OPTIMAL CONTROL ON THE DOUBLY INFINITE CONTINUOUS-TIME AXIS AND COPRIME FACTORIZATIONS

MARK R. OPMEER* AND OLOF J. STAFFANS†

Abstract. We study the problem of existence of weak right or left or strong coprime factorizations in H-infinity over the right half-plane of an analytic function defined in some subset of the right half-plane. We give necessary and sufficient conditions for the existence of such coprime factorizations in terms of an optimal control problem over the doubly infinite continuous-time axis. In particular, we show that an equivalent condition for the existence of a strong coprime factorization is that both the control and the filter algebraic Riccati equation (of an arbitrary realization that need not be well-posed) have a solution (in general unbounded and not even densely defined) and that a coupling condition involving these two solutions is satisfied. The proofs that we give are partly based on corresponding discrete time results which we have recently obtained.

Key words. Riccati equation, linear quadratic optimal control, infinite-dimensional system, coprime factorization, input-output stabilization, state feedback

AMS subject classifications. 49N10; 47N70, 47A48, 47A56, 47A62, 93B28, 93C05, 93C25, 93D15, 93D25

1. Introduction. Linear finite-dimensional time-invariant systems in continuous time are typically modeled by the equations

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \\ x(0) &= x_0, \end{aligned} \tag{1.1}$$

on a triple of finite-dimensional vector spaces, namely, the *input* space $\mathcal{U}$, the *state* space $\mathcal{X}$, and the *output* space $\mathcal{Y}$. We have $u(t) \in \mathcal{U}$, $x(t) \in \mathcal{X}$ and $y(t) \in \mathcal{Y}$. In this article we are interested in the infinite-dimensional generalization of this situation (which models, e.g., evolutionary partial differential equations). The main subject of this article is the interplay between linear quadratic optimal control theory, the factorization approach to control theory, and Riccati equations.

In the remainder of this introduction we describe these connections (highlighting our new contributions) without getting into too much technical detail (which we leave for the main body of the article).

In the standard infinite-dimensional setting the main operator A in (1.1) is unbounded and it generates a C_0 semigroup in $\mathcal{X}$, whereas the control operator B and the observation operator C may be bounded or unbounded, and the feedthrough operator D is sometimes difficult to define. One common approach to discuss this case is to re-write (1.1) into the form

$$\begin{aligned} \begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} &= S \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \\ x(0) &= x_0. \end{aligned} \tag{1.2}$$

where S is an (in general unbounded) operator $\mathrm{dom}(S) \subset \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] \to \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$. In the case where B and D in (1.1) are bounded it is usually possible to split S into the block

* Department of Mathematical Sciences, University of Bath, BA2 7AY, Bath, United Kingdom. m.opmeer@maths.bath.ac.uk.

† Department of Mathematics, Åbo Akademi University, FIN-20500 Åbo, Finland, http://www.abo.fi/~staffans

matrix form $S = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$, where $\operatorname{dom}(A) = \operatorname{dom}(C) \subset \mathcal{X}$, $\operatorname{dom}(B) = \operatorname{dom}(D) = \mathcal{U}$, and $\operatorname{dom}(S) = \operatorname{dom}(A) \times \mathcal{U}$. However, in the most interesting cases (e.g., in boundary control systems) one is forced to use an unbounded operator B, and neither $\operatorname{dom}(S)$ nor S itself is of the above form. In this case we denote $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, where $A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ is the first component of $S \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ and $C\&D \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ is the second component of $S \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$. The above idea is formalized in the notion of an *operator node* (see, e.g., [2,17,28] or [33, Section 4.7]) which we review in Section 2.

One of the most classical problems in control theory, the Linear Quadratic Regulator (LQR) problem, is to minimize the quadratic cost function

$$J_{\text{fut}}(x_0, u) = \int_0^\infty \left(\|u(t)\|_{\mathcal{U}}^2 + \|y(t)\|_{\mathcal{Y}}^2\right) dt, \tag{1.3}$$

subject to the linear dynamics (1.1). This problem goes back to Kalman [11] (at least in this formulation, its origins can be tracked back further) and has now been a textbook subject in control theory for decades. The objective is not just to find an optimal control u^{opt}, but also to prove that it is of feedback form, i.e., that by adding the equation

$$u(t) = Kx(t),$$

to the equations (1.1), the unique solution of this new system of equations is the solution of (1.1) that minimizes the cost (1.3). Moreover, a crucial aspect of the solution of this problem is that the optimal feedback operator K can be obtained from the control Riccati equation as follows: We have

$$K = -W^{-1}(B^*Q + D^*C), \qquad W = 1_{\mathcal{U}} + D^*D, \tag{1.4}$$

where Q is the minimal nonnegative solution of the control Riccati equation

$$QA + A^*Q + C^*C = (B^*Q + D^*C)^*W^{-1}(B^*Q + D^*C), \tag{1.5}$$

which can also be written in the form

$$QA + A^*Q + C^*C = K^*WK. \tag{1.6}$$

We note that (1.5) depends on the individual operators A, B, C and D and not only on the "unsplit" operators $A\&B$ and $C\&D$. This precludes a straightforward generalization of that form of the Riccati equation to operator nodes, except in the special case where B and D are bounded. In particular, the interpretation of the term B^*Q in (1.4) and (1.5) causes significant problems. The Riccati equation has proven to be a notoriously difficult issue for unbounded control operators B; see, e.g., [10, 15, 16, 29, 30, 32, 39, 40] and Remark 5.11.

In order to also include the case of an unbounded control operator B we replace the above "standard" Riccati equation by the Lur'e form of the Riccati equation (which is in common use for singular optimal control problems). Instead of defining K by (1.4) we define K and F by (the sign change in K compared to (1.4) is not significant, but it leads to a slight simplification of the formulas)

$$K = W^{-1/2}(B^*Q + D^*C), \qquad F = W^{1/2}, \qquad W = 1_{\mathcal{U}} + D^*D. \tag{1.7}$$

Then (1.6) is replaced by

$$QA + A^*Q + C^*C = K^*K, \tag{1.8}$$

and equations (1.7) and (1.8) can be rewritten in block matrix form as

$$\begin{bmatrix} QA + A^*Q + C^*C & QB + C^*D \\ B^*Q + D^*C & 1_{\mathcal{U}} + D^*D \end{bmatrix} = \begin{bmatrix} K^*K & K^*F \\ F^*K & F^*F \end{bmatrix}.$$

This in turn can be re-written as (applying the above equation to $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$, taking the inner-product with $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ and re-arranging):

$$\left\langle \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}, Qx \right\rangle + \left\langle Qx, \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \right\rangle + \left\| \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \right\|^2 + \|u\|^2 = \left\| \begin{bmatrix} K & F \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \right\|^2.$$

As we show in Section 5, replacing $\begin{bmatrix} A & B \end{bmatrix}$ by $A\&B$, $\begin{bmatrix} C & D \end{bmatrix}$ by $C\&D$, $\begin{bmatrix} K & F \end{bmatrix}$ by $K\&F$ and requiring this equation to hold only for $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \mathrm{dom}(S)$ (or for $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ in a subspace of $\mathrm{dom}(S)$ if Q is unbounded), this is indeed the correct Riccati equation in the sense that it provides the solution to the optimal control problem. We note that the non-singularity condition that F is invertible in general has to be replaced by another condition (since in general only the combination $K\&F$ exists); this is discussed further in Remark 5.11. We believe that the use of Lure's form of the Riccati equation is crucial in avoiding some of the pitfalls in the existing Riccati equation theory in the case of an unbounded control operator.

The classical companion to the LQR problem is the optimal filtering problem [12,13]. In its original formulation the filtering problem (whose solution is the Kalman filter) amounts to finding the best estimate of the state at some given time t_0 based on measurements of the past values of a signal. If we take $t_0 = 0$, then this leads to a minimization problem over $\mathbb{R}^-$, namely to the problem of minimizing the *past cost function*

$$J_{\mathrm{past}}(x_0, u) = \int_{-\infty}^{0} \left(\|u(t)\|_{\mathcal{U}}^2 + \|y(t)\|_{\mathcal{Y}}^2 \right) dt, \tag{1.9}$$

where this time x_0 is the final value in (1.1). The optimal cost is a quadratic function of the state x_0, which can be written in the form $\langle x_0, Rx_0 \rangle_{\mathcal{X}}$. Here R^{-1} turns out to be the minimal nonnegative solution of the control Riccati equation for the adjoint system, i.e., the system that one gets by replacing $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ by $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^* = \left[\begin{smallmatrix} A^* & C^* \\ B^* & D^* \end{smallmatrix}\right]$. The control Riccati equation for the adjoint system is commonly called the *filter Riccati equation* of the original problem.[1] Observe that in the case of the filter Riccati equation it is the possible unboundedness of the observation operator C that becomes problematic if one wants to rewrite it in "standard form", as the control operator of the adjoint system is C^*.

Since most infinite-dimensional systems cannot be solved backwards in time, the minimization problem on $\mathbb{R}^-$ is a genuinely different problem than the minimization problem on $\mathbb{R}^+$. In particular, it is typical that not every final value x_0 can be reached. This implies that the operator R above will be unbounded in such cases. Another issue is that existence and uniqueness of solutions on $\mathbb{R}^-$ is problematic (see e.g. [36, Section 5]). We overcome this latter problem by considering only linear combinations of exponential trajectories of the form

$$\begin{bmatrix} x(t) \\ u(t) \\ y(t) \end{bmatrix} = \mathrm{e}^{\lambda t} \begin{bmatrix} x_0 \\ u_0 \\ y_0 \end{bmatrix}, \qquad t \in \mathbb{R}^-.$$

[1] The name "filter Riccati equation" is somewhat misleading in the sense that it is the *inverse* of the solution of the filter Riccati equation that solves the original filtering problem, and not the solution itself!

In the finite-dimensional case $x_0 = (\lambda - A)^{-1}Bu_0$ and $y_0 = C(\lambda - A)^{-1}Bu_0 + Du_0$ and in the general case appropriate generalizations of these formulas turn out to hold. It is clear from these formulas that we should restrict λ to lie in the resolvent set of A. It is also clear that to have a meaningful infimization of (1.9), we should consider "enough" trajectories, i.e. we should consider all $\lambda \in \Omega$ for some large enough set Ω. In the standard setting where the main operator A in (1.1) generates a C_0 semigroup we can, e.g., take Ω to be the right half-plane $\mathbb{C}_{\omega(A)} := \{\lambda \in \mathbb{C} \mid \Re\lambda > \omega(A)\}$ where $\omega(A)$ is the growth bound of the semigroup generated by A, but other choices of Ω are also possible (especially in those cases where A does not generate a C_0 semigroup).

To obtain the filter Riccati equation we employ duality, i.e., we relate the optimization problem on $\mathbb{R}^-$ for the system $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ to the optimization problem on $\mathbb{R}^+$ for the system $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$. Because of this it makes sense to minimize (1.3) over the set of trajectories on $\mathbb{R}^+$ which is dual to the set of exponential trajectories considered on $\mathbb{R}^-$. A pair of functions $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ belongs to this set if and only if both u and y belong to $L^2(\mathbb{R}^+)$, and if the Laplace transforms of u and y satisfy

$$\hat{y}(\lambda) = C(\lambda - A)^{-1}x_0 + \big(D + C(\lambda - A)^{-1}B\big)\,\hat{u}(\lambda), \qquad \lambda \in \Omega, \tag{1.10}$$

where Ω is the same set as before. This equation coincides with the equation that one gets by taking formal Laplace transforms in (1.1).

Before commenting on the optimal control problem on $\mathbb{R}$, we indicate how the above considered optimal control problems on $\mathbb{R}^+$ and $\mathbb{R}^-$ relate to the factorization approach to control theory.

The factorization approach [35] leads to a parameterization of all stabilizing controllers and also has links to metrics that measure robustness of controllers. There are well-known connections between the optimal control problem and the factorization approach. See, e.g., [14, 18, 21] for the finite-dimensional case, the bibliography of [5, Chapter 9] for the infinite-dimensional case, and [4, 19, 31] for more recent contributions in the infinite-dimensional case. The characteristic function (or transfer function) of the system (1.1) is defined and analytic on the resolvent set $\rho(A)$ of the main operator A, and it is given by $\mathsf{G}(\lambda) := D + C(\lambda - A)^{-1}B$, $\lambda \in \rho(A)$. It is possible to introduce the notion of the characteristic function G of (1.2) in an analogous way, and it is still defined and analytic on $\rho(A)$, where A is the main operator of (1.2) (see Section 2 for details). For simplicity, let us assume that this transfer function G is *well-posed*, i.e., that $\rho(A)$ contains some open right-half plane Ω, and G is uniformly bounded on this half-plane Ω. The basic idea in the factorization approach is to write the restriction of G to Ω as a quotient $\mathsf{G}(\lambda) = \mathsf{N}(\lambda)\mathsf{M}(\lambda)^{-1}$, $\lambda \in \Omega$, of two analytic functions N and M which are defined and uniformly bounded on the full open right half-plane $\mathbb{C}_+ := \{\lambda \in \mathbb{C} \mid \Re\lambda > 0\}$, and in addition, $\mathsf{M}(\lambda)$ is invertible for all $\lambda \in \Omega$ and M^{-1} is uniformly bounded in Ω. In the finite-dimensional case, G is a rational function and such a factorization always exists. This is not true in the infinite-dimensional case (e.g. when G has an essential singularity in the right half-plane). Given a function G, a system (1.1) with G as its transfer function is called a *realization*. We note that realizations are never unique.

To relate the optimal control problem on $\mathbb{R}^+$ to right factorizations, we recall the *state finite future cost condition*, i.e., the condition that for every $x_0 \in \mathcal{X}$ there exists a control u such that $J_{\text{fut}}(x_0, u) < \infty$. It is known that that if G is well-posed, then G has a right factorization of the type described above if and only if G has a realization of the type (1.2) which is well-posed in the sense of [33], and which satisfies the state finite future cost condition (see [20]). Even though this is a necessary and

sufficient condition, it is still problematic in two respects: first of all, the condition that (1.2) must be well-posed is not a natural assumption for this problem (i.e., in order to define G there is no need to assume that (1.2) is well-posed, only that it is given by an operator node), and secondly, if it turns out that a given realization (1.2) (well-posed or not) of G does not satisfy the state finite future cost condition, then the above result says nothing. In Theorem 5.9 below we improve the above mentioned result from [20] by giving a necessary and sufficient condition for the existence of a right factorization over H^∞ in terms of an *arbitrary operator node* realization of G (which need not be well-posed, and which need not satisfy the state finite future cost condition). In this theorem the state finite future cost condition has been replaced by the weaker *input finite future cost condition* (see Definition 5.7). It follows from our results that for finite-dimensional systems the input finite future cost condition always holds (the state finite future cost condition may not hold) although this is not obvious from the definition. Another necessary and sufficient condition in terms of an arbitrary operator node realization is that the control Riccati equation of this realization has a solution (Theorem 5.9); however this solution may be unbounded. This is the main reason why we have to allow for unbounded solutions of our Riccati equations. In this case it is technically more convenient to work with the sesquilinear form q associated to the operator Q through

$$q[x, y] = \langle Qx, y\rangle, \quad x, y \in \operatorname{dom}(Q),$$

rather than with Q itself, and we therefore formulate our Riccati equations in terms of sesquilinear forms.

The optimal control problem on $\mathbb{R}^-$ is related to left factorizations in a similar manner (Theorem 6.5), and it can be reduced to the the problem on $\mathbb{R}^+$ by means of duality. The dual of the input finite future cost condition, which we call the *output coercive past cost condition* (Definition 6.2), says roughly that the cost of reaching an observable state should have a nonzero lower bound.

A right factorization is called strongly coprime if there exist uniformly bounded analytic functions $\tilde{\mathsf{X}}$ and $\tilde{\mathsf{Y}}$ on the open right half-plane such that $\tilde{\mathsf{X}}\mathsf{M} - \tilde{\mathsf{Y}}\mathsf{N} = 1_{\mathcal{U}}$ on the open right half-plane. Not every transfer function which has a right factorization has a strongly coprime right factorization. It is known that a transfer function has a strongly coprime right factorization if and only if it has a realization for which the state finite future cost condition is satisfied both for the system itself and for the dual system (see [4, 20]). In our context this is also equivalent to the existence of a realization for which both the control and the filter Riccati equation have a bounded solution. For a general realization it is not sufficient to have (unbounded) solutions to both the control and the filter Riccati equation. It turns out that there is an additional coupling condition: we show in Theorem 7.5 that for a general realization the necessary and sufficient condition is that both the control and the filter Riccati equation have a (possibly unbounded) solution Q respectively P, and that there exists a finite constant M such that $Q \leq MP^{-1}$ (in the sense of nonnegative self-adjoint operators, or more generally, in the sense of nonnegative self-adjoint relations). We remark that this is reminiscent of the famous "coupling condition" in H^∞ control [9]. Also the solution of the H^∞ control problem involves the solutions of two algebraic Riccati equations. One of them is of "control" type and the other of "filter" type, and they tend to the LQG/H^2 control and filter Riccati equations that we consider in this article when a certain parameter γ tends to infinity. Let us call these solutions X and Y, respectively. The standard coupling condition in H^∞ control is that $r(XY) < \gamma^2$,

where r denotes the spectral radius. If Y is invertible, then this is equivalent to the condition $X < \gamma^2 Y^{-1}$ in the sense of nonnegative self-adjoint operators (if Y is not invertible, then it is still true in the sense of nonnegative self-adjoint relations). Thus, in this article we prove that in the LQG/H^2 setting the H^∞ coupling condition $X < \gamma^2 Y^{-1}$ must be replaced by the analogous coupling condition $Q \leq MP^{-1}$, where M must be finite, but without any size limitation on M. The reason why this LQG/H^2 coupling condition has not been discovered before is that in the existing literature on LQG/H^2 control it has been hidden in the general setting: if both P and Q are required to be bounded, then automatically $r(PQ) < \infty$, and we may take $M = r(PQ)$.

The optimal control problem related to strongly coprime factorizations is the one given by the *two-sided cost function* obtained by adding the past and future costs, i.e.,

$$J(x_0, u) = \int_{-\infty}^{\infty} \left(\|u(t)\|_{\mathcal{U}}^2 + \|y(t)\|_{\mathcal{Y}}^2 \right) dt, \tag{1.11}$$

where this time x_0 is the intermediate value in (1.1). In this setting the LQG/H^2 coupling condition that we mentioned above is equivalent to the *past cost dominance condition* (Definition 7.2), which says that the future cost should be dominated by the past cost, or equivalently, that the full cost should be dominated by the past cost.

One characteristic feature of this article is that we allow solutions of the control and filter Riccati equations to be unbounded (and not even densely defined). For the results that we present here this amount of generality is unavoidable, since our main theorems would be false in a less general setting. Unbounded solutions of Riccati equations have been considered before, but not in the present setting. In the discrete time setting unbounded solutions appear in, e.g., [1,23,24,25]. Technically the discrete time setting is much easier than the continuous time setting due to the boundedness of the discrete time versions of the operators A, B, C, and D. Unbounded solutions in the continuous time setting with bounded B and D are discussed in [6, 7, 8], and also this special setting is technically much easier than the setting in this article. The most closely related result that we have been able to find in the literature is the study of the continuous time Kalman–Yakubovich–Popov inequality in [3] (which has been the main source of inspiration for this article in addition to [23, 24, 25]). There the setting is essentially the same as our present setting, apart from the fact that there A is throughout required to generate a C_0 semigroup, and our positive cost function is replaced by an indefinite cost function of KYP-type. The connection with the factorization approach is absent from the above references, and the connections with optimal control described there were far more limited.

The discrete time counterpart of the present theory has recently been developed in a series of three papers [23,24,25]. This article is the first in a series of three where we discuss the continuous time case, and it is partly based on the above discrete time results. To round off this introduction, we comment on why we in this first article have chosen to work with systems of the type (1.2) where S is an operator node.

In the discrete time setting it is natural to assume that the operators A, B, C, and D appearing in the discrete time equation

$$\begin{aligned} x_{n+1} &= Ax_n + Bu_n, \\ y_n &= Cx_n + Du_n, \end{aligned} \tag{1.12}$$

are bounded, but in continuous time the choice of the setting in which the problem

is studied is less obvious. Basically we can think the following four alternatives (in increasing order of generality):

A) Operator nodes of the type (1.2) which are *well-posed* in the sense of [33],
B) Systems nodes in the sense of [33], i.e., operator nodes with the extra property that the main operator A generates a C_0 semigroup (see Section 2 for details),
C) Operator nodes for which the main operator A satisfies $\rho(A)\cap\mathbb{C}^+ \neq \emptyset$ (so that the transfer function G is defined at least in some part of the right half-plane $\mathbb{C}^+$),
D) Resolvent linear systems in the sense of [22].

As we mentioned above, out of these possibilities we have in this article chosen to work in the class C). In our opinion, this is the class of systems that the most resembles the standard finite-dimensional set-up (1.1), and thereby the one which is intuitively easiest to comprehend. It is also the class of systems which is easiest to use in applications in the following sense. To show that a given system is of type A) or type B) one must first show that it is of type C), and then one needs additional, often nontrivial, arguments to show that it actually of type A) or type B). By working directly in the class C) those additional arguments are no longer needed. Furthermore, as it turns out, the additional structure present in classes A) and B) does not lead to any significant simplifications of either the statements of our main results nor of their proofs. The setting D) is, of course, much more general than the settings A)–C), but it is very different from A)–C), and in the setting D) both the statements of the main results and their proofs become more complicated. We shall return to the settings A) and B) in our second article in this series, and to class D) in the third.

To model a continuous time system one often starts with a (formal) partial differential equation, a (formal) input operator, and a (formal) output operator, and then tries to show that the system belongs to one of the classes A)–D) above. To do so one must first choose a function space in which to study this partial differential equation, which will then become the state space of the realization. As we observed in the discrete time setting in [23, p. 481]: "Choosing the proper state space is usually considered to be something that has to be done before one can solve control problems. One of the main points of the present series of articles is that it should instead be considered as an *integral part* of the control problem." The same statement is even more true in the continuous time setting. As we shall show in our second continuous time article, the results that we obtain here can in some cases be used in the following way: *we start with a realization of class C), solve the future or past cost minimization problem, and then use the solution of this problem to get a realization of class A) by redefining the norm in the state space.* This is illustrated by two simple examples (a one-dimensional wave equation and a one-dimensional heat equation) in [24, pp. 5089–5090], as well as in the example given in Section 8 below. More precisely, the following additional claims are true (as will be proved in our next paper):

(i) Both the optimal state feedback system that one gets from the future cost minimization problem and the optimal output injection system that one gets from the past cost minimization problem are actually well-posed with respect to the appropriate norm in the state space (i.e., the norm induced by the optimal future and past costs respectively),
(ii) If the transfer function is well-posed (in the sense described earlier), then the open loop system is also well-posed with respect to the appropriate norms derived from the optimization problems.
(iii) The norm derived from the future cost minimization problem is in a certain

sense the weakest possible norm within the class of all operator node realizations that satisfy the state finite future cost condition, and the norm derived from the past cost minimization problem is in a certain sense the strongest possible norm within the class of all operator node realizations that satisfy the state coercive past cost condition.

