is satisfied by Remark 9.5. Therefore by Lemma 9.7 we can restrict S_d to the space of discrete-time finite cost states. Conversely, we see from Lemma 9.7 that the discrete-time input finite future cost condition (which is implied by (iii)) is necessary for the restriction of S_d to the space of discrete-time finite cost states to make sense.

Definition 9.8. *Let S_d be a bounded i/s/o node. Let q be a closed nonnegative sesquilinear symmetric form on $\mathscr{X}$ with domain $\mathscr{X}_1$ and equip $\mathscr{X}_1$ with the inner-product $q[x_1,x_2]+\langle x_1,x_2\rangle_{\mathscr{X}}$ (so that $\mathscr{X}_1$ is a Hilbert space which is continuously embedded in $\mathscr{X}$). We say that q satisfies the* discrete-time Riccati equation *for S_d if*

(i) the part $S_{d,\text{part}}$ of S_d in $\mathscr{X}_1$ is bounded;

(ii) $S_{d,\text{part}}$ has a standard output extension $S^{\text{ext}}_{d,\text{part}}$ with a feedthrough extension which has left-invertible standard part and is such that

$$q[z_d,z_d]-q[x_d,x_d]+\|y_d\|^2+\|u_d\|^2=\|w_d\|^2, \quad \textit{for all} \quad \begin{bmatrix} z_d \\ \begin{bmatrix} w_d \\ y_d \end{bmatrix} \\ x_d \\ u_d \end{bmatrix} \in \operatorname{gph}(S^{\text{ext}}_{d,\text{part}}). \tag{9}$$

Sesquilinear symmetric forms can be ordered as follows: $q_1 \le q_2$ means that $\operatorname{dom}(q_1) \supset \operatorname{dom}(q_2)$ and $q_1[x,x] \le q_2[x,x]$ for all $x \in \operatorname{dom}(q_2)$.

Theorem 9.9. *Let S_d be a bounded i/s/o node for which the discrete-time input finite cost condition holds. Then the discrete-time optimal cost sequilinear form is the smallest solution of the discrete-time Riccati equation. The optimal control is characterized by putting the additional output in $S^{\text{ext}}_{d,\text{part}}$ equal to zero.*

Proof. This is [14, Theorem 3.14] once translated to the current terminology. □

9.2 The continuous-time case

Definition 9.10. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty open subset of $\rho(S)\cap\mathbb{C}_+$. The initial state $x^0 \in \mathscr{X}$ is said to have Ω* finite cost *if the corresponding set of i/o stable Ω trajectories is non-empty. We denote the subspace of Ω finite cost initial states by $\mathscr{X}_{\text{finite}}$.*

By [15, Section 3.1], for every element x^0 of $\mathscr{X}_{\text{finite}}$, a unique minimal norm i/o stable Ω trajectory $(u^{\min},y^{\min})$ exists. This defines a closed nonnegative sesquilinear symmetric form q on $\mathscr{X}$ with domain $\mathscr{X}_{\text{finite}}$ given by

$$q[x_1^0,x_2^0] := \left\langle \begin{bmatrix} u_1^{\min} \\ y_1^{\min} \end{bmatrix}, \begin{bmatrix} u_2^{\min} \\ y_2^{\min} \end{bmatrix} \right\rangle,$$

which we call the Ω *optimal cost sequilinear form*. We will consider $\mathscr{X}_{\text{finite}}$ with the inner-product

$$\langle x_1^0,x_2^0\rangle_{\mathscr{X}_{\text{finite}}} := \langle x_1^0,x_2^0\rangle_{\mathscr{X}} + q[x_1^0,x_2^0].$$

With this, $\mathscr{X}_{\text{finite}}$ is a Hilbert space which is continuously embedded in $\mathscr{X}$.