Unfortunately, if the transfer function is not well-posed, then the open loop system with the appropriate norms derived from the optimization problems does not always seem to be an operator node. However, under some relatively weak assumptions it can be shown that it is a system of the class D). We shall return to this question in our third article on the continuous time cost minimization problem.

The present article consists of 9 sections and an appendix, this introduction being the first section. In Section 2 we collect some background material on operator nodes, the class of systems that we consider in this article. In Section 3 we solve the optimal control problems on $\mathbb{R}^+$, $\mathbb{R}^-$ and $\mathbb{R}$. The main result in this section is Theorem 3.18, which describes the connection between the past cost minimization problem for the original system and the future cost minimization problem for the adjoint system. Section 4 describes the connection between the present continuous time setting and the discrete-time setting used in [23, 24, 25]. In Section 5 we consider the control Riccati equation and right factorizations. That section contains the already mentioned Theorem 5.9, which is the first of our three main results. Section 6 contains results on the filter Riccati equation and left factorizations. The results presented there (including the second of our main results, Theorem 6.5) follow easily from those in Section 5 with the help of Theorem 3.18. Section 7 contains our results on doubly coprime factorizations, including our third main result, Theorem 7.5. This third main result contains the already mentioned new LQG/H^2 coupling condition. In Section 8 we apply our theory to investigate a partial differential equation example originally presented in [6, Section 2.2], and which turns out to have a transfer function which does have right and left factorizations, but no strongly coprime factorization. This example also shows how one may start from a system of type C) and arrive at a system of type A) by solving the forward and backward cost minimization problems. Finally, Section 9 contains an example which illustrates that the minimization and factorization results obtained here depend on the component Ω of $\rho(A) \cap \mathbb{C}^+$ that is chosen in the precise problem formulation.

2. Continuous Time Operator Nodes. In this article we will use a natural continuous time setting (that of operator nodes), earlier used in, e.g., [2, 17, 27, 28, 33] (in slightly different forms). In this section, for easy reference, we collect some results on operator nodes from the literature that we will need in this article.

In the sequel, we think about the block matrix $S = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ as *one single closed (possibly unbounded) linear operator* from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] = \mathcal{X} \oplus \mathcal{U}$ to $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ with dense domain $\operatorname{dom}(S) \subset \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$, and write (1.1) in the form

$$\Sigma\colon \begin{cases} \begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = S \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, & t \in (-\infty, \infty), x(0) \quad = x_0. \end{cases} \tag{2.1}$$

In the infinite-dimensional case such an operator S need not have a four block decomposition corresponding to the decompositions $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ of the domain and range spaces. However, we shall throughout assume that the operator

$$\begin{aligned} Ax &:= P_{\mathcal{X}} S \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right], \\ x \in \operatorname{dom}(A) &:= \{x \in \mathcal{X} \mid \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right] \in \operatorname{dom}(S)\}, \end{aligned} \tag{2.2}$$

is closed and densely defined in $\mathcal{X}$ (here $P_{\mathcal{X}}$ is the orthogonal projection onto $\mathcal{X}$). We define $\mathcal{X}^1 := \operatorname{dom}(A)$ with the graph norm of A, $\mathcal{X}_*^1 := \operatorname{dom}(A^*)$ with the graph norm of A^*, and let $\mathcal{X}^{-1}$ be the dual of $\mathcal{X}_*^1$ when we identify the dual of $\mathcal{X}$ with itself. Then $\mathcal{X}^1 \subset \mathcal{X} \subset \mathcal{X}^{-1}$ with continuous and dense embeddings, and the operator A has a unique extension to an operator $A|_{\mathcal{X}} = (A^*)^* \in \mathcal{B}(\mathcal{X}; \mathcal{X}^{-1})$ (with the same spectrum as A), where we interpret A^* as an operator in $\mathcal{B}(\mathcal{X}_*^1; \mathcal{X})$. Additional assumptions on S will be imposed in Definition 2.1 below.

The remaining blocks of S are only partially defined. The 'block' B will be an operator in $\mathcal{B}(\mathcal{U}; \mathcal{X}^{-1})$. In particular, it may happen that $\operatorname{img}(B) \cap \mathcal{X} = \{0\}$. The 'block' C will be an operator in $\mathcal{B}(\mathcal{X}^1; \mathcal{Y})$. We shall make no attempt to define the 'block' D in general since this can be done only under additional assumptions (see, e.g., [33, Chapter 5] or [34,37,38]). Nevertheless, we still use a modified block notation $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, where $A\&B = P_{\mathcal{X}} S$ and $C\&D = P_{\mathcal{Y}} S$.

DEFINITION 2.1. *By an* operator node *on a triple of Hilbert spaces $(\mathcal{X}, \mathcal{U}, \mathcal{Y})$ we mean a (possibly unbounded) linear operator $S\colon \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] \to \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ with the following properties. We decompose S into $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, where $A\&B = P_{\mathcal{X}} S\colon \operatorname{dom}(S) \to \mathcal{X}$ and $C\&D = P_{\mathcal{Y}} S\colon \operatorname{dom}(S) \to \mathcal{Y}$. We denote $\operatorname{dom}(A) = \left\{x \in \mathcal{X} \;\middle|\; \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right] \in \operatorname{dom}(S)\right\}$, define $A\colon \operatorname{dom}(A) \to \mathcal{X}$ by $Ax = A\&B \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right]$, and require the following conditions to hold:*

(i) S is closed as an operator from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{Y} \end{smallmatrix}\right]$ (with domain $\operatorname{dom}(S)$).

(ii) $A\&B$ is closed as an operator from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ to $\mathcal{X}$ (with domain $\operatorname{dom}(S)$).

(iii) A has a nonempty resolvent set, and $\operatorname{dom}(A)$ is dense in $\mathcal{X}$.

(iv) For every $u \in \mathcal{U}$ there exists a $x \in \mathcal{X}$ with $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \operatorname{dom}(S)$.

We call S a system node *if, in addition, A is the generator of a C_0 semigroup.*

LEMMA 2.2. *Every operator node S on $(\mathcal{X}, \mathcal{U}, \mathcal{Y})$ has the following additional properties:*

(v) The operator $A\&B$ (with $\operatorname{dom}(A\&B) = \operatorname{dom}(S)$) can be extended to an operator $\begin{bmatrix} A|_{\mathcal{X}} & B \end{bmatrix} \in \mathcal{B}(\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]; \mathcal{X}^{-1})$.

(vi) $\operatorname{dom}(S) = \left\{\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right] \;\middle|\; A|_{\mathcal{X}} x + Bu \in \mathcal{X}\right\}$.

(vii) For every $u \in \mathcal{U}$, the set $\left\{x \in \mathcal{X} \;\middle|\; \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \operatorname{dom}(S)\right\}$ is dense in $\mathcal{X}$. (Thus, in particular, $\operatorname{dom}(S)$ is dense in $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$.)

(viii) $C\&D \in \mathcal{B}(\operatorname{dom}(S); \mathcal{Y})$, where we use the graph norm

$$\left\|\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|_{\operatorname{dom}(A\&B)}^2 = \left\|A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|_{\mathcal{X}}^2 + \|x\|_{\mathcal{X}}^2 + \|u\|_{\mathcal{U}}^2 \tag{2.3}$$

of $A\&B$ on $\operatorname{dom}(S)$.

(ix) The graph norm of $A\&B$ on $\operatorname{dom}(S)$ defined above is equivalent to the full graph norm

$$\left\|\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|_{\operatorname{dom}(S)}^2 = \left\|A\&B \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|_{\mathcal{X}}^2 + \left\|C\&D \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\|_{\mathcal{X}}^2 + \|x\|_{\mathcal{X}}^2 + \|u\|_{\mathcal{U}}^2 \tag{2.4}$$

of S on $\operatorname{dom}(S)$.

(x) For every $\alpha \in \rho(A) = \rho(A|_{\mathcal{X}})$, the operator $\left[\begin{smallmatrix} 1_{\mathcal{X}} & -(\alpha - A|_{\mathcal{X}})^{-1} B \\ 0 & 1_{\mathcal{U}} \end{smallmatrix}\right]$ maps $\operatorname{dom}(S)$ one-to-one onto $\left[\begin{smallmatrix} \mathcal{X}_1 \\ \mathcal{U} \end{smallmatrix}\right]$ and it is bounded and invertible on $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$. The inverse of this operator (which maps $\left[\begin{smallmatrix} \mathcal{X}_1 \\ \mathcal{U} \end{smallmatrix}\right]$ one-to-one onto $\operatorname{dom}(S)$ and $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ one-to-one onto itself) is $\left[\begin{smallmatrix} 1_{\mathcal{X}} & (\alpha - A|_{\mathcal{X}})^{-1} B \\ 0 & 1_{\mathcal{U}} \end{smallmatrix}\right]$.

(xi) For each $\alpha \in \rho(A)$ the graph norm of S is equivalent to the norm

$$\left\|\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]\right\| := \left(\|x - (\alpha - A|_{\mathcal{X}})^{-1} Bu\|_{\mathcal{X}_1}^2 + \|u\|_{\mathcal{U}}^2\right)^{1/2}.$$

Proof. See [33, Lemmas 4.7.3 and 4.7.7]. □

Each operator node has a main operator, a control operator, an observation operator, and a transfer function:

DEFINITION 2.3. *Let $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ be an operator node on $(\mathcal{X},\mathcal{U},\mathcal{Y})$.*

(i) The operator A in Definition 2.1 is called the main operator *of S. If S is a system node, then we shall also refer to A as the* semigroup generator *of S.*

(ii) The operator B in Lemma 2.2 is called the control operator *of S.*

(iii) The operator $C\colon \mathcal{X}_1 \to \mathcal{Y}$ defined by $Cx = C\&D\left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right]$ is called the observation operator *of S.*

(iv) The transfer function *of S is the operator-valued function*

$$\widehat{\mathfrak{D}}(\alpha) = C\&D \begin{bmatrix} (\alpha - A|_{\mathcal{X}})^{-1}B \\ 1_{\mathcal{U}} \end{bmatrix}, \qquad \alpha \in \rho(A). \tag{2.5}$$

By the resolvent identity, for any two α, $\beta \in \rho(A)$,

$$\begin{aligned} \widehat{\mathfrak{D}}(\alpha) - \widehat{\mathfrak{D}}(\beta) &= C\&D \begin{bmatrix} (\beta-\alpha)(\alpha - A|_{\mathcal{X}})^{-1}(\beta - A|_{\mathcal{X}})^{-1}B \\ 0 \end{bmatrix} \\ &= (\beta-\alpha)C(\alpha - A)^{-1}(\beta - A|_{\mathcal{X}})^{-1}B. \end{aligned} \tag{2.6}$$

Note that if $B \in \mathcal{B}(\mathcal{U};\mathcal{X})$, then $\operatorname{dom}(S) = \left[\begin{smallmatrix} \mathcal{X}_1 \\ \mathcal{U} \end{smallmatrix}\right]$, and we can define the operator $D \in \mathcal{B}(\mathcal{U};\mathcal{Y})$ by $D = P_{\mathcal{Y}} S|_{\left[\begin{smallmatrix} 0 \\ \mathcal{U} \end{smallmatrix}\right]}$, after which formula (2.5) can be rewritten in the form

$$\widehat{\mathfrak{D}}(\lambda) = D + C(\lambda - A)^{-1}B, \qquad \lambda \in \rho(A). \tag{2.7}$$

Let

$$\begin{aligned} G_\alpha :&= \left(\begin{bmatrix} \alpha & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix} - \begin{bmatrix} A\&B \\ 0 & 0 \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} (\alpha - A)^{-1} & (\alpha - A|_{\mathcal{X}})^{-1}B \\ 0 & 1_{\mathcal{U}} \end{bmatrix}, \quad \alpha \in \rho(A). \end{aligned} \tag{2.8}$$

Then, for all $\alpha \in \rho(A)$, G_α is a bounded bijection from $\left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix}\right]$ onto $\operatorname{dom}(S)$, and

$$\begin{bmatrix} A\&B \\ C\&D \end{bmatrix} G_\alpha = \begin{bmatrix} A(\alpha - A)^{-1} & \alpha(\alpha - A|_{\mathcal{X}})^{-1}B \\ C(\alpha - A)^{-1} & \widehat{\mathfrak{D}}(\alpha) \end{bmatrix}, \qquad \alpha \in \rho(A). \tag{2.9}$$

As shown in [33, Lemma 4.7.6], one way to construct an operator node $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ is to specify a densely defined main operator A with nonempty resolvent set, a control operator $B \in \mathcal{B}(\mathcal{U};\mathcal{X}^{-1})$, and an observation operator $C \in \mathcal{B}(\mathcal{X}_1;\mathcal{Y})$, to fix some $\alpha \in \rho(A)$ and an operator $D_\alpha \in \mathcal{B}(\mathcal{U};\mathcal{Y})$, define $\operatorname{dom}(S)$ by condition (vi) in Lemma 2.2, let $A\&B$ be the restriction of $\begin{bmatrix} A|_{\mathcal{X}} & B \end{bmatrix}$ to $\operatorname{dom}(S)$, and define $C\&D\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ for all $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \operatorname{dom}(S)$ by

$$C\&D \begin{bmatrix} x \\ u \end{bmatrix} := C(x - (\alpha - A|_{\mathcal{X}})^{-1}Bu) + D_\alpha u. \tag{2.10}$$

The transfer function $\widehat{\mathfrak{D}}$ of this operator node satisfies $\widehat{\mathfrak{D}}(\alpha) = D_\alpha$.

LEMMA 2.4. *Let S be an operator node on $(\mathcal{X},\mathcal{U},\mathcal{Y})$ with main operator A, control operator B, observation operator C, and transfer function $\widehat{\mathfrak{D}}$. Then the adjoint S^* of S is an operator node on $(\mathcal{X},\mathcal{Y},\mathcal{U})$. The main operator of S^* is A^*, the control*

operator of S^* is C^*, the observation operator of S^* is B^*, and the transfer function of S^* is $\widehat{\mathfrak{D}}(\overline{\alpha})^*$, $\alpha \in \rho(A^*)$. If S is a system node, then so is S^*. Moreover,

$$\begin{aligned} \big((\alpha - A|_{\mathcal{X}})^{-1}B\big)^* &= B^*(\overline{\alpha} - A^*)^{-1}, \\ \big(C(\alpha - A)^{-1}\big)^* &= (\overline{\alpha} - A|_{\mathcal{X}}^*)^{-1}C^*, \qquad \alpha \in \rho(A). \end{aligned} \tag{2.11}$$

For a proof (and for more details), see, e.g., [2, Section 3], [17, Proposition 2.3], or [33, Lemma 6.2.14].

DEFINITION 2.5. *Let S be an operator node on $(\mathcal{X},\mathcal{U},\mathcal{Y})$. By a* classical trajectory *of the system $\Sigma = (S;\mathcal{X},\mathcal{U},\mathcal{Y})$ on some interval $I \subset \mathbb{R}$ we mean a triple of functions $\left[\begin{smallmatrix} x \\ u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} C^1(I;\mathcal{X}) \\ C(I;\mathcal{U}) \\ C(I;\mathcal{Y}) \end{smallmatrix}\right]$ satisfying*

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = S \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \qquad t \in I.$$

In Section 3 we shall extend this definition by introducing the notion of a *generalized stable trajectory* of Σ in the case where I is one of the intervals $I = \mathbb{R}^-$, $I = \mathbb{R}$, or $I = \mathbb{R}^+$. The state of such a trajectory is defined only at time $t = 0$, and the input and output components u and y belong to $L^2(I)$. This extended notion is not the standard one, but it is the natural one for the problem at hand.

3. The Future, Past, and Full Cost Minimization Problem. In principle, the future, past, and two-sided cost minimization problems are the following:

- In the *future cost minimization problem* we fix an initial state $x_0 \in \mathcal{X}$ and minimize the future cost (1.3) over a suitable set of generalized stable future trajectories $\left[\begin{smallmatrix} x \\ u \\ y \end{smallmatrix}\right]$ of Σ with the given initial state $x(0) = x_0$.
- In the *past cost minimization problem* we fix a final state $x_0 \in \mathcal{X}$ and minimize the past cost (1.9) over a suitable set of generalized stable past trajectories $\left[\begin{smallmatrix} x \\ u \\ y \end{smallmatrix}\right]$ of Σ with the given final state $x(0) = x_0$.
- In the *two-sided cost minimization problem* we fix an intermediate state $x_0 \in \mathcal{X}$ and minimize the two-sided cost (1.11) over all two-sided trajectories $\left[\begin{smallmatrix} x \\ u \\ y \end{smallmatrix}\right]$ of Σ with the given intermediate state $x(0) = x_0$, with the property that the restrictions of these trajectories to $\mathbb{R}^+$ and $\mathbb{R}^-$ are of the type considered above.

The definition of a classical trajectory of a system $\Sigma = (S;\mathcal{X},\mathcal{U},\mathcal{Y})$ is straightforward (see Definition 2.5). However, in the cost minimizations problems we shall use *generalized stable trajectories* of Σ instead of classical trajectories. Our definition of generalized stable trajectories of Σ uses a frequency domain approach. This approach works under minimal assumptions, and it makes it possible to treat a very general class of continuous time systems, namely the class of systems Σ induced by an operator node S introduced in Section 2.

In the definitions of the generalized stable trajectories (future, past, or two-sided) we throughout fix one particular open subset Ω of $\rho(A) \cap \mathbb{C}^+$, where A is the main operator of the operator node S.

Remark 3.1. **Throughout the rest of this article Σ will be induced by an operator node $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ with main operator A, and Ω will be a fixed nonempty open subset of $\rho(A) \cap \mathbb{C}^+$.** *(Thus in particular, we throughout assume that $\rho(A) \cap \mathbb{C}^+ \neq \emptyset$.)*

If A is the generator of a C_0 semigroup, then we may take Ω to be the right half-plane right half-plane $\mathbb{C}_{\omega(A)} := \{\lambda \in \mathbb{C} \mid \Re\lambda > \omega(A)\}$ where $\omega(A)$ is the growth bound of the semigroup generated by A, but other choices of Ω are also possible. The result is actually independent of the choice of Ω under the weak assumption that $\rho(A) \cap \mathbb{C}^+$ is connected, but if $\rho(A) \cap \mathbb{C}^+$ is disconnected then the result may depend on the choice of Ω, as we show with an example in Section 9.

First in Section 3.1 we consider the future cost minimization problem, then in Section 3.2 we consider the past cost minimization problem, and subsequently in Section 3.3 we consider the two-sided cost minimization problem. Finally, in Section 3.4, we consider the duality between the past and future cost minimization problems that plays a crucial role in the remainder of the article.

3.1. The Future Cost Minimization Problem. In the future cost minimization problem we must first define the set of trajectories over which we minimize the cost function $J_{\text{fut}}(x_0, u)$. If the given operator node is well-posed in the sense of [33], then we could define the notion of a generalized future trajectory for every given $u \in L^2(\mathbb{R}^+;\mathcal{U})$ in the standard way, and define this trajectory to have finite future cost if the output y is a function in $L^2(\mathbb{R}^+;\mathcal{Y})$, so that the cost is finite. However, for the purpose of this paper it is more convenient to use a different *frequency domain definition*, which works well also in the non-well-posed case.

To motivate the following definition we look at the Laplace transform of a trajectory satisfying (2.1) on the time interval $\mathbb{R}^+$. Formal Laplace transforms give

$$\hat{y}(\lambda) = \widehat{\mathfrak{D}}(\lambda)\hat{u}(\lambda) + C(\lambda - A)^{-1}x_0,$$

for all $\lambda \in \rho(A)$. However, in the definition below we require the above identity to hold only in the open subset Ω of $\rho(A) \cap \mathbb{C}^+$ that was fixed in Remark 3.1.

Definition 3.2.

(i) By the set of generalized stable future trajectories *of* Σ *we mean the set of all triples* $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{smallmatrix}\right]$ *which satisfy*

$$\hat{y}(\lambda) = \widehat{\mathfrak{D}}(\lambda)\hat{u}(\lambda) + C(\lambda - A)^{-1}x_0, \qquad \lambda \in \Omega, \tag{3.1}$$

where $\hat{u}$ *and* $\hat{y}$ *are the Laplace transforms of* u *and* y*, respectively. We denote this set by* $\mathfrak{W}_+$*, and call* x_0 *the* initial state*,* u *the* input component*, and* y *the* output component *of a triple* $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \mathfrak{W}_+$.

(ii) By the stable future behavior *of* Σ *we mean the set of all pairs* $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{smallmatrix}\right]$ *which satisfy*

$$\hat{y}(\lambda) = \widehat{\mathfrak{D}}(\lambda)\hat{u}(\lambda), \qquad \lambda \in \Omega. \tag{3.2}$$

We denote this set by $\mathfrak{W}_+^0$*, and call* u *the* input component *and* y *the* output component *of a pair* $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathfrak{W}_+^0$.

Note that we here do not actually define the state component $x(t)$ of the trajectory for $t > 0$, but only for $t = 0$. However, the input u and output y are almost everywhere defined L^2-functions. If Σ is well-posed in the sense of [33, Definition 2.2.1] and if we choose Ω to be the right-half plane $\Omega = \mathbb{C}_\omega := \{s \in \mathbb{C} \mid \Re s > \omega\}$, where ω is the maximum of zero and the growth rate of Σ, then the set $\mathfrak{W}_+$ coincides with the set of all triples $\left[\begin{smallmatrix} x(0) \\ u \\ y \end{smallmatrix}\right]$, where $\left[\begin{smallmatrix} x \\ u \\ y \end{smallmatrix}\right]$ is an generalized trajectory on $\mathbb{R}^+$ (in the sense

of [33]) with the property that both u and y are L^2-functions. In the well-posed case the significance of the set $\mathfrak{W}_+$ in the solution of the future time quadratic cost minimization problem is well understood, at least in the case where the state finite future cost condition holds (see Definition 5.7 below).

DEFINITION 3.3. *The* future cost minimization problem *for* Σ *is the following: Given a vector* $x_0 \in \mathcal{X}$*, find the generalized stable future trajectory of* Σ *with initial state* x_0 *which minimizes the future cost* $J_{\text{fut}}(x_0, u)$ *defined in* (1.3).

In the setting described above, it is easy to solve the future cost minimization problem. It is convenient here to use the language of linear relations in the form presented in [25, Appendix A].

LEMMA 3.4. *The set* $\mathfrak{W}_+$ *of all generalized stable future trajectories of* Σ *is a closed subspace of* $\begin{bmatrix} \mathcal{X} \\ L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix}$.

Proof. For each $\lambda \in \Omega$, the set of all triples which satisfy (3.1) is a closed subspace of $\begin{bmatrix} \mathcal{X} \\ L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix}$. The set $\mathfrak{W}_+$ is the intersection over all $\lambda \in \Omega$ of these subspaces, and hence $\mathfrak{W}_+$ is a closed subspace, too. □

DEFINITION 3.5. *By the* stable s/s (state/signal) output map $\mathfrak{C}$ *of* Σ *we mean the relation* $\mathcal{X} \to \begin{bmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix}$ *whose graph is* $\mathfrak{W}_+$.

LEMMA 3.6. *The output map* $\mathfrak{C}$ *defined above is closed. Thus, for each* $z \in \operatorname{dom}(\mathfrak{C})$ *the set* $\mathfrak{C}z$ *is a closed affine subspace of* $\begin{bmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix}$*. In particular, the multi-valued part* $\mathfrak{C}0 = \mathfrak{W}_+^0$ *of* $\mathfrak{C}$ *is a closed subspace of* $\begin{bmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix}$.

Proof. This follows from the fact that the graph of $\mathfrak{C}$ is closed. □

As any closed relation, $\mathfrak{C}$ has an orthogonal decomposition into a an operator part and a multi-valued part (see [25, Appendix A] for details). The multi-valued part is $\operatorname{mul}(\mathfrak{C}) = \mathfrak{W}_+^0$, and the operator part is the operator $\mathfrak{C}_o = P_{[\mathfrak{W}_+^0]^\perp}\mathfrak{C}$. Note that the domain of $\mathfrak{C}_o$ is the same as the domain of $\mathfrak{C}$. Moreover, for every $z \in \operatorname{dom}(\mathfrak{C})$ the vector $\mathfrak{C}_o z$ is the element in $\mathfrak{C}z$ which has the minimal norm. Thus, we immediately get the following solution to the future cost minimization problem.

THEOREM 3.7. *A necessary and sufficient condition for a vector* $x_0 \in \mathcal{X}$ *to have a finite future cost is that* $x_0 \in \operatorname{dom}(\mathfrak{C})$*. The future cost of* x_0 *is then equal to*

$$\|x_0\|_{\text{fut}}^2 := \inf_{\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathfrak{C}x_0} \left(\|u\|_{L^2(\mathbb{R}^+;\mathcal{U})}^2 + \|y\|_{L^2(\mathbb{R}^+;\mathcal{Y})}^2 \right) = \|\mathfrak{C}_o x_0\|_{\begin{bmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix}}^2, \tag{3.3}$$

and it is achieved for the generalized stable future trajectory $\left[\begin{smallmatrix} x_0 \\ \mathfrak{C}_o x_0 \end{smallmatrix}\right]$.

3.2. The Past Cost Minimization Problem. Also in the past cost minimization problem we must first define the set of trajectories over which we minimize the cost function $J_{\text{past}}(x_0, u)$. We again use a frequency domain approach to define a reasonable set of generalized stable past trajectories. Recall that Ω stands for a particular fixed open subset of $\rho(A) \cap \mathbb{C}^+$ (see Remark 3.1).

Above we commented that the significance of the notion of "stable future trajectories" that we introduced in Definition 3.2 is well understood, at least in the well-posed case. The same statement is no longer true about the set of "stable past trajectories" that we shall introduce in Definition 3.8 below. In the discrete time case we solved the past cost minimization problem by interpreting it as the dual of the future cost minimization problem for the adjoint system. In order to be able to solve the continuous time past cost minimization problem in the same way we must choose

the set of stable past trajectories of the system Σ in such a way that this set is "dual" to the set of all stable future trajectories of Σ^* (in the sense described in Lemma 3.16 below). When Σ is replaced by Σ^*, then the fixed subset Ω in Remark 3.1 must be replaced by $\Omega^* = \{\overline{\lambda} \mid \lambda \in \Omega\}$ (recall that $\rho(A^*) = \{\overline{\lambda} \mid \lambda \in \rho(A)\}$). If, for example, $y_* \in L^2(\mathbb{R}^+; \mathcal{Y})$ is the output of the adjoint system Σ^* and $\lambda \in \Omega^*$, then the Laplace transform $\hat{y}_*(\lambda)$ of a y_* evaluated at λ can be interpreted as the inner product of y_* with the function $t \mapsto \mathrm{e}^{-\lambda t}$. This (combined with a time reflection) motivated us to take a closer look at the set of classical past trajectories of Σ the type

$$\begin{bmatrix} x(t) \\ u(t) \\ y(t) \end{bmatrix} = \mathrm{e}^{\lambda t} \begin{bmatrix} x_0 \\ u_0 \\ y_0 \end{bmatrix}, \qquad t \in \mathbb{R}^-,$$

where $\lambda \in \Omega$. A direct substitution into the appropriate equation shows that this triple of functions is a classical trajectory of Σ on $\mathbb{R}^-$ if and only if

$$\lambda x_0 = A|_{\mathcal{X}} x_0 + B u_0 \text{ and } y_0 = \widehat{\mathfrak{D}}(\lambda) u_0,$$

or equivalently,

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} (\lambda - A|_{\mathcal{X}})^{-1} B \\ \widehat{\mathfrak{D}}(\lambda) \end{bmatrix} u_0. \tag{3.4}$$

DEFINITION 3.8. *For each $\lambda \in \mathbb{C}^+$ we denote the function $t \mapsto \mathrm{e}^{\lambda t}$, $t \in \mathbb{R}^-$, by* $\mathbf{e}_\lambda$.

(i) By the set of classical stable past exponential trajectories *of Σ we mean*

$$\mathfrak{V}_- := \operatorname{span} \left\{ \begin{bmatrix} (\lambda - A|_{\mathcal{X}})^{-1} B u_0 \\ \mathbf{e}_\lambda u_0 \\ \mathbf{e}_\lambda \widehat{\mathfrak{D}}(\lambda) u_0 \end{bmatrix} \middle| \, \lambda \in \Omega, \ u_0 \in \mathcal{U} \right\} \subset \begin{bmatrix} \mathcal{X} \\ L^2(\mathbb{R}^-; \mathcal{U}) \\ L^2(\mathbb{R}^-; \mathcal{Y}) \end{bmatrix}. \tag{3.5}$$

We call x_0 the final state*, u the* input component*, and y the* output component *of a triple $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \mathfrak{V}_-$.*

(ii) By the set of generalized stable past trajectories *of Σ we mean the closure in $\left[\begin{smallmatrix} \mathcal{X} \\ L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{smallmatrix}\right]$ of $\mathfrak{V}_-$. We denote this set by $\mathfrak{W}_-$.*

(iii) By the classical exponential past behavior *of Σ we mean*

$$\mathfrak{V}_-^0 := \operatorname{span} \left\{ \begin{bmatrix} \mathbf{e}_\lambda u_0 \\ \mathbf{e}_\lambda \widehat{\mathfrak{D}}(\lambda) u_0 \end{bmatrix} \middle| \, \lambda \in \Omega, \ u_0 \in \mathcal{U} \right\} \subset \begin{bmatrix} L^2(\mathbb{R}^-; \mathcal{U}) \\ L^2(\mathbb{R}^-; \mathcal{Y}) \end{bmatrix}. \tag{3.6}$$

We call u the input component*, and y the* output component *of a pair $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathfrak{V}_-^0$.*

(iv) By the (generalized) stable past behavior *of Σ we mean the closure in $\left[\begin{smallmatrix} L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{smallmatrix}\right]$ of $\mathfrak{V}_-^0$. We denote this set by $\mathfrak{W}_-^0$.*

Observe that we again ignore the values of the state component $x(t)$ for $t \neq 0$.

DEFINITION 3.9. *The* past cost minimization problem *is the following: Given a vector $x_0 \in \mathcal{X}$, find the generalized stable past trajectory with final state x_0 which minimizes the past cost $J_{\text{past}}(x_0, u)$ defined in* (1.9).

As we saw above, the solution of the future cost minimization problem can be expressed in terms of the (possibly multi-valued) s/s output map $\mathfrak{C}$ of Σ. In the

same way the solution of the past cost minimization can be expressed in terms of the (possibly multi-valued) s/s *input map* $\mathfrak{B}$ of Σ.

DEFINITION 3.10. *By the* stable s/s input map $\mathfrak{B}$ *of* Σ *we mean the relation* $\begin{bmatrix} L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{bmatrix} \to \mathcal{X}$ *whose (inverse) graph is* $\mathfrak{W}_-$.

LEMMA 3.11. *The s/s input map* $\mathfrak{B}$ *defined above is closed, and its domain is a dense subspace of the stable past behavior* $\mathfrak{W}_-^0$. *For each* $x \in \operatorname{img}(\mathfrak{B})$ *the inverse image* $(\mathfrak{B})^{-1}x$ *is a closed affine subspace of* $\begin{bmatrix} L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{bmatrix}$. *In particular, the kernel* $\ker(\mathfrak{B}) := (\mathfrak{B})^{-1}0$ *of* $\mathfrak{B}$ *is a closed subspace of* $\mathfrak{W}_-^0$.

Proof. This all follows from the fact that the graph of $\mathfrak{B}$ is closed, except for the density of the domain, which follows from the fact that $\mathfrak{W}_-^0$ is the closure of $\mathfrak{V}_-^0$. ∎

Since the inverse image $(\mathfrak{B})^{-1}x$ of any $x \in \operatorname{img}(\mathfrak{B})$ is is closed and convex, it has an element of minimal norm, namely $P_{[\ker(\mathfrak{B})]^\perp}(\mathfrak{B})^{-1}x$. Thus, the solution to the past cost minimization problem is the following.

THEOREM 3.12. *A necessary and sufficient condition for a vector* $x_0 \in \mathcal{X}$ *to have a finite past cost is that* $x_0 \in \operatorname{img}(\mathfrak{B})$. *The past cost of* x_0 *is then equal to*

$$\begin{aligned} \|x_0\|_{\text{past}}^2 &:= \inf_{\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in (\mathfrak{B})^{-1}x_0} \left(\|u\|^2_{L^2(\mathbb{R}^-;\mathcal{U})} + \|y\|^2_{L^2(\mathbb{R}^-;\mathcal{Y})} \right) \\ &= \|P_{[\ker(B)]^\perp}(\mathfrak{B})^{-1}x_0\|^2_{\begin{bmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{bmatrix}}, \end{aligned} \tag{3.7}$$

and it is achieved for the generalized stable past trajectory $\begin{bmatrix} x_0 \\ P_{[\ker(B)]^\perp}(\mathfrak{B})^{-1}x_0 \end{bmatrix}$.

3.3. The Two-Sided Cost Minimization Problem. DEFINITION 3.13. *By the set of generalized stable two-sided trajectories of* Σ *we mean the set of all triples* $\begin{bmatrix} x_0 \\ u \\ y \end{bmatrix} \in \begin{bmatrix} \mathcal{X} \\ L^2(\mathbb{R};\mathcal{U}) \\ L^2(\mathbb{R};\mathcal{Y}) \end{bmatrix}$ *for which* $\begin{bmatrix} x_0 \\ \pi_+ u \\ \pi_+ y \end{bmatrix}$ *is a generalized stable future trajectory of* Σ *and* $\begin{bmatrix} x_0 \\ \pi_- u \\ \pi_- y \end{bmatrix}$ *is a generalized stable past trajectory of* Σ *(here* π_+ *and* π_- *are the obvious projections). We denote this set by* $\mathfrak{W}$, *and call* x_0 *the* intermediate state, u *the* input component, *and* y *the* output component *of a triple* $\begin{bmatrix} x_0 \\ u \\ y \end{bmatrix} \in \mathfrak{W}$.

DEFINITION 3.14. *The* two-sided cost minimization problem *is the following: Given a vector* $x_0 \in \mathcal{X}$, *find the generalized stable two-sided trajectory with intermediate state* x_0 *which minimizes the two-sided cost* $J(x_0, u)$ *defined in* (1.11).

The solution to the two-sided cost minimization problem can be derived from the future and past cost minimization problems as follows.

THEOREM 3.15. *A necessary and sufficient condition for a vector* $x_0 \in \mathcal{X}$ *to have a finite two-sided cost is that* $x_0 \in \operatorname{dom}(\mathfrak{C}) \cap \operatorname{img}(\mathfrak{B})$. *The two-sided cost of* x_0 *is then equal to the sum of the future and past costs of* x_0, *and it is achieved for the generalized stable two-sided trajectory* $\begin{bmatrix} x_0 \\ \mathfrak{C}_o x_0 + P_{[\ker(B)]^\perp}(\mathfrak{B})^{-1}x_0 \end{bmatrix}$.

3.4. The Duality Between the Past and Future Cost Minimization Problems. The duality of the future and past cost minimization problems depends on the fact that the set of all generalized stable past trajectories of Σ is the annihilator in a certain sense of the set of all generalized stable future trajectories of the adjoint system $\Sigma^\dagger$ with the fixed open subset Ω of $\rho(A) \cap \mathbb{C}^+$ replaced by the reflected set $\Omega^* = \{\overline{\lambda} \mid \lambda \in \Omega\}$ of $\rho(A^*) \cap \mathbb{C}^+$. To describe this connection we denote $\mathfrak{K} := \begin{bmatrix} \mathcal{X} \\ L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{bmatrix}$ and $\mathfrak{K}^\dagger := \begin{bmatrix} \mathcal{X} \\ L^2(\mathbb{R}^+;\mathcal{Y}) \\ L^2(\mathbb{R}^+;\mathcal{U}) \end{bmatrix}$ and identify the dual of $\mathfrak{K}$ with $\mathfrak{K}^\dagger$ by means

of the duality pairing

$$\left\langle \begin{bmatrix} x \\ u \\ y \end{bmatrix}, \begin{bmatrix} x^\dagger \\ y^\dagger \\ u^\dagger \end{bmatrix} \right\rangle_{\mathfrak{K},\mathfrak{K}^\dagger} = \langle x, x^\dagger \rangle_{\mathcal{X}} - \int_{-\infty}^0 \langle u(s), u^\dagger(-s) \rangle_{\mathcal{U}}\, ds + \int_{-\infty}^0 \langle y(s), y^\dagger(-s) \rangle_{\mathcal{Y}}\, ds. \tag{3.8}$$

We further denote $\mathfrak{K}_0 := \begin{bmatrix} L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{bmatrix}$ and $\mathfrak{K}_0^\dagger := \begin{bmatrix} L^2(\mathbb{R}^+;\mathcal{Y}) \\ L^2(\mathbb{R}^+;\mathcal{U}) \end{bmatrix}$, and use the corresponding duality pairing

$$\left\langle \begin{bmatrix} u \\ y \end{bmatrix}, \begin{bmatrix} y^\dagger \\ u^\dagger \end{bmatrix} \right\rangle_{\mathfrak{K}_0,\mathfrak{K}_0^\dagger} = -\int_{-\infty}^0 \langle u(s), u^\dagger(-s) \rangle_{\mathcal{U}}\, ds + \int_{-\infty}^0 \langle y(s), y^\dagger(-s) \rangle_{\mathcal{Y}}\, ds. \tag{3.9}$$

Lemma 3.16.

(i) The annihilator of the set $\mathfrak{V}_-$ of all classical stable past exponential trajectories of Σ with respect to the duality pairing (3.8) is the set $\mathfrak{W}_+^\dagger$ of all generalized stable future trajectories of the adjoint system $\Sigma^\dagger$ induced by the reflected subset Ω^* of $\rho(A^*) \cap \mathbb{C}^+$.

(ii) The annihilator of the set $\mathfrak{W}_+^\dagger$ of all generalized stable future trajectories of the adjoint system $\Sigma^\dagger$ induced by the reflected subset Ω^* of $\rho(A^*) \cap \mathbb{C}^+$ with respect to the duality pairing (3.8) is the set $\mathfrak{W}_-$ of all generalized stable past trajectories of Σ.

(iii) The annihilator of the classical exponential past behavior $\mathfrak{V}_-$ of Σ with respect to the duality pairing (3.9) is the stable future behavior $\mathfrak{W}_+^{0\dagger}$ of the adjoint system $\Sigma^\dagger$ induced by the reflected subset Ω^* of $\rho(A^*) \cap \mathbb{C}^+$.

(iv) The annihilator of the stable future behavior $\mathfrak{W}_+^{0\dagger}$ of the adjoint system $\Sigma^\dagger$ induced by the component Ω^* of $\rho(A^*) \cap \mathbb{C}^+$ is equal to the stable past behavior $\mathfrak{W}_-^0$ of Σ.

Proof. Clearly (ii) follows from (i), and (iv) follows from (iii), so it suffices to prove (i) and (iii). The proof of (iii) is very similar to the proof of (i), so here we only given the proof of (i).

By (3.8), a triple $\begin{bmatrix} x^\dagger \\ y^\dagger \\ u^\dagger \end{bmatrix}$ is orthogonal to $\begin{bmatrix} (\lambda - A|_{\mathcal{X}})^{-1} B u_0 \\ \mathrm{e}^{\lambda t} u_0 \\ \mathrm{e}^{\lambda t} \widehat{\mathfrak{D}}(\lambda) u_0 \end{bmatrix} \in \mathfrak{V}_-$ if and only if

$$\begin{aligned} 0 =& \langle (\lambda - A|_{\mathcal{X}})^{-1} B u_0, x^\dagger \rangle_{\mathcal{X}} - \int_{-\infty}^0 \langle \mathrm{e}^{\lambda s} u_0, u^\dagger(-s) \rangle_{\mathcal{U}}\, ds \\ &+ \int_{-\infty}^0 \langle \mathrm{e}^{\lambda s} \widehat{\mathfrak{D}}(\lambda) u_0, y^\dagger(-s) \rangle_{\mathcal{Y}}\, ds \\ =& \langle u_0, B^*(\overline{\lambda} - A^*)^{-1} x^\dagger \rangle_{\mathcal{U}} - \left\langle u_0, \int_0^\infty \mathrm{e}^{-\overline{\lambda} s} u^\dagger(s)\, ds \right\rangle_{\mathcal{U}} \\ &+ \left\langle u_0, \widehat{\mathfrak{D}}(\lambda)^* \int_0^\infty \mathrm{e}^{-\overline{\lambda} s} y^\dagger(s)\, ds \right\rangle_{\mathcal{U}} \\ =& \langle u_0, B^*(\overline{\lambda} - A^*)^{-1} x^\dagger \rangle_{\mathcal{U}} - \langle u_0, \hat{u}(\overline{\lambda}) \rangle_{\mathcal{U}} + \langle u_0, \widehat{\mathfrak{D}}(\lambda)^* \hat{y}(\overline{\lambda}) \rangle_{\mathcal{U}}. \end{aligned}$$

This is true for all $u_0 \in \mathcal{U}$ if and only if

$$\hat{u}(\overline{\lambda}) = \widehat{\mathfrak{D}}(\lambda)^* \hat{y}(\overline{\lambda}) + B^*(\overline{\lambda} - A^*)^{-1} x^\dagger,$$

which is true for all $\lambda \in \Omega$ if and only if $\begin{bmatrix} x^\dagger \\ y^\dagger \\ u^\dagger \end{bmatrix} \in \mathfrak{W}_+^\dagger$. □

As we shall see below, it follows from Lemma 3.16 that the past cost for the given system Σ is the inverse and the dual (in to be described senses) of the future cost for the adjoint system $\Sigma^\dagger$. The proof of this fact is based on the following lemma (which also describes the notion of inverse and dual of a nonnegative quadratic form). The consequences for the optimal control problems are drawn in Theorem 3.18.

LEMMA 3.17. *Let $\mathcal{X}$ and $\mathcal{W}$ be Hilbert spaces, and let V be a closed subspace of $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{W}\end{smallmatrix}\right] = \mathcal{X} \oplus \mathcal{W}$. We denote the orthogonal projections in $\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{W}\end{smallmatrix}\right]$ onto $\mathcal{X}$ and $\mathcal{W}$ by $P_{\mathcal{X}}$ and $P_{\mathcal{W}}$, respectively. Let*

$$\begin{aligned}\mathcal{X}_V &:= \left\{x \in \mathcal{X} \,\middle|\, \left[\begin{smallmatrix}x\\ w\end{smallmatrix}\right] \in V \text{ for some } w \in \mathcal{W}\right\},\\ \mathcal{X}_{V^\perp} &:= \left\{x^\dagger \in \mathcal{X} \,\middle|\, \left[\begin{smallmatrix}x^\dagger\\ w^\dagger\end{smallmatrix}\right] \in V^\perp \text{ for some } w^\dagger \in \mathcal{W}\right\},\end{aligned}$$

and define

$$\begin{aligned}\|x\|_V &:= \inf\left\{\|w\| \,\middle|\, \left[\begin{smallmatrix}x\\ w\end{smallmatrix}\right] \in V\right\}, && x \in \mathcal{X}_V,\\ \|x^\dagger\|_{V^\perp} &:= \inf\left\{\|w^\dagger\| \,\middle|\, \left[\begin{smallmatrix}x^\dagger\\ w^\dagger\end{smallmatrix}\right] \in V^\perp\right\}, && x^\dagger \in \mathcal{X}_{V^\perp}.\end{aligned} \tag{3.10}$$

Then the following claims are true.

(i) $\|\cdot\|_V^2$ and $\|\cdot\|_{V^\perp}^2$ are closed nonnegative quadratic forms in $\mathcal{X}$.