Remark 9.11. *If Ω is connected, $\alpha\in\Omega$ and S_d is the Cayley transform with parameter α of S, then by Lemma 6.2 we have that the Ω finite cost initial states for S and the discrete-time finite cost initial states of S_d are the same, i.e. that $\mathscr{X}_{\text{finite}} = \mathscr{X}_{d,\text{finite}}$.*

Definition 9.12. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty open subset of $\rho(S)\cap\mathbb{C}_+$. Then S satisfies the Ω* input finite future cost condition *if for all $\lambda\in\Omega$, all initial states in $\operatorname{im}(\widehat{\mathfrak{B}}(\lambda))$ have Ω finite cost.*

Remark 9.13. *The above definition is adapted from [15, Definition 5.7] where the corresponding concept is defined with respect to a fixed $\alpha\in\Omega$ (by [15, Theorem 5.9] this is equivalent to it holding for all $\lambda\in\Omega$ if Ω is connected).*

Lemma 9.14. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty connected open subset of $\rho(S)\cap\mathbb{C}_+$. Then $\widehat{\mathfrak{A}}(\lambda)\mathscr{X}_{\text{finite}}\subset\mathscr{X}_{\text{finite}}$ for all $\lambda\in\Omega$.*

Proof. Let $\alpha\in\Omega$ and let S_d be the Cayley transform of S with parameter α. By Remark 9.11 we have $\mathscr{X}_{\text{finite}}=\mathscr{X}_{d,\text{finite}}$. By Remark 9.5 we have $A_d\mathscr{X}_{d,\text{finite}}\subset\mathscr{X}_{d,\text{finite}}$, which by (2) is equivalent to $\widehat{\mathfrak{A}}(\alpha)\mathscr{X}_{\text{finite}}\subset\mathscr{X}_{\text{finite}}$. Since $\alpha\in\Omega$ was arbitrary, we get the desired result. □

Definition 9.15. *Let S_1 and S be two i/s/o nodes with the same input and output spaces and where $\mathscr{X}_1\subset\mathscr{X}$ and let Ω be a non-empty open subset of $\mathbb{C}$. We call S_1 an Ω-*restriction *of S if the following two conditions hold*

(i) Every frequency domain Ω trajectory of S_1 is also a frequency domain Ω trajectory of S;

(ii) If $(\hat{x},\hat{y},x^0,\hat{u})$ is a frequency domain Ω trajectory of S with $x^0\in\mathscr{X}_1$, then $\hat{x}(\lambda)\in\mathscr{X}_1$ for all $\lambda\in\Omega$ and $(\hat{x},\hat{y},x^0,\hat{u})$ is a frequency domain Ω trajectory of S_1.

Definition 9.15 is adapted from [2, Definition 11.1.44] where it was assumed that $\mathscr{X}_1$ is a closed subspace of $\mathscr{X}$.

Lemma 9.16. *Let S be a resolvable i/s/o node, let $\Omega\subset\rho(S)$ and let $\mathscr{X}_1\subset\mathscr{X}$ be continuously embedded. Then the following are equivalent:*

(i) S has an Ω-restriction to $\mathscr{X}_1$;

(ii) S has a unique resolvable Ω-restriction to $\mathscr{X}_1$ whose resolvent set includes Ω;

(iii) $\widehat{\mathfrak{A}}(\lambda)\mathscr{X}_1\subset\mathscr{X}_1$ and $\widehat{\mathfrak{B}}(\lambda)\mathscr{U}\subset\mathscr{X}_1$ for all $\lambda\in\Omega$;

(iv) S_{part} is resolvable and $\Omega\subset\rho(S_{\text{part}})$.

If these equivalent conditions hold, then the part of S in $\mathscr{X}_1$ is the unique Ω-restriction to $\mathscr{X}_1$ from (ii).

Proof. This is part of [2, Theorem 11.1.51]. There is was assumed that $\mathscr{X}_1\subset\mathscr{X}$ is closed (rather than continuously embedded), but the result remains true (with essentially the same proof) under our weaker assumption. □

If the Ω input finite future cost condition holds for S and Ω is connected, then with $\mathscr{X}_1:=\mathscr{X}_{\text{finite}}$ the condition (iii) in Lemma 9.16 is satisfied by Lemma 9.14. Therefore by Lemma 9.16 we can restrict S to the space of Ω finite cost states. Conversely, we see from Lemma 9.16 that the Ω input finite future cost condition (which is implied by (iii)) is necessary.

Definition 9.17. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty open subset of $\rho(S)\cap\mathbb{C}_+$. Let q be a closed nonnegative sesquilinear symmetric form on $\mathscr{X}$ with domain $\mathscr{X}_1$ and equip $\mathscr{X}_1$ with the inner-product $q[x_1,x_2]+\langle x_1,x_2\rangle_{\mathscr{X}}$ (so that $\mathscr{X}_1$ is a Hilbert space which is continuously embedded in $\mathscr{X}$). We say that q satisfies the* Ω-Riccati equation *for S if*

(i) the part S_{part} of S in $\mathscr{X}_1$ is resolvable and satisfies $\Omega\subset\rho(S_{\text{part}})$;

(ii) S_{part} has a nonstandard output extension $S_{\text{part}}^{\text{ext}}$ with a feedthrough extension which is standard and has left-invertible standard part and is such that

$$q[z,x]+q[x,z]+\|y\|^2+\|u\|^2=\|w\|^2,\quad \textit{for all}\quad \begin{bmatrix} z\\ \begin{bmatrix} w\\ y\end{bmatrix}\\ x\\ u\end{bmatrix}\in\operatorname{gph}(S_{\text{part}}^{\text{ext}});\tag{10}$$

Theorem 9.18. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty connected open subset of $\rho(S)\cap\mathbb{C}_+$ for which the Ω input finite future cost condition holds. Then the Ω optimal cost sequilinear form is the smallest solution of the Ω-Riccati equation. The optimal control is characterized by putting the additional output in $S_{\text{part}}^{\text{ext}}$ equal to zero.*

Proof. This is proven similarly to Theorem 7.1 using Remark 9.2. □

9.3 Impulse controllability

To relate our results to available results on Riccati equations for DAEs, we briefly discuss the concept of impulse controllability in the i/s/o node framework.

Definition 9.19. *The* classical state space $\mathscr{X}_0$ *of the i/s/o node S equals*

$$\mathscr{X}_0:=\left\{x\in\mathscr{X}:\exists z,y,u \text{ such that } \begin{bmatrix} z\\ y\\ x\\ u\end{bmatrix}\in\operatorname{gph}(S)\right\}=\left\{x\in\mathscr{X}:\exists u \text{ such that } \begin{bmatrix} x\\ u\end{bmatrix}\in\operatorname{dom}(S)\right\}.$$

The above definition is adapted from [2, Definition 2.1.15].

Remark 9.20. *We have for $\lambda\in\mathbb{C}$*

$$\begin{aligned}\mathscr{X}_0&=\left\{x\in\mathscr{X}:\exists z,y,u \text{ such that } \begin{bmatrix} x\\ y\\ -z-\lambda\\ u\end{bmatrix}\in\operatorname{gph}(\widehat{\mathfrak{G}}(\lambda))\right\}\\ &=\left\{x\in\mathscr{X}:\exists \widetilde{z},y,u \text{ such that } \begin{bmatrix} x\\ y\\ \widetilde{z}\\ u\end{bmatrix}\in\operatorname{gph}(\widehat{\mathfrak{G}}(\lambda))\right\}\\ &=\operatorname{im}\begin{bmatrix}\widehat{\mathfrak{A}}(\lambda) & \widehat{\mathfrak{B}}(\lambda)\end{bmatrix}.\end{aligned}$$