(ii) The forms $\|\cdot\|_V^2$ and $\|\cdot\|_{V^\perp}^2$ are inverses of each other in the following sense. If we denote the self-adjoint relations in $\mathcal{X}$ that induce the quadratic forms $\|\cdot\|_V^2$ and $\|\cdot\|_{V^\perp}^2$ by Q_V and $Q_{V^\perp}$, respectively, i.e.,

$$\begin{aligned}\|x\|_V^2 &= \langle x, Q_V x\rangle_{\mathcal{X}}, \quad x \in \operatorname{dom}(Q_V) \subset \operatorname{dom}\left(Q_V^{1/2}\right) = \mathcal{X}_V,\\ \|x\|_{V^\perp}^2 &= \langle x, Q_{V^\perp} x\rangle_{\mathcal{X}}, \quad x \in \operatorname{dom}(Q_{V^\perp}) \subset \operatorname{dom}\left(Q_{V^\perp}^{1/2}\right) = \mathcal{X}_{V^\perp},\end{aligned}$$

then $Q_{V^\perp} = Q_V^{-1}$.

(iii) The forms $\|\cdot\|_V^2$ and $\|\cdot\|_{V^\perp}^2$ are dual to each other in the sense that

$$\begin{aligned}\operatorname{dom}(\|\cdot\|_V) &= \left\{x \in \mathcal{X} \,\middle|\, \sup_{x^\dagger \in \mathcal{X}_{V^\perp}, \|x^\dagger\|_{V^\perp} \le 1} |\langle x, x^\dagger\rangle_{\mathcal{X}}| < \infty\right\},\\ \operatorname{dom}(\|\cdot\|_{V^\perp}) &= \left\{x^\dagger \in \mathcal{X} \,\middle|\, \sup_{x \in \mathcal{X}_V, \|x\|_V \le 1} |\langle x, x^\dagger\rangle_{\mathcal{X}}| < \infty\right\},\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}\|x\|_V &= \sup_{x^\dagger \in \mathcal{X}_{V^\perp}, \|x^\dagger\|_{V^\perp} \le 1} |\langle x, x^\dagger\rangle_{\mathcal{X}}|, && x \in \mathcal{X}_V,\\ \|x^\dagger\|_{V^\perp} &= \sup_{x \in \mathcal{X}_V, \|x\|_V \le 1} |\langle x, x^\dagger\rangle_{\mathcal{X}}|, && x^\dagger \in \mathcal{X}_{V^\perp}.\end{aligned} \tag{3.12}$$

The proof of Lemma 3.17 is given in Appendix A.

THEOREM 3.18. *Let $\|\cdot\|_{\text{past}}^2$ be the past cost defined in Theorem 3.12, and let $\|\cdot\|_{\text{fut}^\dagger}^2$ be the future cost defined in Theorem 3.7 applied to the adjoint system $\Sigma^\dagger$. Then $\|\cdot\|_{\text{past}}^2$ and $\|\cdot\|_{\text{fut}^\dagger}^2$ are closed nonnegative quadratic forms in $\mathcal{X}$ which are dual to each other and inverses of each other in the senses described in Lemma 3.17.*

Proof. We replace $\mathcal{W}$ in Lemma 3.17 by $\left[\begin{smallmatrix}L^2(\mathbb{R}^-;\mathcal{U})\\ L^2(\mathbb{R}^-;\mathcal{Y})\end{smallmatrix}\right]$ and take $V = \mathfrak{W}_-$. By that lemma, the costs $\|\cdot\|_V^2$ and $\|\cdot\|_{V^\perp}^2$ are duals and inverses of each other. It follows

immediately from the definition of $\|\cdot\|_V^2$ and Theorem 3.12 that $\|\cdot\|_V^2 = \|\cdot\|_{\text{past}}^2$. We claim that $\|\cdot\|_{V^\perp}^2 = \|\cdot\|_{\text{fut}^\dagger}^2$. To see that this is true it suffices to observe that by definition $\|\cdot\|_{V^\perp}^2$ is the cost induced by the subspace $V^\perp = \mathfrak{W}_-^\perp$ of $\mathcal{X} \oplus \left[\begin{smallmatrix} L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{smallmatrix}\right]$, whereas, by Theorem 3.12, $\|\cdot\|_{\text{fut}^\dagger}^2$ is the cost induced by the subspace $\mathfrak{W}_+^\dagger$ of $\mathcal{X} \oplus \left[\begin{smallmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{smallmatrix}\right]$ for the adjoint system. By Lemma 3.16, $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \mathfrak{W}_-^\perp$ if and only if $\left[\begin{smallmatrix} x_0 \\ -\boldsymbol{\mathfrak{R}} u \\ \boldsymbol{\mathfrak{R}} y \end{smallmatrix}\right] \in \mathfrak{W}_+$, where $\boldsymbol{\mathfrak{R}}$ is the reflection operator $(\boldsymbol{\mathfrak{R}}u)(t) = u(-t)$, $t \in \mathbb{R}^-$. Since the norm of $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ in $\left[\begin{smallmatrix} L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{smallmatrix}\right]$ is equal to the norm of $\left[\begin{smallmatrix} -\boldsymbol{\mathfrak{R}} u \\ \boldsymbol{\mathfrak{R}} y \end{smallmatrix}\right]$ in $\left[\begin{smallmatrix} L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{smallmatrix}\right]$, we find that, indeed, $\|\cdot\|_{V^\perp}^2 = \|\cdot\|_{\text{fut}^\dagger}^2$. □

4. The Connection to the Discrete Time Cost Minimization Problem. The continuous time future and past cost minimization problems described above for the system $\Sigma = (\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right], \mathcal{X}, \mathcal{U}, \mathcal{Y})$ can be connected to discrete time problems in the following way. We fix some $\alpha \in \Omega$, define $\mathbf{A}_\alpha$, $\mathbf{B}_\alpha$, $\mathbf{C}_\alpha$, and $\mathbf{D}_\alpha$ by

$$\begin{aligned} \mathbf{A}_\alpha &= (\overline{\alpha} + A)(\alpha - A)^{-1}, & \mathbf{B}_\alpha &= \sqrt{2\Re\alpha}\,(\alpha - A|_{\mathcal{X}})^{-1}B, \\ \mathbf{C}_\alpha &= \sqrt{2\Re\alpha}\,C(\alpha - A)^{-1}, & \mathbf{D}_\alpha &= \widehat{\mathfrak{D}}(\alpha), \end{aligned} \tag{4.1}$$

and consider the discrete time system

$$\boldsymbol{\Sigma}_\alpha : \begin{cases} x_{n+1} = \mathbf{A}_\alpha x_n + \mathbf{B}_\alpha u_n, \\ y_n = \mathbf{C}_\alpha x_n + \mathbf{D}_\alpha u_n, \end{cases} \tag{4.2}$$

with these coefficients. We denote this system by $\boldsymbol{\Sigma}_\alpha$ and call it the (internal) Cayley transform (with parameter α) of Σ. The (discrete time) *transfer function* of $\boldsymbol{\Sigma}_\alpha$ is the function

$$\widehat{\mathbf{D}}_\alpha(z) = z\mathbf{C}_\alpha(1_{\mathcal{X}} - z\mathbf{A}_\alpha)^{-1}\mathbf{B}_\alpha + \mathbf{D}_\alpha, \qquad z \in \Lambda(\mathbf{A}_\alpha), \tag{4.3}$$

where $\Lambda(\mathbf{A}_\alpha)$ is the Fredholm resolvent set of $\mathbf{A}_\alpha$, i.e., the set of point $z \in \mathbb{C}$ for which the operator $1_{\mathcal{X}} - z\mathbf{A}_\alpha$ has a bounded inverse. The transform $\lambda \mapsto z$, where

$$\lambda = \frac{\alpha - \overline{\alpha}z}{1 + z}, \quad z = \frac{\alpha - \lambda}{\overline{\alpha} + \lambda}, \quad \lambda \in \rho(A), \quad z \in \Lambda(\mathbf{A}_\alpha), \tag{4.4}$$

maps $\rho(A)$ one-to-one onto $\Lambda(\mathbf{A}_\alpha)$ if A is unbounded, and it maps $\rho(A) \cup \{\infty\}$ one-to-one onto $\Lambda(\mathbf{A}_\alpha)$ if A is bounded. The connection between the resolvent of A and the Fredholm resolvent of $\mathbf{A}_\alpha$ is

$$z(1_{\mathcal{X}} - z\mathbf{A}_\alpha)^{-1} = \frac{\alpha - \lambda}{2\Re\alpha}(\alpha - A)(\lambda - A)^{-1}, \quad \lambda \in \rho(A), \quad z \in \Lambda(\mathbf{A}_\alpha), \tag{4.5}$$

and the continuous time transfer function $\widehat{\mathfrak{D}}$ is related to the discrete time transfer function $\widehat{\mathbf{D}}_\alpha$ by

$$\widehat{\mathbf{D}}_\alpha(z) = \widehat{\mathfrak{D}}(\lambda), \quad \lambda \in \rho(A), \quad z \in \Lambda(\mathbf{A}_\alpha), \tag{4.6}$$

The same transformation maps $\mathbb{C}^+$ one-to-one onto the unit disk $\mathbb{D} := \{z \in \mathbb{C} \mid |z| < 1\}$. Thus

$$\Lambda(\mathbf{A}_\alpha) \cap \mathbb{D} = \left\{ \frac{\alpha - \lambda}{\overline{\alpha} + \lambda} \middle| \, \lambda \in \rho(A) \cap \mathbb{C}^+ \right\}.$$

We denote the image under this transformation of the subset Ω in Remark 3.1 by Ω_α. Then

$$\{0\} \subset \Omega_\alpha = \left\{ \frac{\alpha - \lambda}{\overline{\alpha} + \lambda} \middle| \lambda \in \Omega \right\} \subset \Lambda(\mathbf{A}_\alpha) \cap \mathbb{D}.$$

Moreover, Ω_α is connected if and only if Ω is connected. See [33, Chapter 12] for more details. Note, however, that the Cayley transform in [33] is defined in a slightly different way, so that to pass from the formulas given in [33] to the formulas used here one needs to replace z by $1/z$ (and to replace the exterior of the unit disk by the unit disk itself).

4.1. The Discrete Time Future Cost Minimization Problem. We begin by solving (4.2) with a given initial state $x_0 \in \mathcal{X}$ and a given input sequence $u \in \mathcal{U}^{\mathbb{Z}^+}$. By solving (4.2) recursively we get

$$y_n = \mathbf{D}_\alpha u_n + \sum_{k=0}^{n-1} \mathbf{C}_\alpha \mathbf{A}_\alpha^k \mathbf{B}_\alpha u_{n-k-1} + \mathbf{C}_\alpha \mathbf{A}_\alpha^n x_0, \qquad n \in \mathbb{Z}^+. \tag{4.7}$$

Definition 4.1.

(i) By the set of stable future trajectories of $\mathbf{\Sigma}_\alpha$ we mean the set of all triples $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \ell^2(\mathbb{Z}^+;\mathcal{U}) \\ \ell^2(\mathbb{Z}^+;\mathcal{Y}) \end{smallmatrix}\right]$ which satisfy (4.7). We denote this set by $\mathfrak{W}_{\alpha+}$, and call x_0 the initial state*, u the* input component*, and y the* output component *of a triple $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \mathfrak{W}_{\alpha+}$.*

(ii) By the stable future behavior *of $\mathbf{\Sigma}_\alpha$ we mean the set of all pairs $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \ell^2(\mathbb{Z}^+;\mathcal{U}) \\ \ell^2(\mathbb{Z}^+;\mathcal{Y}) \end{smallmatrix}\right]$ which satisfy (4.7) with $x_0 = 0$. We denote this set by $\mathfrak{W}^0_{\alpha+}$, and call u the* input component *and y the* output component *of a pair $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathfrak{W}^0_{\alpha+}$.*

Definition 4.2. *The* future cost minimization problem *for $\mathbf{\Sigma}_\alpha$ is the following: Given a vector $x_0 \in \mathcal{X}$, find the stable future trajectory of $\mathbf{\Sigma}_\alpha$ with initial state x_0 which minimizes*

$$\mathbf{J}_\alpha^{\mathrm{fut}}(x_0, u) := \|u\|^2_{\ell^2(\mathbb{Z}^+;\mathcal{U})} + \|y\|^2_{\ell^2(\mathbb{Z}^+;\mathcal{Y})}.$$

Lemma 4.3. *If Ω is connected, then the set $\mathfrak{W}^+_\alpha$ of all stable future trajectories of $\mathbf{\Sigma}_\alpha$ has the following alternative characterization: $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \ell^2(\mathbb{Z}^+;\mathcal{U}) \\ \ell^2(\mathbb{Z}^+;\mathcal{Y}) \end{smallmatrix}\right]$ is a stable future trajectory of $\mathbf{\Sigma}_\alpha$ if and only if*

$$\hat{y}(z) = \widehat{\mathbf{D}}_\alpha(z)\hat{u}(z) + \mathbf{C}_\alpha(1_\mathcal{X} - z\mathbf{A}_\alpha)^{-1}x_0, \qquad z \in \Omega_\alpha, \tag{4.8}$$

where $\hat{u}$ and $\hat{z}$ are the Z-transforms of u and y, defined for all $z \in \mathbf{D}$ by $\hat{u}(z) = \sum_{k\in\mathbb{Z}^+} z^k u_k$ and $\hat{y}(z) = \sum_{k\in\mathbb{Z}^+} z^k y_k$.

Proof. Let $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right]$ be a stable future trajectory of $\mathbf{\Sigma}_\alpha$. Then x is power bounded, and the Z-transform $\hat{x}(z)$ of x converges for all z in some (sufficiently small) neighborhood $\mathcal{O}$ of the origin. By multiplying the two equations in (4.2) by z^n, adding over $n \in \mathbb{Z}^+$, and simplifying the result we find that (4.8) holds for all $z \in \mathcal{O}$. Since Ω is connected also Ω_α is connected, and since both sides of (4.8) are analytic in Ω_α, the same identity must then hold for all $z \in \Omega_\alpha$.

The converse direction follows from the fact that a sequence in $\ell^2(\mathbb{Z}^+)$ is uniquely determined by its Z-transform. □

LEMMA 4.4. *If Ω is connected, then the following conditions are equivalent:*

(i) The triple $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ L^2(\mathbb{R}^+;\mathcal{U}) \\ L^2(\mathbb{R}^+;\mathcal{Y}) \end{smallmatrix}\right]$ is a generalized stable future trajectory of Σ.

(ii) For some $\alpha \in \Omega$, the triple $\left[\begin{smallmatrix} x_0 \\ \mathcal{L}_\alpha u \\ \mathcal{L}_\alpha y \end{smallmatrix}\right]$ is a stable future trajectory of $\mathbf{\Sigma}_\alpha$, where $\mathcal{L}_\alpha$ is the Laguerre transform with parameter α (see [33, Definition 12.3.2]).

(iii) For all $\alpha \in \Omega$, the triple $\left[\begin{smallmatrix} x_0 \\ \mathcal{L}_\alpha u \\ \mathcal{L}_\alpha y \end{smallmatrix}\right]$ is a stable future trajectory of $\mathbf{\Sigma}_\alpha$.

Proof. Trivially (iii) ⇒ (ii). That (i) ⇒ (iii) and that (ii) ⇒ (i) follows from Definition 3.2, Lemma 4.3, equations 4.1, (4.5), (4.6), and [33, Theorem 12.3.1]. □

THEOREM 4.5. *If Ω is connected, then for each $x_0 \in \mathcal{X}$, the future continuous time cost of x_0 is finite if and only if the future discrete time α-cost is finite for some, or equivalently, for all $\alpha \in \Omega$. Moreover, the optimal costs for all these problems are the same.*

Proof. This follows from Lemma 4.4 and the fact that the Laguerre transform is a unitary map of $\left[\begin{smallmatrix} L^2(\mathbb{R}^+,\mathcal{U}) \\ L^2(\mathbb{R}^+,\mathcal{Y}) \end{smallmatrix}\right]$ onto $\left[\begin{smallmatrix} \ell^2(\mathbb{Z}^+,\mathcal{U}) \\ \ell^2(\mathbb{Z}^+,\mathcal{Y}) \end{smallmatrix}\right]$, and hence the cost of a generalized continuous time trajectory $\left[\begin{smallmatrix} z \\ u \\ y \end{smallmatrix}\right]$ of Σ coincides with the cost of the transformed trajectory $\left[\begin{smallmatrix} z \\ \mathcal{L}_\alpha u \\ \mathcal{L}_\alpha y \end{smallmatrix}\right]$ of $\mathbf{\Sigma}_\alpha$. □

4.2. The Discrete Time Past Cost Minimization Problem. In the discrete time past cost minimization problem we start with the case where we have a trajectory of (4.2) on $\mathbb{Z}^-$ whose support is bounded to the left. Using the fact that $x_n = 0$ and $u_n = 0$ for sufficiently large negative n we can solve x_0 and y from (4.2) to get

$$\begin{aligned} x_0 &= \sum_{k=0}^{\infty} \mathbf{A}_\alpha^k \mathbf{B}_\alpha u_{-k-1}, \\ y_n &= \mathbf{D}_\alpha u_n + \sum_{k=0}^{\infty} \mathbf{C}_\alpha \mathbf{A}_\alpha^k \mathbf{B}_\alpha u_{n-k-1}, \qquad n \in \mathbb{Z}^-. \end{aligned} \tag{4.9}$$

DEFINITION 4.6.

(i) By the set of compactly supported past trajectories of $\mathbf{\Sigma}_\alpha$ we mean the set of all triples $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \left[\begin{smallmatrix} \mathcal{X} \\ \ell^2(\mathbb{Z}^-;\mathcal{U}) \\ \ell^2(\mathbb{Z}^-;\mathcal{Y}) \end{smallmatrix}\right]$, where the supports of y and u are bounded to the left, which satisfy (4.9). *We denote this set by $\mathfrak{V}_{\alpha-}$, and call x_0 the* final state, *u the* input component, *and y the* output component *of a triple $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right] \in \mathfrak{V}_{\alpha-}$.*

(ii) By the set of generalized stable past trajectories of $\mathbf{\Sigma}_\alpha$ we mean the closure in $\left[\begin{smallmatrix} \mathcal{X} \\ \ell^2(\mathbb{Z}^-;\mathcal{U}) \\ \ell^2(\mathbb{Z}^-;\mathcal{Y}) \end{smallmatrix}\right]$ of $\mathfrak{V}_{\alpha-}$. We denote this set by $\mathfrak{W}_{\alpha-}$.

(iii) By the compactly supported past behavior *of $\mathbf{\Sigma}_\alpha$ we mean the set of compactly supported sequences in $\left[\begin{smallmatrix} \ell^2(\mathbb{Z}^-;\mathcal{U}) \\ \ell^2(\mathbb{Z}^-;\mathcal{Y}) \end{smallmatrix}\right]$ which satisfy the second equation in* (4.9). *We denote this set by $\mathfrak{V}^0_{\alpha-}$, and call u the* input component *and y the* output component *of a pair $\left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right] \in \mathfrak{V}^0_{\alpha-}$.*

(iv) By the (generalized) stable past behavior *of $\mathbf{\Sigma}_\alpha$ we mean the closure in $\left[\begin{smallmatrix} \ell^2(\mathbb{Z}^-;\mathcal{U}) \\ \ell^2(\mathbb{Z}^-;\mathcal{Y}) \end{smallmatrix}\right]$ of $\mathfrak{V}^0_{\alpha-}$. We denote this set by $\mathfrak{W}^0_{\alpha-}$.*

DEFINITION 4.7. *The* past cost minimization problem *for* $\mathbf{\Sigma}_\alpha$ *is the following: Given a vector* $x_0 \in \mathcal{X}$*, find the generalized stable past trajectory of* $\mathbf{\Sigma}_\alpha$ *with final state* x_0 *which minimizes*

$$\mathbf{J}_\alpha^{\text{past}}(x_0, u) := \|u\|^2_{\ell^2(\mathbb{Z}^-;\mathcal{U})} + \|y\|^2_{\ell^2(\mathbb{Z}^-;\mathcal{Y})}.$$

Remark 4.8. *We note that in [24] the compactly support past behavior was denoted by* $\mathscr{G}_c$ *and the generalized stable past behavior by* $\mathscr{G}$*. The operator denoted* $\mathcal{J}$ *in that article is the discrete-time equivalent of the s/s input map* $\mathfrak{B}$ *defined in Definition 3.10.*

4.3. The Duality Between the Discrete Time Past and Future Cost Minimization Problems. As we shall see below, the discrete time past cost minimization problem is dual to the future cost minimization problem in a well-defined sense (this is the discrete-time analogue of Lemma 3.16). To describe this connection we denote $\mathfrak{K} := \begin{bmatrix} \mathcal{X} \\ \ell^2(\mathbb{Z}^-;\mathcal{U}) \\ \ell^2(\mathbb{Z}^-;\mathcal{Y}) \end{bmatrix}$ and $\mathfrak{K}^\dagger := \begin{bmatrix} \mathcal{X} \\ \ell^2(\mathbb{Z}^+;\mathcal{Y}) \\ \ell^2(\mathbb{Z}^+;\mathcal{U}) \end{bmatrix}$ and identify the dual of $\mathfrak{K}$ with $\mathfrak{K}^\dagger$ by means of the duality pairing

$$\left\langle \begin{bmatrix} x \\ u \\ y \end{bmatrix}, \begin{bmatrix} x^\dagger \\ y^\dagger \\ u^\dagger \end{bmatrix} \right\rangle_{\mathfrak{K},\mathfrak{K}^\dagger} = \langle x, x^\dagger\rangle_{\mathcal{X}} - \sum_{n=0}^{\infty} \langle u_{-n-1}, u_n^\dagger\rangle_{\mathcal{U}} + \sum_{n=0}^{\infty} \langle y_{-n-1}, y_n^\dagger\rangle_{\mathcal{Y}}. \tag{4.10}$$

We further denote $\mathfrak{K}_0 := \begin{bmatrix} \ell^2(\mathbb{Z}^-;\mathcal{U}) \\ \ell^2(\mathbb{Z}^-;\mathcal{Y}) \end{bmatrix}$ and $\mathfrak{K}^{0\dagger} := \begin{bmatrix} \ell^2(\mathbb{Z}^+;\mathcal{Y}) \\ \ell^2(\mathbb{Z}^+;\mathcal{U}) \end{bmatrix}$, and use the corresponding duality pairing

$$\left\langle \begin{bmatrix} u \\ y \end{bmatrix}, \begin{bmatrix} y^\dagger \\ u^\dagger \end{bmatrix} \right\rangle_{\mathfrak{K}^0,\mathfrak{K}^{0\dagger}} = -\sum_{n=0}^{\infty} \langle u_{-n-1}, u_n^\dagger\rangle_{\mathcal{U}} + \sum_{n=0}^{\infty} \langle y_{-n-1}, y_n^\dagger\rangle_{\mathcal{Y}}. \tag{4.11}$$

LEMMA 4.9.

(i) The annihilator of the set $\mathfrak{V}_{\alpha-}$ *of all compactly supported past trajectories of* $\mathbf{\Sigma}_\alpha$ *with respect to the duality pairing* (4.10) *is the set* $\mathfrak{W}^\dagger_{\alpha+}$ *of all generalized stable future trajectories of the adjoint system* $\mathbf{\Sigma}^\dagger_\alpha$*.*

(ii) The annihilator of the set $\mathfrak{W}^\dagger_{\alpha+}$ *of all generalized stable future trajectories of the adjoint system* $\mathbf{\Sigma}^\dagger_\alpha$ *with respect to the duality pairing* (4.10) *is the set* $\mathfrak{W}_{\alpha-}$ *of all stable past trajectories of* $\mathbf{\Sigma}_\alpha$*.*

(iii) The annihilator of the compactly supported past behavior $\mathfrak{V}^0_{\alpha-}$ *of* $\mathbf{\Sigma}_\alpha$ *with respect to the duality pairing* (4.11) *is the stable future behavior* $\mathfrak{W}^{0\dagger}_{\alpha+}$ *of the adjoint system* $\mathbf{\Sigma}^\dagger_\alpha$*.*

(iv) The annihilator of the stable future behavior $\mathfrak{W}^{0\dagger}_{\alpha+}$ *of the adjoint system* $\mathbf{\Sigma}^\dagger_\alpha$ *is the stable past behavior* $\mathfrak{V}^0_{\alpha-}$ *of* $\mathbf{\Sigma}_\alpha$*.*

Proof. As in the case of Lemma 3.16 we only prove (i) and leave the remaining proofs to the reader.

By (4.10), a triple $\begin{bmatrix} x^\dagger \\ u^\dagger \\ y^\dagger \end{bmatrix}$ is orthogonal to $\begin{bmatrix} x_0 \\ u \\ y \end{bmatrix} \in \mathfrak{W}_{\alpha-}$ if and only if

$$\begin{aligned}
0 &= \sum_{k=0}^{\infty} \left\langle \mathbf{A}_\alpha^k \mathbf{B}_\alpha u_{-k-1}, x^\dagger \right\rangle_{\mathcal{X}} - \sum_{n=0}^{\infty} \langle u_{-n-1}, u_n^\dagger \rangle_{\mathcal{U}} \\
&\qquad + \sum_{n=0}^{\infty} \langle \mathbf{D}_\alpha u_{-n-1}, y_n^\dagger \rangle_{\mathcal{Y}} + \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \langle \mathbf{C}_\alpha \mathbf{A}_\alpha^k \mathbf{B}_\alpha u_{-2-n-k}, y_n^\dagger \rangle_{\mathcal{Y}} \\
&= \sum_{n=0}^{\infty} \left\langle u_{-n-1}, \mathbf{B}_\alpha^* \mathbf{A}_\alpha^{*n} x^\dagger + \mathbf{D}_\alpha^* y_n^\dagger - u_n^\dagger \right\rangle_{\mathcal{X}} \\
&\qquad + \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} \langle u_{-n-1}, \mathbf{B}_\alpha^* \mathbf{A}_\alpha^{*k} \mathbf{C}_\alpha^* y_{n-k-1}^\dagger \rangle_{\mathcal{Y}} \\
&= \sum_{n=0}^{\infty} \left\langle u_{-n-1}, \mathbf{B}_\alpha^* \mathbf{A}_\alpha^{*n} x^\dagger + \mathbf{D}_\alpha^* y_n^\dagger - u_n^\dagger + \sum_{k=0}^{n-1} \mathbf{B}_\alpha^* \mathbf{A}_\alpha^{*k} \mathbf{C}_\alpha^* y_{n-k-1}^\dagger \right\rangle_{\mathcal{X}}.
\end{aligned}$$

This is true for all sequences u with finite support if and only if

$$u_n^\dagger = \mathbf{D}_\alpha^* y_n^\dagger + \sum_{k=0}^{n-1} \mathbf{B}_\alpha^* \mathbf{A}_\alpha^{*k} \mathbf{C}_\alpha^* y_{n-k-1}^\dagger + \mathbf{B}_\alpha^* \mathbf{A}_\alpha^{*n} x^\dagger,$$

which is equivalent to the condition $\begin{bmatrix} x^\dagger \\ y^\dagger \\ u^\dagger \end{bmatrix} \in \mathfrak{W}_{\alpha+}^\dagger$. □

LEMMA 4.10. *If Ω is connected, then the following conditions are equivalent:*

(i) The triple $\begin{bmatrix} x_0 \\ u \\ y \end{bmatrix} \in \begin{bmatrix} \mathcal{X} \\ L^2(\mathbb{R}^-;\mathcal{U}) \\ L^2(\mathbb{R}^-;\mathcal{Y}) \end{bmatrix}$ is a generalized stable past trajectory of Σ.

(ii) For some $\alpha \in \Omega$, the triple $\begin{bmatrix} x_0 \\ \mathcal{L}_\alpha u \\ \mathcal{L}_\alpha y \end{bmatrix}$ is a stable past trajectory of $\mathbf{\Sigma}_\alpha$, where $\mathcal{L}_\alpha$ is the Laguerre transform with parameter α (see [33, Definition 12.3.2]).

(iii) For all $\alpha \in \Omega$, the triple $\begin{bmatrix} x_0 \\ \mathcal{L}_\alpha u \\ \mathcal{L}_\alpha y \end{bmatrix}$ is a stable past trajectory of $\mathbf{\Sigma}_\alpha$.

Proof. We map $L^2(\mathbb{R}^-;\mathcal{U})$ onto $\ell^2(\mathbb{Z}^-;\mathcal{U})$ and $L^2(\mathbb{R}^-;\mathcal{Y})$ onto $\ell^2(\mathbb{Z}^-;\mathcal{Y})$ by using the Laguerre transform $\mathcal{L}_\alpha$ (restricted to negative time), and we map $L^2(\mathbb{R}^+;\mathcal{U})$ onto $\ell^2(\mathbb{Z}^+;\mathcal{U})$ and $L^2(\mathbb{R}^+;\mathcal{Y})$ onto $\ell^2(\mathbb{Z}^+;\mathcal{Y})$ by using the Laguerre transform $\mathcal{L}_{\overline{\alpha}}$ (restricted to positive time). These two transforms are unitary, and they map the duality pairing (3.8) onto the duality pairing (4.10). We know from Lemma 4.4 that the image of the set of all generalized stable future trajectories of $\Sigma^\dagger$ induced by the set Ω^* is equal to the set of all generalized stable future trajectories of $\mathbf{\Sigma}_\alpha^\dagger$. Consequently, by Lemmas 3.16 and 4.9, the set of all generalized stable past trajectories of Σ induced by the set Ω is equal to the set of all generalized stable past trajectories of $\mathbf{\Sigma}_\alpha$. □

THEOREM 4.11. *If Ω is connected, then for each $x_0 \in \mathcal{X}$, the past continuous time cost of x_0 is finite if and only if the past discrete time α-cost is finite for some, or equivalently, for all $\alpha \in \Omega$. Moreover, the optimal costs for all these problems are the same.*

Proof. This follows from Lemma 4.10 and the fact that the Laguerre transform is a unitary map of $\begin{bmatrix} L^2(\mathbb{R}^-,\mathcal{U}) \\ L^2(\mathbb{R}^-,\mathcal{Y}) \end{bmatrix}$ onto $\begin{bmatrix} \ell^2(\mathbb{Z}^-,\mathcal{U}) \\ \ell^2(\mathbb{Z}^-,\mathcal{Y}) \end{bmatrix}$, and hence to cost of a generalized continuous time future trajectory $\begin{bmatrix} x_0 \\ u \\ y \end{bmatrix}$ of Σ coincides with the cost of the transformed past trajectory $\begin{bmatrix} x_0 \\ \mathcal{L}_\alpha u \\ \mathcal{L}_\alpha y \end{bmatrix}$ of $\mathbf{\Sigma}_\alpha$. □

5. The Control Riccati Equation and Right Factorizations. In this section we consider the control Riccati equation satisfied by the optimal future cost quadratic form $\|\cdot\|^2_{\mathrm{fut}}$. This equation contains a parameter $\alpha \in \Omega$, where Ω is the open subset of $\rho(A) \cap \mathbb{C}^+$ that was fixed in Remark 3.1. After the formal definition (Definition 5.1), we make the connection with the discrete-time Riccati equation (Theorem 5.6) and, in the main theorem of this section, with right factorizations (Theorem 5.9). In Remark 5.11 we comment further on how the results in this section relate to known finite-dimensional results.

As in the discrete time case, considered in [23], we want to show that the minimizing trajectory can be written in feedback form. The underlying idea is as follows. In the discrete time case we introduced a "state feedback" in the following way: we take the input u to be given by $u = Kx - v$, where K is a (unknown and possibly unbounded) feedback operator, and v is a new disturbance. The minimizing control u is given by $u = Kx$, i.e., $v = 0$. (The minus sign in front of v is not significant, but it leads to a slight simplification of some later formulas.)

In order to apply the same idea in the continuous time case we reinterpret the above procedure as follows: We first create an extra output to the original equation, namely $v = Kx - u$, and then we require this output to be zero. This interpretation can be applied also in the continuous time case. To the original set of equations

$$\Sigma : \begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A\&B \\ C\&D \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \qquad t \in \mathbb{R}^+, \tag{5.1}$$

we add one more output

$$\Sigma : \begin{bmatrix} \dot{x}(t) \\ y(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} A\&B \\ C\&D \\ K\&F \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \qquad t \in \mathbb{R}^+. \tag{5.2}$$

For $v = 0$, i.e., for

$$K\&F \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} = 0, \qquad t \in \mathbb{R}^+, \tag{5.3}$$

we expect this set of equations to give us the optimal control u which minimizes the future cost if $K\&F$ is chosen appropriately. However, as in the discrete time case, this will not be true, in general, for all possible initial states x_0, but only for a certain subset of initial states. In the discrete time case this critical set of initial states are those that can be reached in finite time [23]. It turns out (Theorem 5.9) that in the continuous time case the critical set of initial states is those that can be reached by means of a classical stable past exponential trajectory.

After this digression, we now introduce the control Riccati equation.

DEFINITION 5.1. *Let* $\Sigma := \left(\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix} \right]; \mathcal{X}, \mathcal{U}, \mathcal{Y} \right)$ *be an operator node with main operator* A*, and control operator* B*, and let* $\alpha \in \rho(A) \cap \mathbb{C}^+$*. By an* α*-normalized solution of the (generalized) continuous time* control Riccati equation *induced by* $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix} \right]$ *we mean a form* q *on* $\mathcal{X}$ *with the following properties:*

(i) q *is a closed nonnegative sesquilinear symmetric form on* $\mathcal{X}$ *with domain* $\mathcal{Z}$*;*
(ii) $(\alpha - A)^{-1}\mathcal{Z} \subset \mathcal{Z}$*;*
(iii) $(\alpha - A|_{\mathcal{X}})^{-1}B\mathcal{U} \subset \mathcal{Z}$*;*
(iv) There exists an operator $[K\&F]_\alpha \colon \left[\begin{smallmatrix} \mathcal{X} \\ \mathcal{U} \end{smallmatrix} \right] \to \mathcal{U}$ *with*

$$\mathrm{dom}\,([K\&F]_\alpha) = \left\{ \left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix} \right] \in \mathrm{dom}\,(A\&B) \middle| \, x_0 \in \mathcal{Z} \text{ and } A\&B \left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix} \right] \in \mathcal{Z} \right\} \tag{5.4}$$

and a self-adjoint operator $W_\alpha \in \mathcal{B}(\mathcal{U})$ such that (here $\Re z$ denotes the real part of the complex number z)

$$\begin{aligned} 2\Re q\left[[A\&B]\begin{bmatrix}x_0\\u_0\end{bmatrix}, x_0\right] + \left\|C\&D\begin{bmatrix}x_0\\u_0\end{bmatrix}\right\|_{\mathcal{Y}}^2 + \|u_0\|_{\mathcal{U}}^2 \\ = \left\langle [K\&F]_\alpha\begin{bmatrix}x_0\\u_0\end{bmatrix}, W_\alpha[K\&F]_\alpha\begin{bmatrix}x_0\\u_0\end{bmatrix}\right\rangle_{\mathcal{U}}, \quad \left[\begin{smallmatrix}x_0\\u_0\end{smallmatrix}\right] \in \operatorname{dom}\left([K\&F]_\alpha\right), \end{aligned} \tag{5.5}$$

and

$$[K\&F]_\alpha\begin{bmatrix}(\alpha - A|_{\mathcal{X}})^{-1}B\\ 1_{\mathcal{U}}\end{bmatrix} = -1_{\mathcal{U}}. \tag{5.6}$$

Here the term $[K\&F]_\alpha\left[\begin{smallmatrix}x_0\\u_0\end{smallmatrix}\right]$ can alternatively be written in the form

$$[K\&F]_\alpha\left[\begin{smallmatrix}x_0\\u_0\end{smallmatrix}\right] = K_\alpha\big(x_0 - (\alpha - A|_{\mathcal{X}})^{-1}Bu_0\big) - u_0, \tag{5.7}$$

where

$$K_\alpha x_0 := [K\&F]_\alpha\left[\begin{smallmatrix}x_0\\0\end{smallmatrix}\right], \qquad \operatorname{dom}(K_\alpha) := \{x_0 \in \operatorname{dom}(A) \cap \mathcal{Z} |\, Ax_0 \in \mathcal{Z}\}. \tag{5.8}$$

Note that $\operatorname{dom}([K\&F]_\alpha)$ and $\operatorname{dom}(K_\alpha)$ do not depend on α, but only on $\mathcal{Z}$, $A\&B$, and $\operatorname{dom}(S) = \operatorname{dom}(A\&B)$.

Remark 5.2. *We remark that condition* (5.6) *above could be weakened, without loss of generality, to the condition*

$$\text{The operator } \mathbf{F}_\alpha := [K\&F]_\alpha\left[\begin{smallmatrix}(\alpha - A|_{\mathcal{X}})^{-1}B\\ 1_{\mathcal{U}}\end{smallmatrix}\right] \text{ is invertible in } \mathcal{B}(\mathcal{U}). \tag{5.9}$$

Indeed, if (5.9) *holds, then we get a solution of the generalized control Riccati equation in the sense of Definition 5.1 by replacing $[K\&F]_\alpha$ by $-\mathbf{F}_\alpha^{-1}[K\&F]_\alpha$ and replacing W_α by $\mathbf{F}_\alpha^* W_\alpha \mathbf{F}_\alpha$. See also Remark 5.11 below.*

As we shall see below, the continuous time control Riccati equation is essentially equivalent to the discrete time Riccati equation that one gets by applying the Cayley transform to the continuous time Riccati equation.

DEFINITION 5.3. *Let $\left[\begin{smallmatrix}\mathbf{A}_\alpha & \mathbf{B}_\alpha\\ \mathbf{C}_\alpha & \mathbf{D}_\alpha\end{smallmatrix}\right] \in \mathcal{B}\left(\left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{U}\end{smallmatrix}\right]; \left[\begin{smallmatrix}\mathcal{X}\\ \mathcal{Y}\end{smallmatrix}\right]\right)$. By a solution of the (generalized) discrete time* control Riccati equation *induced by $\left[\begin{smallmatrix}\mathbf{A}_\alpha & \mathbf{B}_\alpha\\ \mathbf{C}_\alpha & \mathbf{D}_\alpha\end{smallmatrix}\right]$ we mean a form q on $\mathcal{X}$ with the following properties:*

(i) q is a closed nonnegative sesquilinear symmetric form on $\mathcal{X}$ with domain $\mathcal{Z}$;
(ii) $\mathbf{A}_\alpha\mathcal{Z} \subset \mathcal{Z}$;
(iii) $\operatorname{img}(\mathbf{B}_\alpha) \subset \mathcal{Z}$;
(iv) There exists an operator $\mathbf{K}_\alpha\colon \mathcal{X} \to \mathcal{U}$ with $\operatorname{dom}(\mathbf{K}_\alpha) = \mathcal{Z}$ and a self-adjoint operator $W_\alpha \in \mathcal{B}(\mathcal{U})$ such that

$$\begin{aligned} q[\mathbf{A}_\alpha z_0 + \mathbf{B}_\alpha u_0, \mathbf{A}_\alpha z_0 + \mathbf{B}_\alpha u_0] + \|\mathbf{C}_\alpha z_0 + \mathbf{D}_\alpha u_0\|_{\mathcal{Y}}^2 + \|u_0\|_{\mathcal{U}}^2 \\ = q[z_0, z_0] + \langle \mathbf{K}_\alpha z_0 - u_0, W_\alpha(\mathbf{K}_\alpha z_0 - u_0)\rangle_{\mathcal{U}}, \quad z_0 \in \mathcal{Z},\ u_0 \in \mathcal{U}. \end{aligned} \tag{5.10}$$

We shall also need the following alternative version of the discrete time control Riccati equation.

LEMMA 5.4. *Equation* (5.10) *in Definition 5.3 can be replaced by the equivalent*

(Lure) version, still valid for all $z_0 \in \mathcal{Z}$ and $u_0 \in \mathcal{U}$,

$$\begin{aligned} q[\mathbf{A}_\alpha z_0, \mathbf{A}_\alpha z_0] + \|\mathbf{C}_\alpha z_0\|_{\mathcal{Y}}^2 &= q[z_0, z_0] + \langle \mathbf{K}_\alpha z_0, W_\alpha \mathbf{K}_\alpha z_0 \rangle_{\mathcal{U}}, \\ \langle u_0, W_\alpha u_0 \rangle_{\mathcal{U}} &= \langle u_0, u_0 \rangle_{\mathcal{U}} + \langle \mathbf{D}_\alpha u_0, \mathbf{D}_\alpha u_0 \rangle_{\mathcal{Y}} + q[\mathbf{B}_\alpha u_0, \mathbf{B}_\alpha u_0], \\ \langle \mathbf{K}_\alpha z_0, W_\alpha u_0 \rangle_{\mathcal{U}} &= -q[\mathbf{A}_\alpha z_0, \mathbf{B}_\alpha u_0] - \langle \mathbf{C}_\alpha z_0, \mathbf{D}_\alpha u_0 \rangle_{\mathcal{U}}. \end{aligned} \tag{5.11}$$

Moreover, W_α is always boundedly invertible.

Proof. The equivalence is proved in [23, Lemma A.2]. That W_α must have a bounded inverse follows from the self-adjointness of W_α and the fact that, according to (5.11), $W_\alpha \geq 1_{\mathcal{U}}$. □

In order to connect the continuous and discrete time Riccati equations to each other we need the following lemma.

LEMMA 5.5 ([3, Lemma 4.4]). *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be an operator node with $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, main operator A, and control operator B, and let $\mathcal{Z}$ be a subspace of $\mathcal{X}$. Let $\alpha \in \rho(A)$ and define G_α as in (2.8).*

(i) $\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]$ is invariant under G_α if and only if

$$(\alpha - A)^{-1}\mathcal{Z} \subset \mathcal{Z}, \quad (\alpha - A|_{\mathcal{X}})^{-1} B\mathcal{U} \subset \mathcal{Z}. \tag{5.12}$$

(ii) If (5.12) holds, then $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]$ belongs to the range of $G_\alpha|_{\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]}$ if and only if $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \operatorname{dom}([K\&F]_\alpha)$ defined in (5.4).

In particular, the range of $G_\alpha|_{\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]}$ does not depend on the particular $\alpha \in \rho(A)$, as long as $\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]$ is invariant under G_α.

Proof. This lemma is identical to [3, Lemma 4.4], except for the fact that there A was supposed to generate a C_0 semigroup. However, that additional assumption was not used in the proof given in [3]. □

The following theorem connects the continuous-time Riccati equation for an operator node to the discrete-time Riccati equation for its Cayley transform.

THEOREM 5.6. *Let $\Sigma := (S; \mathcal{X}, \mathcal{U}, \mathcal{Y}) = (\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right], \mathcal{X}, \mathcal{U}, \mathcal{Y})$ be an operator node with main operator A, let $\alpha \in \rho(A) \cap \mathbb{C}^+$, and let $\left[\begin{smallmatrix} \mathbf{A}_\alpha & \mathbf{B}_\alpha \\ \mathbf{C}_\alpha & \mathbf{D}_\alpha \end{smallmatrix}\right]$ be the Cayley transform of $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$. Then the following claims hold.*

(i) q is a solution of the α-normalized continuous time control Riccati equation induced by $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ if and only if q is a solution of the discrete time control Riccati equation induced by $\left[\begin{smallmatrix} \mathbf{A}_\alpha & \mathbf{B}_\alpha \\ \mathbf{C}_\alpha & \mathbf{C}_\alpha \end{smallmatrix}\right]$.

(ii) The operators $[K\&F]_\alpha$, W_α, and $\mathbf{K}_\alpha$ in parts (iv) of Definitions 5.1 and 5.3 as well as the operator K_α defined in (5.8) are uniquely determined by $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$, q, and α, and the discrete time version of W_α coincides with the continuous time version of W_α. Moreover, W_α always has a bounded inverse.

(iii) The restriction of the operator G_α to $\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]$ maps $\left[\begin{smallmatrix} \mathcal{Z} \\ \mathcal{U} \end{smallmatrix}\right]$ one-to-one onto $\operatorname{dom}([K\&F]_\alpha)$, and the operators $[K\&F]_\alpha$ and $\mathbf{K}_\alpha$ can be recovered from each other by

$$\begin{aligned} \mathbf{K}_\alpha x_0 &:= \sqrt{2\Re\alpha}\,[K\&F]_\alpha \begin{bmatrix} (\alpha - A)^{-1} x_0 \\ 0 \end{bmatrix}, \\ [K\&F]_\alpha &= \left[\begin{smallmatrix} \frac{1}{\sqrt{2\Re\alpha}}\mathbf{K}_\alpha & -1_{\mathcal{U}} \end{smallmatrix}\right] \left(\begin{bmatrix} \alpha & 0 \\ 0 & 1_{\mathcal{U}} \end{bmatrix} - \begin{bmatrix} A\&B \\ 0 \quad 0 \end{bmatrix} \right). \end{aligned} \tag{5.13}$$

Proof. We begin by proving that if q is a solution of the continuous time control Riccati equation, then q is also a solution of the discrete time Riccati equation. Thus, let us suppose that q is a solution of the continuous time control Riccati equation,

let $[K\&F]_\alpha$ and W_α be the operators in part (iv) of Definition 5.1, and define $\mathbf{K}_\alpha$ by the first equation in (5.13). By Lemma 5.5, the restriction of G_α to $\left[\begin{smallmatrix}\mathcal{Z}\\ \mathcal{U}\end{smallmatrix}\right]$ maps $\left[\begin{smallmatrix}\mathcal{Z}\\ \mathcal{U}\end{smallmatrix}\right]$ onto $\operatorname{dom}([K\&F]_\alpha)$, and this operator is injective. In particular, this implies that $\operatorname{dom}(\mathbf{K}_\alpha)=\mathcal{Z}$, as required by condition (iv) in Definition 5.3.