Remark 9.21. *From [4, Remark 4.1] for a finite-dimensional conventional DAE with* $\det(s\mathbf{E}-\mathbf{A})$ *not the zero polynomial (i.e. the corresponding i/s/o node is resolvable), the DAE being* controllable at infinity *is equivalent to* $\operatorname{im}\begin{bmatrix}\mathbf{E} & \mathbf{B}\end{bmatrix} = \mathscr{X}$. *This in turn is equivalent to* $\operatorname{im}\begin{bmatrix}\widehat{\mathfrak{A}}(\lambda) & \widehat{\mathfrak{B}}(\lambda)\end{bmatrix} = \mathscr{X}$. *Combining this with Remark 9.20 we see that the condition* $\mathscr{X}_0 = \mathscr{X}$ *coincides with controllability at infinity.*

Definition 9.22. *The* multivalued part $\mathscr{Z}_0$ *of the i/s/o node S equals*

$$\mathscr{Z}_0 := \left\{ z \in \mathscr{X} : \begin{bmatrix} z \\ 0 \\ 0 \\ 0 \end{bmatrix} \in \operatorname{gph}(S) \right\}.$$

The above definition is adapted from [2, Definition 2.1.15].

We have for $\lambda \in \mathbb{C}$

$$\mathscr{Z}_0 = \left\{ x \in \mathscr{X} : \begin{bmatrix} 0 \\ 0 \\ x \\ 0 \end{bmatrix} \in \operatorname{gph}(\widehat{\mathfrak{G}}(\lambda)) \right\} = N\left(\begin{bmatrix} \widehat{\mathfrak{A}}(\lambda) \\ \widehat{\mathfrak{C}}(\lambda) \end{bmatrix} \right).$$

From this we see that $x^0 \in \mathscr{Z}_0$ precisely when $(0,0,x^0,0)$ is a frequency domain Ω trajectory of the i/s/o node. In particular, every initial condition in $\mathscr{Z}_0$ has zero optimal cost.

For a conventional DAE we have $\mathscr{Z}_0 = N(\mathbf{E})$. Therefore for a conventional DAE we have $\mathscr{Z}_0 = N(\widehat{\mathfrak{A}}(\lambda))$ (which is not true for a general i/s/o node). From [4, Remark 4.6] for a finite-dimensional conventional DAE with $\det(s\mathbf{E}-\mathbf{A})$ not the zero polynomial (i.e. the corresponding i/s/o node is resolvable), the DAE being *impulse controllable* is equivalent to $\operatorname{im}\begin{bmatrix}\widehat{\mathfrak{A}}(\lambda) & \widehat{\mathfrak{B}}(\lambda)\end{bmatrix} + N(\widehat{\mathfrak{A}}(\lambda)) = \mathscr{X}$. Using also Remark 9.21, this is $\mathscr{X}_0 + \mathscr{Z}_0 = \mathscr{X}$. Since for a general i/s/o node this notion also involves the output (for example though the state/output resolvent $\widehat{\mathfrak{C}}(\lambda)$ which appears in the characterization of $\mathscr{Z}_0$), referring to this as a controllability property would in general be a misnomer.

For an input/state system (i.e. with output space $\mathscr{Y} = \{0\}$) we have that $\mathscr{X}_0 + \mathscr{Z}_0 = \mathscr{X}$ is equivalent to $\operatorname{im}\begin{bmatrix}\widehat{\mathfrak{A}}(\lambda) & \widehat{\mathfrak{B}}(\lambda)\end{bmatrix} + N(\widehat{\mathfrak{A}}(\lambda)) = \mathscr{X}$, i.e. to impulse controllability of the input/state system.

9.4 Comparison with [17]

We compare Theorem 9.18 to the Riccati equation obtained in [17]. We first note that [17] is for finite-dimensional conventional DAEs whereas our results hold more generally for (possibly infinite-dimensional) i/s/o nodes. Another difference is that [17] considers the linear quadratic optimal control problem “with state stability” whereas we do not consider stability of the state.