The two invariance conditions (ii) and (iii) in Definition 5.3 are equivalent to the corresponding invariance conditions (ii) and (iii) in Definition 5.1.

By Lemma 5.5, we can replace the parameter $\left[\begin{smallmatrix}x_0\\ u_0\end{smallmatrix}\right]$ in (5.4) by $G_\alpha\left[\begin{smallmatrix}\sqrt{2\Re\alpha}z_0\\ u_0\end{smallmatrix}\right]$, where $\left[\begin{smallmatrix}z_0\\ u_0\end{smallmatrix}\right]$ is a free parameter in $\left[\begin{smallmatrix}\mathcal{Z}\\ \mathcal{U}\end{smallmatrix}\right]$. Doing so the different terms in (5.10) can be rewritten in the following form:

$$\begin{aligned}
x_0 &= \frac{1}{\sqrt{2\Re\alpha}}\big(\mathbf{A}_\alpha z_0 + z_0 + \mathbf{B}_\alpha u_0\big),\\
A\&B\begin{bmatrix}x_0\\ u_0\end{bmatrix} &= \frac{1}{\sqrt{2\Re\alpha}}\big(\alpha\mathbf{A}_\alpha z_0 - \overline{\alpha}z_0 + \alpha\mathbf{B}_\alpha u_0\big),\\
2\Re q\left[[A\&B]\begin{bmatrix}x_0\\ u_0\end{bmatrix}, x_0\right] &= q[\mathbf{A}_\alpha z_0 + \mathbf{B}_\alpha u_0, \mathbf{A}_\alpha z_0 + \mathbf{B}_\alpha u_0] - q[z_0, z_0],\\
\begin{bmatrix}C\&D\\ [K\&F]_\alpha\end{bmatrix}\begin{bmatrix}x_0\\ u_0\end{bmatrix} &= \begin{bmatrix}\mathbf{C}_\alpha & \mathbf{D}_\alpha\\ \mathbf{K}_\alpha & -1_{\mathcal{U}}\end{bmatrix}\begin{bmatrix}z_0\\ u_0\end{bmatrix}.
\end{aligned}$$

Substituting this into (5.5) we get (5.10).

It follows from Lemma 5.4 that W_α and $\mathbf{K}_\alpha$ are determined uniquely by $\left[\begin{smallmatrix}\mathbf{A}_\alpha & \mathbf{B}_\alpha\\ \mathbf{B}_\alpha & \mathbf{C}_\alpha\end{smallmatrix}\right]$ and q, and hence by $\left[\begin{smallmatrix}A\&B\\ C\&D\end{smallmatrix}\right]$, q, and α. According to the same lemma, W_α always has a bounded inverse. By construction, $\left[\begin{smallmatrix}\mathbf{A}_\alpha & \mathbf{B}_\alpha\\ \mathbf{B}_\alpha & \mathbf{C}_\alpha\\ \mathbf{K}_\alpha & -1_{\mathcal{U}}\end{smallmatrix}\right]$ is the Cayley transform of $\left[\begin{smallmatrix}A\&B\\ C\&D\\ [K\&F]_\alpha\end{smallmatrix}\right]$. By applying the inverse Cayley transform we see that $[K\&F]_\alpha$ is determined uniquely by $\left[\begin{smallmatrix}A\&B\\ C\&D\end{smallmatrix}\right]$, q, and α.

We have now proved (ii) and one half of (i) and (iii). To prove the remaining claims we assume that q is a solution of the discrete time control Riccati equation induced by $\left[\begin{smallmatrix}\mathbf{A}_\alpha & \mathbf{B}_\alpha\\ \mathbf{C}_\alpha & \mathbf{C}_\alpha\end{smallmatrix}\right]$. Let $\mathbf{K}_\alpha$ and W_α be as in part (iv) of Definition 5.3. As above we find that the restriction of the operator G_α^{-1} to $\operatorname{dom}([K\&F]_\alpha)$ maps $\operatorname{dom}([K\&F]_\alpha)$ one-to-one onto $\left[\begin{smallmatrix}\mathcal{Z}\\ \mathcal{U}\end{smallmatrix}\right]$, so that we can define $[K\&F]_\alpha$ by the second equation in (5.13). We can then replace the free parameter $\left[\begin{smallmatrix}z_0\\ u_0\end{smallmatrix}\right]$ in (5.10) by $\left[\begin{smallmatrix}\frac{1}{\sqrt{2\Re\alpha}} & 0\\ 0 & 1_{\mathcal{U}}\end{smallmatrix}\right]G_\alpha^{-1}\left[\begin{smallmatrix}x_0\\ u_0\end{smallmatrix}\right]$ where $\left[\begin{smallmatrix}x_0\\ u_0\end{smallmatrix}\right]$ is a free parameter in $\operatorname{dom}([K\&F]_\alpha)$. All the computations that we did above to pass from (5.5) to get (5.10) are reversible, and by carrying out the same computation backwards we get from (5.10) to (5.5). □

The following conditions are important to connect the Riccati equation to the future cost minimization problem (which will be done in Theorem 5.9).

Definition 5.7.

(i) The system Σ satisfies the input finite future cost condition *at the point $\alpha\in\Omega$ if $(\alpha - A|_{\mathcal{X}})^{-1}Bu_0$ has a finite future cost for every $u_0\in\mathcal{U}$.*

(ii) The system Σ satisfies the state finite future cost condition *if every initial state in $\mathcal{X}$ has a finite future cost.*

Note that in part (i) of the above definition the vector $x_0 := (\alpha - A|_{\mathcal{X}})^{-1}Bu_0$ is the state of Σ at time zero corresponding to the input $u(t) = \mathrm{e}^{\alpha t}u_0$; cf. (3.4).

Definition 5.8. *Let $\widehat{\mathfrak{D}}$ be the transfer function of the system Σ.*

(i) $\widehat{\mathfrak{D}}$ has a right $H^\infty(\mathbb{C}^+)$ factorization valid in Ω *if there exist two functions $\mathsf{M}\in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{U}))$ and $\mathsf{N}\in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{U};\mathcal{Y}))$ such that $\mathsf{M}(\lambda)$ has a bounded inverse and $\widehat{\mathfrak{D}}(\lambda)=\mathsf{N}(\lambda)\mathsf{M}(\lambda)^{-1}$ for all $\lambda\in\Omega$.*

(ii) The factorization in (i) is normalized *if the multiplication operator*

$$\hat{u} \mapsto \begin{bmatrix} \mathsf{M} \\ \mathsf{N} \end{bmatrix} \hat{u} \colon H^2(\mathbb{C}^+;\mathcal{U}) \to \begin{bmatrix} H^2(\mathbb{C}^+;\mathcal{U}) \\ H^2(\mathbb{C}^+;\mathcal{Y}) \end{bmatrix}$$

is isometric.

(iii) The factorization in (i) is weakly coprime *if the range of the multiplication operator in (ii) is equal to the Laplace transform of the future behavior* $\mathfrak{W}^0_+$ *defined in Definition 3.2.*

Note that by [20, Section 2] the above definition of weakly coprime is equivalent to several seemingly different notions that go by that name in the literature.

The following theorem is the main result in this section.

THEOREM 5.9. *If* Ω *is connected, then the following conditions are equivalent for the system* Σ*:*

(i) Σ *satisfies the input finite future cost condition at some point* $\alpha \in \Omega$*.*

(ii) Σ *satisfies the input finite future cost condition at every point* $\alpha \in \Omega$*.*

(iii) The control Riccati equation for Σ *has an* α*-normalized nonnegative solution* q *for some* $\alpha \in \Omega$*.*

(iv) The control Riccati equation for Σ *has an* α*-normalized nonnegative solution* q *for every* $\alpha \in \Omega$*.*

(v) The transfer function $\widehat{\mathfrak{D}}$ *of* Σ *has a right* H^∞*-factorization valid in some open subset of* Ω*.*

(vi) The transfer function $\widehat{\mathfrak{D}}$ *of* Σ *has a normalized weakly coprime right* H^∞*-factorization valid in* Ω*.*

If these equivalent conditions hold, then the optimal future cost is equal to the minimal α*-normalized nonnegative solution of the continuous time control Riccati equation for all* $\alpha \in \Omega$*, and it is also the minimal nonnegative solution of the corresponding discrete time control Riccati equations for all* $\alpha \in \Omega$*. In particular, these minimal solutions do not depend on the value of* $\alpha \in \Omega$*.*

Proof. Proof of (i) $\Leftrightarrow$ *(ii):* This follows from Theorem 4.5.

Proof of (i) $\Leftrightarrow$ *(iii) and (ii)* $\Leftrightarrow$ *(iv):* This follows from Theorem 5.6 and [23, Theorem 6.3].

Proof of (i) $\Rightarrow$ *(vi):* If (i) holds, then by Theorem 4.5, the discrete time system $\left[\begin{smallmatrix} \mathbf{A}_\alpha & \mathbf{B}_\alpha \\ \mathbf{C}_\alpha & \mathbf{C}_\alpha \end{smallmatrix}\right]$ satisfies the condition which in [25] was called the finite future incremental cost condition. By [25, Corollary 2.7], the discrete time transfer function $\widehat{\mathbf{D}}_\alpha$ has a weakly right coprime H^∞ factorization over the unit disc $\mathbb{D}$ which is valid in Ω_α (this set was defined at the beginning of Section 4). When this factorization is mapped into continuous time by replacing the discrete time frequency variable z by the continuous time frequency variable λ according to the formula in in (4.4) we get a weakly right coprime H^∞ factorization of $\widehat{\mathfrak{D}}$ over the right half-plane valid in Ω. For weak coprimeness, note that the restriction of the Laguerre transform to the stable future behavior $\mathfrak{W}^0_+$ of Σ maps $\mathfrak{W}^0_+$ unitarily onto the stable future behavior $\mathfrak{W}^0_{\alpha+}$ of $\mathbf{\Sigma}_\alpha$.

Proof of (vi) $\Rightarrow$ *(v):* This implication is trivial.

Proof of (v) $\Rightarrow$ *(i):* This is essentially the same proof as the proof of the implication (i) $\Rightarrow$ (vi) carried out backwards. □

It is also possible to use the minimal solution of the control Riccati equation to compute a normalized weakly right coprime factorization of $\widehat{\mathfrak{D}}$ of the type mentioned in Theorem 5.9. These generalize the well-known finite-dimensional formulas.

THEOREM 5.10. *Suppose that* Ω *is connected, let the equivalent conditions in Theorem 5.9 hold, and let* q *be the optimal future cost sesquilinear form obtained from*

the quadratic form $\|\cdot\|^2_{\mathrm{fut}}$ (so that q is the minimal solution of the α-normalized control Riccati equation for every $\alpha \in \Omega$). For each $\alpha \in \Omega$, let $[K\&F]_\alpha$, K_α, W_α, and $\mathbf{K}_\alpha$ be the operators in part (ii) of Theorem 5.6, and for all α, $\beta \in \Omega$ define $\mathbf{K}_\alpha(\beta)$ and $\mathbf{F}_\alpha(\beta)$ by

$$\begin{bmatrix}\mathbf{K}_\alpha(\beta) & \mathbf{F}_\alpha(\beta)\end{bmatrix} = [K\&F]_\alpha \begin{bmatrix}(\beta-A)^{-1} & (\beta-A|_{\mathcal{X}})^{-1}B \\ 0 & 1_{\mathcal{U}}\end{bmatrix} \begin{bmatrix}\sqrt{2\Re\beta} & 0 \\ 0 & 1_{\mathcal{U}}\end{bmatrix}. \tag{5.14}$$

(i) For all α, $\beta \in \Omega$ the operator $\mathbf{F}_\alpha(\beta)$ has a bounded inverse, $\mathbf{F}_\alpha(\alpha) = -1_{\mathcal{U}}$, and $\mathbf{K}_\alpha(\alpha) = \mathbf{K}_\alpha$.

(ii) Fix $\alpha \in \Omega$ and define

$$\mathsf{M}_\alpha(\lambda) := -[W_\alpha^{1/2}\mathbf{F}_\alpha(\lambda)]^{-1}, \qquad \mathsf{N}_\alpha(\lambda) := \widehat{\mathfrak{D}}(\beta)\mathsf{M}_\alpha(\lambda), \qquad \lambda \in \Omega. \tag{5.15}$$

Then M_α and N_α can be extended to H^∞-functions over $\mathbb{C}^+$, and $\widehat{\mathfrak{D}} = \mathsf{N}_\alpha\mathsf{M}_\alpha^{-1}$ is a normalized weakly right coprime H^∞ factorization of $\widehat{\mathfrak{D}}$ valid in Ω.

(iii) For all α, $\beta \in \Omega$ we have

$$\begin{aligned}
[K\&F]_\beta &= -\mathbf{F}_\alpha(\beta)^{-1}[K\&F]_\alpha, \\
\mathbf{K}_\beta(\lambda) &= -\mathbf{F}_\alpha(\beta)^{-1}\mathbf{K}_\alpha(\lambda), \\
\mathbf{F}_\beta(\lambda) &= -\mathbf{F}_\alpha(\beta)^{-1}\mathbf{F}_\alpha(\lambda), \\
\mathbf{K}_\beta &= -\mathbf{F}_\alpha(\beta)^{-1}\mathbf{K}_\alpha, \\
W_\beta &= \mathbf{F}_\alpha(\beta)^* W_\alpha \mathbf{F}_\alpha(\beta).
\end{aligned} \tag{5.16}$$

In particular, the factors M_α and N_α in (ii) differ from the factors M_β and N_β only by the multiplication to the right by a unitary operator (which may depend on α and β).

Proof. Proof of (i) and (ii). The condition $\mathbf{F}_\alpha(\alpha) = -1_{\mathcal{U}}$ is equivalent to (5.6). That $\mathbf{F}_\alpha(\beta)$ is invertible and that (ii) holds follows from [25, Corollary 2.7] by mapping that result back into continuous time using the transformation $z \mapsto \lambda$ given in (4.4).

Proof of (iii). We repeat the first part of the proof of Theorem 5.6, starting from the α-normalized continuous time Riccati equation, but this time we replace the free parameter $\left[\begin{smallmatrix}x_0\\u_0\end{smallmatrix}\right]$ in (5.4) by $G_\beta\left[\begin{smallmatrix}\sqrt{2\Re\beta}z_0\\u_0\end{smallmatrix}\right]$, where $\left[\begin{smallmatrix}z_0\\u_0\end{smallmatrix}\right]$ is a free parameter in $\left[\begin{smallmatrix}\mathcal{Z}\\\mathcal{U}\end{smallmatrix}\right]$. The result remains the same, except for that α is replaced by β, and that the identity $[K\&F]_\alpha = \left[\begin{smallmatrix}\mathbf{K}_\alpha & -1_{\mathcal{U}}\end{smallmatrix}\right]\left[\begin{smallmatrix}z_0\\u_0\end{smallmatrix}\right]$ is replaced by the identity $[K\&F]_\beta = \left[\begin{smallmatrix}\mathbf{K}_\alpha(\beta) & \mathbf{F}_\alpha(\beta)\end{smallmatrix}\right]\left[\begin{smallmatrix}z_0\\u_0\end{smallmatrix}\right]$. Thus, instead of the β-normalized control Riccati equation (5.5) (with α replaced by β) we get the slightly modified equation

$$\begin{aligned}
&q[\mathbf{A}_\alpha z_0 + \mathbf{B}_\alpha u_0, \mathbf{A}_\alpha z_0 + \mathbf{B}_\alpha u_0] + \|\mathbf{C}_\alpha z_0 + \mathbf{D}_\alpha u_0\|^2_{\mathcal{Y}} + \|u_0\|^2_{\mathcal{U}} \\
&\qquad = q[z_0, z_0] + \langle \mathbf{K}_\alpha(\beta)z_0 + \mathbf{F}_\alpha(\beta)u_0, W_\alpha(\mathbf{K}_\alpha(\beta)z_0 + \mathbf{F}_\alpha(\beta)u_0)\rangle_{\mathcal{U}}, \quad z_0 \in \mathcal{Z},\ u_0 \in \mathcal{U}.
\end{aligned} \tag{5.17}$$

Comparing this equation to the equation (5.10) with α replaced by β and using Remark 5.2 and the uniqueness claim in part (ii) of Theorem 5.6 we find that $\mathbf{K}_\beta = \mathbf{F}_\alpha(\beta)^{-1}\mathbf{K}_\alpha$ and $W_\beta = \mathbf{F}_\alpha(\beta)^* W_\alpha \mathbf{F}_\alpha(\beta)$. Once this is known the proofs of the remaining claims are straightforward. ∎

The next remark indicates how our results in this section relate to known finite-dimensional results.

Remark 5.11. *In the finite-dimensional setting, if Q, K, and S satisfy* (1.4)–(1.6)*, and if we add the equation $v(t) = \begin{bmatrix}K & -1_{\mathcal{U}}\end{bmatrix}\left[\begin{smallmatrix}x(t)\\u(t)\end{smallmatrix}\right] = Kx(t) - u(t)$ to* (1.1)*,*

then the transfer functions from $-v$ to u and y are given by

$$\begin{aligned}\mathsf{M}(\lambda) &:= [K(\lambda - A - BK)^{-1}B + 1_{\mathcal{U}}],\\ \mathsf{N}(\lambda) &:= [(C + DK)(\lambda - A - BK)^{-1}B + D],\end{aligned}$$

respectively (cf. Remark 8.2 below). Thus, if we denote the transfer function from u to y by G, then this gives the right factorization $\mathsf{G}(\lambda) = \mathsf{N}(\lambda)\mathsf{M}(\lambda)^{-1}$. Note that the transfer function from u to $-v$ is $-K(\lambda - A)^{-1}B + 1_{\mathcal{U}}$, and thus $\mathsf{M}(\lambda)$ is invertible for all $\lambda \in \rho(A)$ with inverse $\mathsf{M}(\lambda)^{-1} = -K(\lambda - A)^{-1}B + 1_{\mathcal{U}}$. If the system in observable and if Q is the minimal nonnegative solution of (1.5)*, then the above factorization is right coprime. It is "almost" normalized in the sense that it suffices to multiply both factors to the right by $S^{-1/2}$ to get the normalized right coprime factorization $\mathsf{G}(\lambda) = \mathsf{N}_S(\lambda)\mathsf{M}_S(\lambda)^{-1}$, where*

$$\begin{aligned}\mathsf{M}_S(\lambda) &:= [K(\lambda - A - BK)^{-1}B + 1_{\mathcal{U}}]S^{-1/2},\\ \mathsf{N}_S(\lambda) &:= [(C + DK)(\lambda - A - BK)^{-1}B + D]S^{-1/2},\end{aligned}$$

This factorization is determined uniquely (among all normalized right coprime factorizations of G) by the fact that

$$\lim_{\lambda \to +\infty} \mathsf{M}_S(\lambda) = S^{-1/2} = (1_{\mathcal{U}} + D^*D)^{-1/2}. \tag{5.18}$$

Under the same assumptions, if we define $K\&F = \begin{bmatrix} K & -1_{\mathcal{U}} \end{bmatrix}$ then the quadratic form $q[\cdot,\cdot] = \langle \cdot, Q\cdot \rangle_{\mathcal{X}}$ satisfies the Riccati equation in Definition 5.1 with $[K\&F]_\alpha$ and W_α replaced by $K\&F$ and S, respectively, except for the normalization condition (5.6)*. Thus, as we saw in Remark 5.2, if we define*

$$\begin{aligned}\mathbf{F}_\alpha &:= K\&F \begin{bmatrix} (\alpha - A)^{-1}B \\ 1_{\mathcal{U}} \end{bmatrix} = K(\alpha - A)^{-1}B - 1_{\mathcal{U}},\\ [K\&F]_\alpha &:= -\mathbf{F}_\alpha^{-1} \begin{bmatrix} K & -1_{\mathcal{U}} \end{bmatrix} = \begin{bmatrix} -\mathbf{F}_\alpha^{-1}K & \mathbf{F}_\alpha^{-1} \end{bmatrix},\\ W_\alpha &= \mathbf{F}_\alpha^* W_\alpha \mathbf{F}_\alpha,\end{aligned}$$

then we get a solution of the Riccati equation in Definition 5.1 which also satisfies the normalization condition (5.6)*.*

In the general infinite-dimensional case it is not possible to use the normalization (5.18) *due to the fact that the limit in* (5.18) *need not exist. For the same reason one cannot expect it to always be possible to rewrite the Riccati equation* (5.5) *in the form* (1.4)–(1.6) *by reversing the steps described above. Moreover, as was first noticed in [32], even if the limit in* (5.18) *does exist, the formula for the operator S should still contain an extra correction term, namely*

$$S = 1_{\mathcal{U}} + D^*D + \lim_{\lambda \to \infty} \overline{B^*}Q(\lambda - A|_{\mathcal{X}})^{-1}B,$$

where $\overline{B^*}$ is a certain extension of B^* (in the finite-dimensional case and also in some infinite-dimensional cases the above correction term vanishes). Since we cannot always normalize M by fixing the value of M at infinity we have instead chosen to use the different "α-normalization" (5.6)*, which together with* (5.15) *results in the normalization $\mathsf{M}_\alpha(\alpha) = W_\alpha^{-1/2}$.*

6. The Filter Riccati Equation and Left Factorizations. In this section we consider the filter Riccati equation satisfied by the inverse of the optimal past cost quadatic form $\|\cdot\|^2_{\text{past}}$.

DEFINITION 6.1. *Let* $\Sigma := \left(\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]; \mathcal{X}, \mathcal{U}, \mathcal{Y}\right)$ *be an operator node, and let* $\alpha \in \rho(A) \cap \mathbb{C}^+$. *By an* α*-normalized solution of the (generalized) continuous time* filter Riccati equation *induced by* $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ *we mean a* $\overline{\alpha}$*-normalized solution of the continuous time control Riccati equation induced by the adjoint system* $\Sigma^\dagger = (\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]^*; \mathcal{X}, \mathcal{Y}, \mathcal{U})$.