In [17] the Riccati equation is considered on the “system space” $\mathscr{V}_{\text{sys}}$. Using [16, Proposition 3.3] we can describe this space inductively in i/s/o node terms. Let $S_1 = S$ and define S_{k+1} as the part of S_k in the classical state space of S_k. It follows from [16, Proposition 3.3] that (for a finite-dimensional conventional DAE) there exists a k_0 such that $S_k = S_{k_0}$ for all $k \geq k_0$ and that $\mathrm{dom}(S_{k_0}) = \mathscr{V}_{\mathrm{sys}}$. The restriction of S to $\mathscr{V}_{\mathrm{sys}}$ has graph (here $\mathscr{Z}_0$ is the multi-valued part of S)

$$\mathrm{gph}(S_{k_0}) + \begin{bmatrix} \mathscr{Z}_0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

In [17] it is this restriction which is considered rather than the part S_{k_0}. However, since elements in the second component of the above sum contribute zero terms to the Riccati equation (10), this difference is immaterial.

By [16, Proposition 2.9], without loss of generality, a conventional DAE can be considered in feedback equivalence form. Focusing on the non-differential part only, and making the dimension the smallest integer large enough so that the system is not impulse controllable (in the sense of [4, Definition 2.1]) we arrive at the examples in Section 9.5 which further illustrate the connection with [17].

9.5 An example

Example 9.23. *Consider the conventional DAE (with no output, i.e. $\mathscr{Y} = \{0\}$) with $\mathscr{X} = \mathbb{R}^3$ and $\mathscr{U} = \mathbb{R}$*

$$\mathbf{E} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

This gives

$$\mathrm{gph}(S) = \left\{ \begin{bmatrix} z \\ x \\ u \end{bmatrix} : z_2 = x_1 + u,\ z_3 = x_2,\ x_3 = 0 \right\}.$$

Then $\rho(S) = \mathbb{C}$ (and we choose $\Omega = \mathbb{C}_+$) and

$$(\lambda \mathbf{E} - \mathbf{A})^{-1} = \begin{bmatrix} -1 & -\lambda & -\lambda^2 \\ 0 & -1 & \lambda \\ 0 & 0 & -1 \end{bmatrix}, \quad \widehat{\mathfrak{A}}(\lambda) = \begin{bmatrix} 0 & 1 & -\lambda \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \widehat{\mathfrak{B}}(\lambda) = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}.$$

We have

$$\mathscr{X}_0 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\}, \qquad \mathscr{Z}_0 = \left\{ \begin{bmatrix} z_1 \\ 0 \\ 0 \end{bmatrix} : z_1 \in \mathbb{R} \right\}.$$

Since $\mathscr{X}_0 + \mathscr{Z}_0 \neq \mathscr{X}$, this system is not impulse controllable. Since this system is in feedback equivalence form, the system space from [16] can be calculated using [16, (3.2)] and equals

$$\mathscr{V}_{\mathrm{sys}} = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ u \end{bmatrix} : x_2 = x_3 = 0, x_1 = -u \right\}.$$

However, it is instructive to calculate this inductively as in Section 9.4. The part of S in its classical state space equals (compared to $\mathrm{gph}(S)$ *we have* $z_3 = 0$ *since z must belong to the classical state space)*

$$\mathrm{gph}(S_2) = \left\{ \begin{bmatrix} z \\ x \\ u \end{bmatrix} : z_2 = x_1 + u,\ 0 = z_3 = x_2,\ x_3 = 0 \right\}.$$

The classical state space of S_2 *is*

$$\mathscr{X}_2 = \left\{ \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} : x_1 \in \mathbb{R} \right\},$$

and the part of S_2 *in* $\mathscr{X}_2$ *equals (compared to* $\mathrm{gph}(S_2)$ *we have* $z_2 = 0$ *since* $z \in \mathscr{X}_2$*)*

$$\mathrm{gph}(S_3) = \left\{ \begin{bmatrix} z \\ x \\ u \end{bmatrix} : 0 = z_2 = x_1 + u,\ 0 = z_3 = x_2,\ x_3 = 0 \right\}.$$