DEFINITION 6.2.

(i) The system Σ *satisfies the* output coercive past cost condition *at the point* $\alpha \in \Omega$ *if there exists a constant* $M > 0$ *such that*

$$\|C(\alpha - A)^{-1}x_0\|^2_{\mathcal{Y}} \le M\big(\|u\|^2_{L^2(\mathbb{R}^-;\mathcal{U})} + \|y\|^2_{L^2(\mathbb{R}^-;\mathcal{Y})}\big) \tag{6.1}$$

for every generalized stable past trajectory $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right]$ *of* Σ.

(ii) The system Σ *satisfies the* state coercive past cost condition *at the point* $\alpha \in \Omega$ *if there exists a constant* $M > 0$ *such that*

$$\|x_0\|^2_{\mathcal{X}} \le M\big(\|u\|^2_{L^2(\mathbb{R}^-;\mathcal{U})} + \|y\|^2_{L^2(\mathbb{R}^-;\mathcal{Y})}\big) \tag{6.2}$$

for every generalized stable past trajectory $\left[\begin{smallmatrix} x_0 \\ u \\ y \end{smallmatrix}\right]$ *of* Σ.

LEMMA 6.3.

(i) Σ *satisfies the output coercive past cost condition at some point* $\alpha \in \Omega$ *if and only if the adjoint system* $\Sigma^\dagger = (\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]^*; \mathcal{X}, \mathcal{Y}, \mathcal{U})$ *satisfies the input finite future cost condition at the point* $\overline{\alpha} \in \Omega^*$.

(ii) Σ *satisfies the state coercive past cost condition at some point* $\alpha \in \Omega$ *if and only if the adjoint system* $\Sigma^\dagger$ *satisfies the state finite future cost condition at the point* $\overline{\alpha} \in \Omega^*$.

Proof. By Theorem 4.5 the future continuous time cost of an initial state x_0 is equal to its future discrete time cost, and by Theorem 4.11, the past continuous time cost of a final state x_0 is equal to its past discrete time cost. Both the claims above therefore follow from (4.1) and [24, Theorem 6.4] and [24, Lemma 6.3 and Remark 3.3]. ∎

DEFINITION 6.4. *Let* $\widehat{\mathfrak{D}}$ *be the transfer function of the system* Σ.

(i) $\widehat{\mathfrak{D}}$ *has a* left $H^\infty(\mathbb{C}^+)$ factorization valid in Ω *if there exist two functions* $\widetilde{\mathsf{M}} \in H^\infty(\mathbb{C}^+; \mathcal{B}(\mathcal{Y}))$ *and* $\widetilde{\mathsf{N}} \in H^\infty(\mathbb{C}^+; \mathcal{B}(\mathcal{U}; \mathcal{Y}))$ *such that* $\widetilde{\mathsf{M}}(\lambda)$ *has a bounded inverse and* $\widehat{\mathfrak{D}}(\lambda) = \widetilde{\mathsf{M}}(\lambda)^{-1}\widetilde{\mathsf{N}}(\lambda)$ *for all* $\lambda \in \Omega$.

(ii) The factorization in (i) is normalized *if the operator*

$$\begin{bmatrix} \hat{u} \\ \hat{y} \end{bmatrix} \mapsto P_{H^2(\mathbb{C}^-;\mathcal{Y})} \begin{bmatrix} -\widetilde{\mathsf{N}} & \widetilde{\mathsf{M}} \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{y} \end{bmatrix} : \begin{bmatrix} H^2(\mathbb{C}^-;\mathcal{U}) \\ H^2(\mathbb{C}^-;\mathcal{Y}) \end{bmatrix} \to H^2(\mathbb{C}^-;\mathcal{Y})$$

is co-isometric.

(iii) The factorization in (i) is weakly coprime *if the kernel of the operator in (ii) coincides with the (past time) Laplace transform of the past behavior* $\mathfrak{W}^0_-$.

The following is the main theorem of this section (we note that the inverse of a quadratic form is understood here in the sense of part (ii) of Lemma 3.17).

THEOREM 6.5. *If* Ω *is connected, the following conditions are equivalent for the system* Σ*:*

(i) Σ *satisfies the output coercive past cost condition at some point* $\alpha \in \Omega$.

(ii) Σ *satisfies the output coercive past cost condition at every point* $\alpha \in \Omega$.
(iii) The filter Riccati equation for Σ *has an* α*-normalized nonnegative solution for some* $\alpha \in \Omega$.
(iv) The filter Riccati equation for Σ *has a* α*-normalized nonnegative solution for every* $\alpha \in \Omega$.
(v) The transfer function $\widehat{\mathfrak{D}}$ *of* Σ *has a left* H^∞*-factorization valid in some open subset of* Ω.
(vi) The transfer function $\widehat{\mathfrak{D}}$ *of* Σ *has a weakly coprime left* H^∞*-factorization valid in* Ω.

If these equivalent conditions hold, then the optimal past cost is the inverse of the minimal α*-normalized nonnegative solution of the continuous time filter Riccati equation for all* $\alpha \in \Omega$*, and it is also the inverse of the minimal nonnegative solution of the corresponding discrete time filter Riccati equations for all* $\alpha \in \Omega$*. In particular, these minimal solutions do not depend on the value of* $\alpha \in \Omega$.

Proof. This follows from Lemmas 6.3 and 4.9(iv), Theorems 3.18 and 5.9, the fact that the transfer function $\widehat{\mathfrak{D}}^d$ of the adjoint system $\Sigma^\dagger$ is given by $\widehat{\mathfrak{D}}^d(\lambda) = \widehat{\mathfrak{D}}(\overline{\lambda})^*$, that $\widehat{\mathfrak{D}} = \mathsf{N}\mathsf{M}^{-1}$ is a right H^∞ factorization of $\widehat{\mathfrak{D}}$ if and only if $\widehat{\mathfrak{D}}^d = \widetilde{\mathsf{M}}^{-1}\widetilde{\mathsf{N}}$ is a left H^∞ factorization of $\widehat{\mathfrak{D}}^d$ where $\widetilde{\mathsf{M}}(\lambda) = \mathsf{M}(\overline{\lambda})^*$ and $\widetilde{\mathsf{N}}(\lambda) = \mathsf{N}(\overline{\lambda})^*$, and that one of these factorizations is normalized or weakly coprime if and only if the other is normalized or weakly coprime. ∎

There is also an analogue of Theorem 5.10 for the past cost minimization problem and left factorizations. We leave the formulation and proof of this result to the reader.

7. Doubly Coprime Factorizations and Past Cost Dominance. In this section we look at the case where both the input finite future cost condition and the output coercive past cost conditions hold, and the future cost is dominated by the past cost. This last notion is made precise in the following two definitions.

Definition 7.1. *Let* q *and* r *be two closed symmetric nonnegative sesquilinear forms on the Hilbert space* $\mathcal{X}$*. Then we say that* r dominates q *if* $\operatorname{dom}(r) \subset \operatorname{dom}(q)$ *and there exists a constant* $M > 0$ *such that* $q[x,x] \le Mr[x,x]$ *for all* $x \in \operatorname{dom}(r)$.

Definition 7.2. *The system* Σ *satisfies the* past cost dominance condition *(with respect to* Ω*) if the optimal future cost* $\|\cdot\|^2_{\mathrm{fut}}$ *is dominated by the optimal past cost* $\|\cdot\|^2_{\mathrm{past}}$.

Lemma 7.3. *If the system* Σ *satisfies the past cost dominance condition (with respect to* Ω*), then it satisfies both the input finite future cost condition and the output coercive past cost condition (with respect to* Ω*). Thus, in particular, the past cost dominance condition implies that both the control Riccati equation and the filter Riccati equation for* Σ *have nonnegative solutions.*

Proof. This follows from Theorems 4.5 and 4.11 and the fact that the corresponding statement is true for discrete time systems according to [25, Lemma 4.2]. ∎

Definition 7.4. *Let* $\widehat{\mathfrak{D}}$ *be the transfer function of the system* Σ.
(i) A right $H^\infty(\mathbb{C}^+)$ *factorization* $\left[\begin{smallmatrix}\mathsf{M}\\ \mathsf{N}\end{smallmatrix}\right]$ *valid in* Ω *is* strongly coprime *if there exist two functions* $\widetilde{\mathsf{X}} \in H^\infty(\mathbb{C}^+; \mathcal{B}(\mathcal{U}))$ *and* $\widetilde{\mathsf{Y}} \in H^\infty(\mathbb{C}^+; \mathcal{B}(\mathcal{Y};\mathcal{U}))$ *such that* $\widetilde{\mathsf{X}}(\lambda)\mathsf{M}(\lambda) - \widetilde{\mathsf{Y}}(\lambda)\mathsf{N}(\lambda) = 1_\mathcal{U}$ *for all* $\lambda \in \mathbb{C}^+$.
(ii) A left $H^\infty(\mathbb{C}^+)$ *factorization* $[\widetilde{\mathsf{M}}, \widetilde{\mathsf{N}}]$ *valid in* Ω *is* strongly coprime *if there exist two functions* $\mathsf{X} \in H^\infty(\mathbb{C}^+; \mathcal{B}(\mathcal{Y}))$ *and* $\mathsf{Y} \in H^\infty(\mathbb{C}^+; \mathcal{B}(\mathcal{U};\mathcal{Y}))$ *such that* $\widetilde{\mathsf{M}}(\lambda)\mathsf{X}(\lambda) - \widetilde{\mathsf{N}}(\lambda)\mathsf{Y}(\lambda) = 1_\mathcal{Y}$ *for all* $\lambda \in \mathbb{C}^+$.
(iii) $\widehat{\mathfrak{D}}$ *has a* doubly coprime $H^\infty(\mathbb{C}^+)$ factorization valid in Ω *if there exist func-*

tions $\mathsf{M} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{U}))$, $\mathsf{N} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{U};\mathcal{Y}))$, $\widetilde{\mathsf{X}} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{U}))$, $\widetilde{\mathsf{Y}} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{Y};\mathcal{U}))$, $\widetilde{\mathsf{M}} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{Y}))$, $\widetilde{\mathsf{N}} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{U};\mathcal{Y}))$, $\mathsf{X} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{Y}))$ *and* $\mathsf{Y} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{U};\mathcal{Y}))$ *such that* $\left[\begin{smallmatrix}\mathsf{M}\\ \mathsf{N}\end{smallmatrix}\right]$ *is a right* $H^\infty(\mathbb{C}^+)$ *factorization valid in* Ω, $[\widetilde{\mathsf{M}},\widetilde{\mathsf{N}}]$ *is a left* $H^\infty(\mathbb{C}^+)$ *factorization valid in* Ω *and*

$$\begin{bmatrix} \mathsf{M} & \mathsf{Y} \\ \mathsf{N} & \mathsf{X} \end{bmatrix} \begin{bmatrix} \tilde{\mathsf{X}} & -\tilde{\mathsf{Y}} \\ -\tilde{\mathsf{N}} & \tilde{\mathsf{M}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathsf{X}} & -\tilde{\mathsf{Y}} \\ -\tilde{\mathsf{N}} & \tilde{\mathsf{M}} \end{bmatrix} \begin{bmatrix} \mathsf{M} & \mathsf{Y} \\ \mathsf{N} & \mathsf{X} \end{bmatrix} = \begin{bmatrix} 1_{\mathcal{U}} & 0 \\ 0 & 1_{\mathcal{Y}} \end{bmatrix}, \quad (7.1)$$

on $\mathbb{C}^+$.

It is well-know that any strongly coprime factorization is weakly coprime in the corresponding sense (right/left) and that a transfer function has a strongly right coprime factorization if and only if it has a strongly left coprime factorization if and only if it has a doubly coprime factorization (all over $H^\infty(\mathbb{C}^+)$), see e.g. [20].

The following theorem is the final result of this article. It involves the notion of the inverse of a quadratic form as defined in part (ii) of Lemma 3.17.

THEOREM 7.5. *If* Ω *is connected, then the following conditions are equivalent for the system* Σ*:*

(i) Σ *satisfies the past cost dominance condition with respect to* Ω.

(ii) Σ *satisfies both the input finite future cost condition and the output coercive past condition at every point* $\alpha \in \Omega$*, and the optimal future cost is dominated by the optimal past cost.*

(iii) For some $\alpha \in \Omega$ *the control Riccati equation for* Σ *has an* α*-normalized nonnegative solution* q *and the filter Riccati equation for* Σ *has an* α*-normalized nonnegative solution* p *and these are such that* q *is dominated by the inverse of* p.

(iv) For all $\alpha \in \Omega$ *the control Riccati equation for* Σ *has an* α*-normalized solution* q *and the filter Riccati equation for* Σ *has an* α*-normalized solution* p *and these are such that* q *is dominated by the inverse of* p.

(v) The transfer function $\widehat{\mathfrak{D}}$ *of* Σ *has a doubly coprime* H^∞*-factorization valid in some open subset of* Ω.

(vi) The transfer function $\widehat{\mathfrak{D}}$ *of* Σ *has a doubly coprime* H^∞*-factorization valid in* Ω.

Proof. Proof of (i) $\Leftrightarrow$ *(vi):* This follows from the corresponding discrete-time result [25, Theorem 4.6], the fact that the existence of a strongly coprime factorization implies the existence of a doubly coprime factorization and the fact that all statements translate to discrete-time (as in the proof of Theorem 5.9).

Proof of (i) $\Rightarrow$ *(ii):* This follows from Lemma 7.3.

Proof of (ii) $\Rightarrow$ *(iv):* Existence of solutions follows from Theorems 5.9 and 6.5. It is the minimal solutions that we consider. The dominance then follows from the fact, again from Theorems 5.9 and 6.5, that the optimal future cost is given by the minimal solution of the control Riccati equation and the optimal past cost by the inverse of the minimal solution of the filter Riccati equation.

Proof of (iii) $\Leftrightarrow$ *(i):* This follows from the corresponding discrete-time result [25, Theorem 4.4] and the fact that all statements translate to discrete-time (as in the proof of Theorem 5.9).

Proof of (v) $\Rightarrow$ *(i):* This follows from the corresponding discrete-time result [25, Theorem 4.6] and the fact that all statements translate to discrete-time (as in the proof of Theorem 5.9).

(iv) $\Rightarrow$ (iii) and (vi) $\Rightarrow$ (v) are trivial.

□

8. Example (transfer function without doubly coprime factorization). We illustrate our theory further by considering an example. We investigate this example from two different perspectives presented in this article. In Section 8.1 we directly consider the optimal control problem to verify the input finite future cost condition and the output coercive past condition. We also obtain formulas for weakly left and right coprime factorizations and for a state space on which the state finite future cost condition is satisfied. In Section 8.2 we make an additional compactness assumption and consider the Riccati equations. This approach allows us to prove more, namely that the past cost dominance condition is not satisfied and that therefore a doubly coprime factorization does not exist. This approach also allows us to precisely characterize the spaces of finite cost states.

In Section 8.3 we consider a slight modification of the above mentioned example where the transfer function does not even have left and right factorizations.

The basic example we consider in this section is somewhat academic. In Section 8.4, we consider an example that is physically more relevant. The computations for that example become too burdensome to carry out algebraically in all detail, but we indicate how the behavior is the same as that of our basic example.

Our example is the one originally presented in [6, Section 2.2]. Given is the following second order differential equation with input u and output y:

$$\ddot{w}(t) + (-2+T)\dot{w}(t) - Tw(t) = u(t), \quad y(t) = w(t). \tag{8.1}$$

Here $T\colon \operatorname{dom}(T) \subset \mathcal{H} \to \mathcal{H}$ is a nonnegative self-adjoint unbounded operator with a bounded inverse on the infinite-dimensional Hilbert space $\mathcal{H}$ (e.g., minus the Dirichlet Laplacian on L^2 of some bounded domain). For $\alpha \geq 0$ we let $\mathcal{H}_\alpha = \operatorname{dom}(T^\alpha)$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_\alpha} = \langle T^\alpha \cdot, T^\alpha \cdot \rangle_{\mathcal{H}}$, and let $\mathcal{H}_{-\alpha}$ be the corresponding extrapolation space (i.e., the dual of $\mathcal{H}_\alpha$ with $\mathcal{H}$ as pivot space). Then T maps $\mathcal{H}_{\alpha+1}$ one-to-one onto $\mathcal{H}_\alpha$ for all $\alpha \geq 0$, and it can be extended to an operator that maps $\mathcal{H}_{\alpha+1}$ one-to-one onto $\mathcal{H}_\alpha$ for all $\alpha < 0$. We denote this extended operator by the same letter T.

The input space $\mathcal{U}$ and output space $\mathcal{Y}$ are both taken to be $\mathcal{H}$, and the state space will be a suitable subspace of $\left[\begin{smallmatrix}\mathcal{H}\\ \mathcal{H}\end{smallmatrix}\right]$. The second order system (8.1) can be written as a first order system in several different ways. As is usually done we throughout take the state to be $\left[\begin{smallmatrix}x_1\\ x_2\end{smallmatrix}\right] := \left[\begin{smallmatrix}w\\ \dot{w}\end{smallmatrix}\right]$. This gives the (formal) equation

$$\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t), \tag{8.2}$$

where

$$A \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 & 1_{\mathcal{H}} \\ T & 2_{\mathcal{H}} - T \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \tag{8.3}$$

$$B = \begin{bmatrix} 0 \\ 1_{\mathcal{H}} \end{bmatrix}, \tag{8.4}$$

$$C = \begin{bmatrix} 1_{\mathcal{H}} & 0 \end{bmatrix}. \tag{8.5}$$

There are several possible choices of state space for which the above system is described by an operator node. In all cases this operator node is the restriction of the

operator $\overline{S}$: $\left[\begin{smallmatrix}\mathcal{H}\\ \mathcal{H}\\ \mathcal{H}\end{smallmatrix}\right] \to \left[\begin{smallmatrix}\mathcal{H}\\ \mathcal{H}_{-1}\\ \mathcal{H}\end{smallmatrix}\right]$ defined by

$$\overline{S} = \begin{bmatrix} 0 & 1_{\mathcal{H}} & 0 \\ T & 2_{\mathcal{H}} - T & 1_{\mathcal{H}} \\ 1_{\mathcal{H}} & 0 & 0 \end{bmatrix}. \tag{8.6}$$

One possibility is to take the state space to be $\mathcal{X} := \left[\begin{smallmatrix}\mathcal{H}_{1/2}\\ \mathcal{H}\end{smallmatrix}\right]$, in which case we take the domain of the operator node S to be

$$\operatorname{dom}(S) = \left\{ \begin{bmatrix} z_1 \\ z_2 \\ u \end{bmatrix} \in \begin{bmatrix} \mathcal{H}_{1/2} \\ \mathcal{H}_{1/2} \\ \mathcal{U} \end{bmatrix} \middle| \, z_1 - z_2 \in \mathcal{H}_1 \right\}, \tag{8.7}$$

and the domain of the main operator A becomes

$$\operatorname{dom}(A) = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \begin{bmatrix} \mathcal{H}_{1/2} \\ \mathcal{H}_{1/2} \end{bmatrix} \middle| \, z_1 - z_2 \in \mathcal{H}_1 \right\}. \tag{8.8}$$

In this setting both the control operator B and the observation operator C are bounded. Another choice is to take the state space to be $\widetilde{\mathcal{X}} := \left[\begin{smallmatrix}\mathcal{H}\\ \mathcal{H}\end{smallmatrix}\right]$, in which case (8.7) and (8.8) are replaced by

$$\operatorname{dom}(S) = \left\{ \begin{bmatrix} z_1 \\ z_2 \\ u \end{bmatrix} \in \begin{bmatrix} \mathcal{H} \\ \mathcal{H} \\ \mathcal{U} \end{bmatrix} \middle| \, z_1 - z_2 \in \mathcal{H}_1 \right\}, \tag{8.9}$$

$$\operatorname{dom}(A) = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \begin{bmatrix} \mathcal{H} \\ \mathcal{H} \end{bmatrix} \middle| \, z_1 - z_2 \in \mathcal{H}_1 \right\}. \tag{8.10}$$

Also in this setting both B and C are bounded. A third option is to take the state space to be $\mathcal{X}_{1/2} := \left[\begin{smallmatrix}\mathcal{H}_1\\ \mathcal{H}_{1/2}\end{smallmatrix}\right]$, in which case (8.7) and (8.8) are replaced by

$$\operatorname{dom}(S) = \left\{ \begin{bmatrix} z_1 \\ z_2 \\ u \end{bmatrix} \in \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_1 \\ \mathcal{U} \end{bmatrix} \middle| \, T(z_1 - z_2) + u \in \mathcal{H}_{1/2} \right\}, \tag{8.11}$$

$$\operatorname{dom}(A) = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_1 \end{bmatrix} \middle| \, z_1 - z_2 \in \mathcal{H}_{3/2} \right\}. \tag{8.12}$$

In this setting B is unbounded but C is still bounded.

It is easy to see that in the three cases above the spectrum of A coincides with the set of points $s \in \mathbb{C}$ where the operator $s^2 + (-2+T)s - T$ does not have a bounded inverse which maps $\mathcal{H}$ into $\mathcal{H}_1$, and that the resolvent of A is given by

$$(s-A)^{-1} = \begin{bmatrix} \big(s - 2 + T\big)\widehat{\mathfrak{D}}(s) & \widehat{\mathfrak{D}}(s) \\ T\widehat{\mathfrak{D}}(s) & s\widehat{\mathfrak{D}}(s) \end{bmatrix}, \qquad s \in \rho(A), \tag{8.13}$$

where

$$\widehat{\mathfrak{D}}(s) := [s^2 + (-2+T)s - T]^{-1}, \qquad s \in \rho(A) \tag{8.14}$$

is the transfer function of the operator node. From this it follows that $1 \in \sigma(A)$, and that $\sigma(A) \subset (-\infty, 0) \cup [1, 2)$. In particular, the system is unstable (since $\sigma(A) \cap \mathbb{C}^+ \neq \emptyset$), but $\rho(A) \cap \mathbb{C}^+$ is connected, and we may take $\Omega = \rho(A) \cap \mathbb{C}^+$. Note that $2 \in \Omega$, and that

$$(2 - A)^{-1} = \begin{bmatrix} 1_{\mathcal{H}} & T^{-1} \\ 1_{\mathcal{H}} & 2T^{-1} \end{bmatrix} \tag{8.15}$$

(where the operator $1_{\mathcal{H}}$ has been restricted to the appropriate subspace of $\mathcal{H}$).