The classical state space $\mathscr{X}_3$ *of* S_3 *equals* $\mathscr{X}_2$ *and therefore the induction stops. We have*

$$\mathrm{dom}(S_3) = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} : 0 = x_1 + u,\ x_2 = x_3 = 0 \right\},$$

which is indeed $\mathscr{V}_{\mathrm{sys}}$ *as determined above.*

Example 9.24. *We continue Example 9.23. We now consider* $\mathscr{Y} = \mathbb{R}^3$ *and the output* $y = x$*. With this output, the cost being finite implies stability of the state, so there is no difference between the problem "with state stability" and without this stability requirement. Let*

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Then

$$\widehat{\mathfrak{C}}(\lambda) = \widehat{\mathfrak{A}}(\lambda), \qquad \widehat{\mathfrak{D}}(\lambda) = \widehat{\mathfrak{B}}(\lambda).$$

Frequency domain trajectories satisfy

$$\hat{y}_1(\lambda) = x_2^0 - \lambda x_3^0 - \hat{u}(\lambda), \qquad \hat{y}_2(\lambda) = -x_3^0, \qquad \hat{y}_3(\lambda) = 0. \tag{11}$$

The condition that $\hat{y}_1$ *and* $\hat{u}$ *are Laplace transforms of* $L^2(0,\infty)$ *functions implies that* $x_2^0 - \lambda x_3^0$ *must be as well; this implies* $x_2^0 = x_3^0 = 0$*. Therefore the space of finite cost states is*

$$\mathscr{X}_{\mathrm{finite}} = \left\{ \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} : x_1 \in \mathbb{R} \right\}.$$

Since $\mathrm{im}(\widehat{\mathfrak{B}}(\lambda)) \subset \mathscr{X}_{\text{finite}}$*, the input finite future cost condition is satisfied (but the finite future cost condition is not). The part of S in* $\mathscr{X}_{\text{finite}}$ *is*

$$\mathrm{gph}(S_{\text{finite}}) = \left\{ \begin{bmatrix} z \\ x \\ u \end{bmatrix} : 0 = z_2 = x_1 + u,\ 0 = z_3 = x_2,\ x_3 = 0 \right\}.$$

We see that S_{finite} *equals* S_3 *from Example 9.23 and that* $\mathrm{dom}(S_{\text{finite}}) = \mathscr{V}_{\text{sys}}$ *(as expected since* $y = x$*).*

For an initial condition in $\mathscr{X}_{\text{finite}}$ *we have from* (11) *that* $y_1 = -u$*,* $y_2 = y_3 = 0$ *no matter what the initial condition in* $\mathscr{X}_{\text{finite}}$ *is. Therefore clearly* $u = 0$ *is the optimal control and the optimal cost equals zero. We indeed see that* $q = 0$ *is a solution of the Riccati equation* (10) *with the standard output extension* $w = \sqrt{2}u$*. From Theorem 9.18 we then indeed conclude (since zero must be the smallest solution) that the optimal cost is zero and the optimal control satisfies* $\sqrt{2}u = 0$ *(i.e. is zero).*

More generally than just verifying that the obvious candidate solution solves the Riccati equation, we can write down the Riccati equation using an image representation of S_{finite}*. Since many components are known to be zero, we only need consider* z_1*,* x_1*,* y_1 *and u. We then obtain the image representation (the idea being that* z_1 *and u are "free" and uniquely determine the other variables)*

$$F = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad K = L_{\text{out}} = \begin{bmatrix} 0 & -1 \end{bmatrix}, \qquad L_{\text{in}} = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

This gives the Riccati equation (as in (8)*)*

$$\begin{bmatrix} 0 & -X \\ -X & 2 \end{bmatrix} = \begin{bmatrix} L_1^2 & L_1 L_2 \\ L_1 L_2 & L_2^2 \end{bmatrix},$$

which gives $L_1 = 0$*,* $X = 0$ *and* $L_2 = \sqrt{2}$*, which is consistent with what we obtained above.*

Using that

$$\mathscr{V}_{\text{sys}} = \mathrm{im} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

the Riccati equation in [17] is

$$\begin{bmatrix} -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A}^* P \mathbf{E} + \mathbf{E}^* P \mathbf{A} + \mathbf{C}^* \mathbf{C} & \mathbf{E}^* P \mathbf{B} + \mathbf{C}^* \mathbf{D} \\ \mathbf{B}^* P \mathbf{E} + \mathbf{D}^* \mathbf{C} & I + \mathbf{D}^* \mathbf{D} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{K}^* \mathbf{K} & \mathbf{K}^* \mathbf{L} \\ \mathbf{L}^* \mathbf{K} & \mathbf{L}^* \mathbf{L} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

which gives

$$\begin{bmatrix} -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{K}^* \mathbf{K} & \mathbf{K}^* \mathbf{L} \\ \mathbf{L}^* \mathbf{K} & \mathbf{L}^* \mathbf{L} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

which has as solution $P \in \mathbb{R}^{3\times 3}$ an arbitrary symmetric matrix and

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{L} = \sqrt{2}.$$

The stability condition from [17] is also satisfied for this solution since for all $\lambda \in \mathbb{C}$

$$\operatorname{rank} \begin{bmatrix} -\lambda \mathbf{E} + \mathbf{A} & \mathbf{B} \\ \mathbf{K} & \mathbf{L} \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 1 & -\lambda & 0 & 1 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix} = 4.$$

According to [17, Theorem 5.7(a)] the optimal cost equals $\langle P\mathbf{E}x^0, \mathbf{E}x^0 \rangle$ for all $x^0 \in \mathscr{V}_{\text{diff}}$. In this example $\mathscr{V}_{\text{diff}} = \mathscr{X}_{\text{finite}}$ and we therefore see that the optimal cost equals zero (independent of the initial condition $x^0 \in \mathscr{V}_{\text{diff}}$). According to [17, Theorem 5.7(a)] the optimal control and optimal state satisfy $\mathbf{K}x + \mathbf{L}u = 0$, which gives $u = 0$. We conclude that (as should be the case) we have consistency between our results and those of [17].

Example 9.25. *We continue Example 9.23 with a different output than in Example 9.24. Consider now $\mathscr{Y} = \mathbb{R}$ and $y = x_3$. Then*

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{D} = 0,$$

and

$$\widehat{\mathfrak{C}}(\lambda) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \qquad \widehat{\mathfrak{D}}(\lambda) = 0.$$

We see that for all initial conditions and all inputs the output is zero. Therefore the space of finite cost states equals the whole state space: $\mathscr{X}_{\text{finite}} = \mathscr{X}$. It is also clear that the optimal cost is zero and the optimal control is zero (independent of the initial condition). We indeed see that $q = 0$ is a solution of the Riccati equation (10) *with the standard output extension $w = u$. From Theorem 9.18 we then indeed conclude (since zero must be the smallest solution) that the optimal cost is zero and the optimal control satisfies $u = 0$ (i.e. is zero). Note that the system space $\mathscr{V}_{\text{sys}}$ plays no role in this example and [17] is not applicable.*

10 Conclusion

We considered linear quadratic optimal control for infinite dimensional differential-algebraic equations (more specifically, for future-resolvable input/state/output nodes) and obtained an algebraic Riccati equation for the quadratic form which gives the optimal cost and which characterizes the optimal control.

For simplicity of exposition, we only considered the most standard cost function for the linear quadratic optimal control problem and the associated algebraic Riccati equation, however the method is applicable to general quadratic cost functions.

Also for simplicity of exposition, we considered input/state/output nodes rather than state/signal nodes (i.e. we assumed that the signal component is a priori split into an input and an output). However, the cost in linear quadratic optimal control is an input/output invariant notion in the sense of [2, Section 5.6.1] and therefore does not depend on the decomposition of the signal component into an input and an output. The objective of writing the input as a state feedback is however not an input/output invariant notion and as illustrated in Example 8.5, it can be beneficial to take a state/signal perspective. With the relevant input/state/output results from [2] replaced by the corresponding state/signal results from [2], state/signal equivalents of our results can be obtained.

References

[1] Ala' Alalabi and Kirsten Morris. Finite-time linear-quadratic optimal control of partial differential-algebraic equations, 2024. arXiv:2404.03766.

[2] Damir Z. Arov and Olof J. Staffans. *Linear state/signal systems*, volume 183 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2022.

[3] Douglas J. Bender and Alan J. Laub. The linear-quadratic optimal regulator for descriptor systems. *IEEE Trans. Automat. Control*, 32(8):672–688, 1987.

[4] Thomas Berger and Timo Reis. Controllability of linear differential-algebraic systems—a survey. In *Surveys in differential-algebraic equations. I*, Differ.-Algebr. Equ. Forum, pages 1–61. Springer, Heidelberg, 2013.

[5] D. Cobb. Descriptor variable systems and optimal state regulation. *IEEE Trans. Automat. Control*, 28(5):601–611, 1983.

[6] Mehmet Erbay, Birgit Jacob, Kirsten Morris, Timo Reis, and Caren Tischendorf. Index concepts for linear differential-algebraic equations in finite and infinite dimensions, 2024. arXiv:2401.01771.

[7] Ton Geerts. Linear-quadratic control with and without stability subject to general implicit continuous-time systems: coordinate-free interpretations of the optimal costs in terms of dissipation inequality and linear matrix inequality; existence and uniqueness of optimal controls and state trajectories. *Linear Algebra Appl.*, 203/204:607–658, 1994.

[8] Hannes Gernandt and Timo Reis. A pseudo-resolvent approach to abstract differential-algebraic equations, 2023. arXiv:2312.02303.

[9] Hannes Gernandt and Timo Reis. Linear-quadratic optimal control for abstract differential-algebraic equations, 2024. arXiv:2402.08762.

[10] Birgit Jacob and Kirsten Morris. On solvability of dissipative partial differential-algebraic equations. *IEEE Control Systems Letters*, 6:3188–3193, 2022.

[11] T. Katayama and K. Minamino. Linear quadratic regulator and spectral factorization for continous-time descriptor systems. In *Proceedings of the 31st IEEE Conference on Decision and Control*, pages 967–972 vol.1, 1992.

[12] Volker Mehrmann and Hans Zwart. Abstract dissipative Hamiltonian differential-algebraic equations are everywhere, 2024. arXiv:2311.03091.

[13] Mark R. Opmeer. Infinite-dimensional linear systems: a distributional approach. *Proc. London Math. Soc. (3)*, 91(3):738–760, 2005.

[14] Mark R. Opmeer and Olof J. Staffans. Optimal state feedback input-output stabilization of infinite-dimensional discrete time-invariant linear systems. *Complex Anal. Oper. Theory*, 2(3):479–510, 2008.

[15] Mark R. Opmeer and Olof J. Staffans. Optimal control on the doubly infinite continuous time axis and coprime factorizations. *SIAM J. Control Optim.*, 52(3):1958–2007, 2014.

[16] Timo Reis, Olaf Rendel, and Matthias Voigt. The Kalman-Yakubovich-Popov inequality for differential-algebraic systems. *Linear Algebra Appl.*, 485:153–193, 2015.

[17] Timo Reis and Matthias Voigt. Linear-quadratic optimal control of differential-algebraic systems: the infinite time horizon problem with zero terminal state. *SIAM J. Control Optim.*, 57(3):1567–1596, 2019.