The function $\widehat{\mathfrak{D}}$ is well-posed in the sense described in the introduction. More precisely, $\widehat{\mathfrak{D}}$ is uniformly bounded in the half-plane $\mathbb{C}_2 = \{s \in \mathbb{C} \mid \Re s \geq 2\}$, which can be seen as follows. For $\Re s \geq 2$ we have $\widehat{\mathfrak{D}}(s) = (s-1)^{-1}(z(s) + T)^{-1}$, where $z(s) = s(s-2)/(s-1)$. Here $|(s-1)^{-1}| \leq 1$ for $\Re s \geq 2$, and for $\Re s \geq 2$ it can be shown by a direct computation that $\Re z(s) \geq 0$, which implies that $\|(z(s)+T)^{-1}\| \leq \|T^{-1}\| < \infty$. Thus, $\|\widehat{\mathfrak{D}}(s)\| \leq \|T^{-1}\|$ for $\Re s \geq 2$.

Remark 8.1. *We note that the second order differential equation* (8.1) *is not written out explicitly in [6], and that the first order system in [6] looks different from* (8.2)–(8.5). *However, the system in [6] is simply the first order form of* (8.1) *with state variables* $y_1 := w$ *and* $y_2 := \dot{w} - 2w$ *and* T *equal to minus the Dirichlet Laplacian on* $L^2(0, \pi)$. *Note also that there is a misprint in the (four) boundary conditions [6, (39) and (40)]; the correct boundary conditions that correspond to the domain [6, (43)] of the main operator* Q *in [6] are* $y_1(0, t) + y_2(0, t) = 0$ *and* $y_1(\pi, t) + y_2(\pi, t) = 0$.

8.1. The optimal control problem. As our first approach to this problem, we consider the optimal control problem directly. In Section 8.1.1, we first identify a stable 'target system' that we will convert our unstable system into by using (an unbounded) feedback. In Section 8.1.2 we implement this feedback on our original example and draw some conclusions. Section 8.1.3 contains formulas for left and right factorizations obtained by employing this feedback.

8.1.1. A stable second order system. Consider the abstract second order differential equation

$$\ddot{w}(t) + (\beta + T)\dot{w}(t) + \gamma T w(t) = 0. \tag{8.16}$$

(Note that the system (8.1) when considered without input and output is the special case $\beta = -2$ and $\gamma = -1$.) As in the preceding section we can rewrite this as a first order system with state variables $\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] := \left[\begin{smallmatrix} w \\ \dot{w} \end{smallmatrix}\right]$ to get

$$\dot{x}(t) = A_{\mathrm{cl}} x(t),$$

where A_{cl} is the restriction to the appropriate subspace of the operator $\overline{A}_{\mathrm{cl}} \colon \left[\begin{smallmatrix} \mathcal{H} \\ \mathcal{H} \end{smallmatrix}\right] \to \left[\begin{smallmatrix} \mathcal{H} \\ \mathcal{H}_{-1} \end{smallmatrix}\right]$ defined by

$$\overline{A}_{\mathrm{cl}} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 & 1_{\mathcal{H}} \\ -\gamma T & -\beta - T \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}. \tag{8.17}$$

It is easy to check that by the Lumer–Phillips theorem, if we take the state space to be $\mathcal{X} = \left[\begin{smallmatrix} \mathcal{H}_{1/2} \\ \mathcal{H} \end{smallmatrix}\right]$ and take $\mathrm{dom}(A_{\mathrm{cl}})$ to be

$$\mathrm{dom}(A_{\mathrm{cl}}) = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \begin{bmatrix} \mathcal{H}_{1/2} \\ \mathcal{H}_{1/2} \end{bmatrix} \middle| \, \gamma z_1 + z_2 \in \mathcal{H}_1 \right\}, \tag{8.18}$$

then A_{cl} generates an exponentially stable strongly continuous contraction semigroup on $\mathcal{X}$ provided that $\beta > 0$ and $\gamma > 0$. (Recall that the coefficients in (8.3) do *not* satisfy these conditions, and that the operator A in (8.3) is unstable.)

We can also consider the same system with state space $\mathcal{X}_{1/2} := \left[\begin{smallmatrix} \mathcal{H}_1 \\ \mathcal{H}_{1/2} \end{smallmatrix}\right]$. In that case the domain of A_{cl} becomes

$$\mathrm{dom}(A_{\mathrm{cl}}) = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_1 \end{bmatrix} \middle| \, \gamma z_1 + z_2 \in \mathcal{H}_{3/2} \right\}. \tag{8.19}$$

Again, by the Lumer–Phillips theorem, A_{cl} generates an exponentially stable strongly continuous contraction semigroup on $\mathcal{X}_{1/2}$ provided that $\beta > 0$ and $\gamma > 0$.

It is easy to see that in the two cases above the spectrum of A_{cl} coincides with the set of points $s \in \mathbb{C}$ where the operator $s^2 + (\beta + T)s + \gamma T$ does not have a bounded inverse which maps $\mathcal{H}$ into $\mathcal{H}_1$, and that the resolvent of A_{cl} is given by

$$(s - A_{\mathrm{cl}})^{-1} = \begin{bmatrix} \big(s + \beta + T\big)\mathsf{N}(s) & \mathsf{N}(s) \\ -\gamma T \mathsf{N}(s) & s\mathsf{N}(s) \end{bmatrix}, \qquad s \in \rho(A), \tag{8.20}$$

where

$$\mathsf{N}(s) := [s^2 + (\beta + T)s + \gamma T]^{-1}, \qquad s \in \rho(A). \tag{8.21}$$

8.1.2. The cost conditions. We now return to the original example (8.1)–(8.5). With the input

$$\begin{aligned} u(t) &= -(\beta + 2)\dot{w}(t) - (\gamma + 1)Tw(t) \\ &= -(\beta + 2)x_2(t) - (\gamma + 1)Tx_1(t), \end{aligned} \tag{8.22}$$

the system (8.1) becomes (8.16). (The respective first order systems also correspond.) This implies that if $x(0) = \left[\begin{smallmatrix} x_1(0) \\ x_2(0) \end{smallmatrix}\right] \in \mathcal{X} = \left[\begin{smallmatrix} \mathcal{H}_{1/2} \\ \mathcal{H} \end{smallmatrix}\right]$ and if we choose u as in (8.22) with $\beta > 0$ and $\gamma > 0$, then the solution $x = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$ of (8.2) satisfies $x \in L^2(0, \infty; \mathcal{X})$, i.e., $x_1 \in L^2(0, \infty; \mathcal{H}_{1/2})$ and $x_2 \in L^2(0, \infty; \mathcal{H})$. From this it follows that $y \in L^2(0, \infty; \mathcal{H})$, but it does not follow that $u \in L^2(0, \infty; \mathcal{H})$, only that $u \in L^2(0, \infty; \mathcal{H}_{-1})$.

If we have more smoothness so that $x(0) := \left[\begin{smallmatrix} x_1(0) \\ x_2(0) \end{smallmatrix}\right] \in \mathcal{X}_{1/2} = \left[\begin{smallmatrix} \mathcal{H}_1 \\ \mathcal{H}_{1/2} \end{smallmatrix}\right]$, then $x \in L^2(0, \infty; \mathcal{X}_{1/2})$, i.e., $x_1 \in L^2(0, \infty; \mathcal{H}_1)$ and $x_2 \in L^2(0, \infty; \mathcal{H}_{1/2})$. In particular, both $u \in L^2(0, \infty; \mathcal{H})$ and $y \in L^2(0, \infty; \mathcal{H})$. This implies that the subspace $\mathcal{X}_{1/2} = \left[\begin{smallmatrix} \mathcal{H}_1 \\ \mathcal{H}_{1/2} \end{smallmatrix}\right]$ is contained in the set of finite future cost states. We note that

$$(2 - A)^{-1}B = \begin{bmatrix} T^{-1} \\ 2T^{-1} \end{bmatrix},$$

which maps $\mathcal{H}$ into $\left[\begin{smallmatrix} \mathcal{H}_1 \\ \mathcal{H}_1 \end{smallmatrix}\right]$, so that the input finite future cost condition at the point $\alpha = 2$ is satisfied.

By Theorem 5.9, the transfer function $\widehat{\mathfrak{D}}$ (8.14) has a normalized weakly coprime right H^∞-factorization valid in $\rho(A) \cap \mathbb{C}^+$. Since the restriction of $\widehat{\mathfrak{D}}$ to $(0, \infty)$ is self-adjoint, it follows that the transfer function $\widehat{\mathfrak{D}}^d$ of the adjoint system coincides with $\widehat{\mathfrak{D}}$, and therefore $\widehat{\mathfrak{D}}$ also has a normalized weakly coprime left H^∞-factorization valid in $\rho(A) \cap \mathbb{C}$. By Theorem 6.5, this means that our example also satisfies the output coercive past condition (at every point $\alpha \in \Omega$).

We finally remark that if we take $\mathcal{X}_{1/2}$ as state space instead of $\mathcal{X}$, then we still have an operator node with the same transfer function, and in that case the state finite future cost condition is satisfied.

8.1.3. Weakly coprime factorizations of the transfer function. Formula (8.22) with $\beta > 0$ and $\gamma > 0$ suggests that we can stabilize the original system (8.3)–(8.5) with the (formal) state feedback

$$K = \begin{bmatrix} -(\gamma+1)T & -(\beta+2) \end{bmatrix}. \tag{8.23}$$

Note that this state feedback operator is unbounded if we choose the state space to be $\mathcal{X} := \begin{bmatrix} \mathcal{H}_{1/2} \\ \mathcal{H} \end{bmatrix}$ or $\widetilde{\mathcal{X}} := \begin{bmatrix} \mathcal{H} \\ \mathcal{H} \end{bmatrix}$, but it is bounded if we choose the state space to be $\mathcal{X}_{1/2} := \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_{1/2} \end{bmatrix}$. We also consider the output injection

$$L = \delta \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \tag{8.24}$$

where $\delta > 2$. This output injection operator is unbounded if we choose the state space to be $\mathcal{X} := \begin{bmatrix} \mathcal{H}_{1/2} \\ \mathcal{H} \end{bmatrix}$ or $\mathcal{X}_{1/2} := \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_{1/2} \end{bmatrix}$, but it is bounded if we choose the state space to be $\widetilde{\mathcal{X}} := \begin{bmatrix} \mathcal{H} \\ \mathcal{H} \end{bmatrix}$. Note that for all of the three considered state spaces either K or L is unbounded (or both are).

Remark 8.2. *For a finite-dimensional system with node* $\left[\begin{smallmatrix} A & B \\ C & 0 \end{smallmatrix}\right]$*, the usual formulas for a doubly coprime factorization in terms of a stabilizing state feedback* K *and a stabilizing output injection operator* L *are as follows. A right factorization* $\left[\begin{smallmatrix} \mathsf{M} \\ \mathsf{N} \end{smallmatrix}\right]$ *is obtained as the transfer function of*

$$\left[\begin{array}{c|c} A+BK & B \\ \hline K & 1_{\mathcal{U}} \\ C & 0 \end{array}\right],$$

a left factorization $\begin{bmatrix} \widetilde{\mathsf{M}} & \widetilde{\mathsf{N}} \end{bmatrix}$ *is obtained as the transfer function of*

$$\left[\begin{array}{c|cc} A-LC & -L & B \\ \hline C & 1_{\mathcal{Y}} & 0 \end{array}\right],$$

Bezout factors $\begin{bmatrix} \widetilde{\mathsf{X}} & \widetilde{\mathsf{Y}} \end{bmatrix}$ *for the right factorization are the transfer functions of*

$$\left[\begin{array}{c|cc} A-LC & -B & L \\ \hline K & 1_{\mathcal{U}} & 0 \end{array}\right],$$

and Bezout factors $\left[\begin{smallmatrix} \mathsf{X} \\ \mathsf{Y} \end{smallmatrix}\right]$ *for the left factorization are the transfer functions of*

$$\left[\begin{array}{c|c} A+BK & L \\ \hline C & 1_{\mathcal{Y}} \\ K & 0 \end{array}\right].$$

By formally applying the state feedback K in (8.23) to our unstable system (8.3)–(8.5) we get from Remark 8.2 the (formal) right factorization $\left[\begin{smallmatrix} \mathsf{M} \\ \mathsf{N} \end{smallmatrix}\right]$ with

$$\begin{aligned} \mathsf{M}(s) &= [s^2 - (2-T)s - T]\,[s^2 + (\beta+T)s + \gamma T]^{-1}, \\ \mathsf{N}(s) &= [s^2 + (\beta+T)s + \gamma T]^{-1}. \end{aligned}$$

To see that this actually is an H^∞ factorization (which is at the same time both a right and a left factorization since the two factors commute) we can argue as follows.

By the boundedness of the resolvent of A_{cl} in the state space $\mathcal{X} = \left[\begin{smallmatrix}\mathcal{H}_{1/2}\\ \mathcal{H}\end{smallmatrix}\right]$ and the explicit formula (8.20) for this resolvent, $\mathsf{N} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{H};\mathcal{H}_{1/2}))$, $s \mapsto s\mathsf{N}(s) \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{H}))$, and $s \mapsto (s+\beta+T)\mathsf{N}(s) \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{H}_{1/2}))$. Since $\mathcal{H}_{1/2}$ is continuously embedded in $\mathcal{H}$ and all the involved operators commute with $T^{1/2}$, it follows that $\mathsf{N} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{H}))$ and $s \mapsto (s+\beta+T)\mathsf{N}(s) \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{H}))$, which implies that furthermore $T\mathsf{N} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{H}))$. Since

$$\mathsf{M}(s) = 1_{\mathcal{H}} - ((\beta+2)s + (\gamma+1)T)\mathsf{N}(s), \qquad s \in \mathbb{C}^+,$$

it follows that also $\mathsf{M} \in H^\infty(\mathbb{C}^+;\mathcal{B}(\mathcal{H}))$.

If we instead apply the output injection L in (8.24) we get the (formal) left factorization $\begin{bmatrix}\widetilde{\mathsf{M}} & \widetilde{\mathsf{N}}\end{bmatrix}$ with

$$\begin{aligned}\widetilde{\mathsf{M}}(s) &= [s^2 - (2-T)s - T]\ [s^2 + (T+\delta-2)s + (\delta-1)T]^{-1},\\ \widetilde{\mathsf{N}}(s) &= [s^2 + (\delta-2+T)s + (\delta-1)T]^{-1}.\end{aligned}$$

This is the same factorization as we obtained above, with $\beta = \delta - 2 > 0$ and $\gamma = \delta - 1 > 0$ (and hence it is a H^∞ factorization).

However, the Bezout factors obtained from Remark 8.2 will not be bounded. For example, for the Bezout factor $\widetilde{\mathsf{Y}}$ we formally obtain

$$\begin{aligned}\widetilde{\mathsf{Y}}(s) = -\delta\left[((\gamma+1)T + (2\beta+4))\,s + (\gamma+1)T^2 + (\beta+2)T\right]\\ \times [s^2 + (T+(\delta-2))\,s + (\delta-1)T]^{-1}.\end{aligned}$$

Due to the presence of the term T^2, it is clear that $\widetilde{\mathsf{Y}}(s)$ is not a bounded operator for any s in the open right half-plane. So the obvious candidate for a Bezout factor is in fact not a Bezout factor. Similarly it can be seen that the formal equation for $\mathsf{Y}(s)$ gives an unbounded operator for all s in the open right half-plane.

8.2. The Riccati equations. For consideration of the Riccati equations we follow [6] and use the state space $\widetilde{\mathcal{X}} := \left[\begin{smallmatrix}\mathcal{H}\\ \mathcal{H}\end{smallmatrix}\right]$. (This is not important, but it leads to simple computations.)

We will make the additional assumption on T that it has a compact inverse. Then there exists an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors $\{\varphi_k : k \in \mathbb{N}\}$ of T. Denote the corresponding eigenvalues by $\{\lambda_k : k \in \mathbb{N}\}$ and note that $\lambda_k \to \infty$ as $k \to \infty$. The space $\widetilde{\mathcal{X}}$ has an orthonormal basis of eigenvectors $\{\left[\begin{smallmatrix}\varphi_k\\ 0\end{smallmatrix}\right], \left[\begin{smallmatrix}0\\ \varphi_k\end{smallmatrix}\right] : k \in \mathbb{N}\}$. With respect to that basis, the operators A, B and C are block-diagonal (with the size of the blocks equal to two). It follows that the sesquilinear forms giving the optimal cost and the optimal feedback pairs are block diagonal as well.

An elementary calculation using the α-normalized control Riccati equation then gives that for these diagonal blocks we have

$$Q_k = \begin{bmatrix}Q_{k,1} & Q_{k,0}\\ Q_{k,0} & Q_{k,2}\end{bmatrix}, \quad [K\&F]_{\alpha,k} = [K_{k,1}\ K_{k,2}\ F_k], \quad W_\alpha = W_k,$$

with

$$\begin{aligned}
Q_{k,0} &= \sqrt{\lambda_k^2+1}+\lambda_k,\\
Q_{k,1} &= \sqrt{2\sqrt{\lambda_k^2+1}+\lambda_k^2-2\lambda_k+4}-\lambda_k+2,\\
Q_{k,2} &= \sqrt{\lambda_k^2+1}\sqrt{2\sqrt{\lambda_k^2+1}+\lambda_k^2-2\lambda_k+4}+\lambda_k^2-2\lambda_k,\\
F_k &= -\left[1+\frac{Q_{k,0}+\alpha Q_{k,2}}{\alpha^2-2+\lambda_k(\alpha-1)}\right]^{-1},\\
K_{k,1} &= FQ_{k,0},\\
K_{k,2} &= FQ_{k,2},\\
W_k &= F^{-2},
\end{aligned}$$

here the optimal cost sesquilinear form q is defined on a dense subset of its domain through $q[x,z]=\langle Qx,z\rangle$ for $x,z\in\operatorname{dom}(Q)$.

The asymptotic behavior of the above terms can be seen to be

$$\begin{aligned}
Q_k &= \lambda_k^2\begin{bmatrix}2&0\\0&0\end{bmatrix}+\lambda_k\begin{bmatrix}-2&2\\2&0\end{bmatrix}+\begin{bmatrix}\frac{5}{2}&0\\0&2\end{bmatrix}+O\left(\frac{1}{\lambda_k}\right),\\
W_k &= \left(\frac{1+\alpha}{1-\alpha}\right)^2+O\left(\frac{1}{\lambda_k}\right),\\
F_k &= \frac{1-\alpha}{1+\alpha}+O\left(\frac{1}{\lambda_k}\right).
\end{aligned}$$

It follows that $W_\alpha\in\mathcal{B}(\mathcal{U})$ and that

$$Q_k^{1/2}=\sqrt{2}\left(\lambda_k\begin{bmatrix}1&0\\0&0\end{bmatrix}+\begin{bmatrix}-1/2&1\\1&0\end{bmatrix}\right)+O\left(\frac{1}{\lambda_k}\right),$$

from which we can conclude that $\operatorname{dom}(q)=\operatorname{dom}(Q^{1/2})=\left[\begin{smallmatrix}\mathcal{H}_1\\\mathcal{H}\end{smallmatrix}\right]$. So the set of finite future cost states Ξ_f equals $\left[\begin{smallmatrix}\mathcal{H}_1\\\mathcal{H}\end{smallmatrix}\right]$.

For the solution of the dual optimal control problem we utilize the filter Riccati equation. A similar calculation shows that the minimal nonnegative solution corresponds to the block diagonal operator with blocks that have the asymptotic behavior

$$P_k=\begin{bmatrix}2&2\\2&2\end{bmatrix}+O\left(\frac{1}{\lambda_k}\right).$$

As Theorem 6.5 indicates we are however interested in the inverse of this operator. The diagonal blocks of this have the asymptotic expansion

$$P_k^{-1}=\lambda_k\begin{bmatrix}2&-2\\-2&2\end{bmatrix}+\begin{bmatrix}5/2&0\\0&-2\end{bmatrix}+O\left(\frac{1}{\lambda_k}\right).$$

It follows that

$$P_k^{-1/2}=\sqrt{\lambda_k}\begin{bmatrix}1&-1\\-1&1\end{bmatrix}+O\left(\frac{1}{\sqrt{\lambda_k}}\right),$$

and therefore we have for the set Ξ_p of finite past cost states

$$\Xi_p = \operatorname{dom}(p^{-1}) = \operatorname{dom}(P^{-1/2}) = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \begin{bmatrix} \mathcal{H} \\ \mathcal{H} \end{bmatrix} \middle| -z_1 + z_2 \in \mathcal{H}_{1/2} \right\}.$$

Note that since $\Xi_p \not\subseteq \Xi_f$, the past cost dominance condition is not satisfied. This can also be seen in the following ways. Firstly, from the asymptotic expansions of Q and P we have for the diagonal blocks of the product QP:

$$Q_k P_k = \lambda_k^2 \begin{bmatrix} 4 & 4 \\ 0 & 0 \end{bmatrix} + O(\lambda_k).$$

It follows that QP is unbounded and that therefore the past cost dominance condition is not satisfied. Secondly we can see from the expressions for Q_k and P_k^{-1} that Q is not dominated by a constant times P^{-1}. From this we can also conclude that the past cost dominance condition is not satisfied.

Since the past cost dominance condition is not satisfied, by Theorem 7.5, the transfer function $\widehat{\mathfrak{D}}$ in (8.14) does not have a doubly coprime factorization.

In the introduction we listed the additional conclusions (i)–(iii) which are valid also in this particular example since $\widehat{\mathfrak{D}}$ is well-posed. If we choose the state space to be the space $\Xi_f = \left[\begin{smallmatrix} \mathcal{H}_1 \\ \mathcal{H} \end{smallmatrix} \right]$ of all vectors with finite future cost, then with this state space both the original system and the optimal state feedback system are well-posed. This is the weakest possible norm for which the system satisfies the state finite future cost condition. We may also choose the state space to be the space of Ξ_p of finite past cost states described above. Also with this norm both the original system and the optimal output injection system are well-posed. This is the strongest possible norm for which the system satisfies the state coercive past cost condition.

8.3. A slightly different example. With the same assumptions on T as in Section 8.2, we now consider

$$\ddot{w}(t) + (-2+T)\dot{w}(t) - Tw(t) = hu(t), \quad y(t) = \langle w(t), h \rangle, \tag{8.25}$$

where $h \in \mathcal{H}$ is nonzero. The transfer function then is

$$\widehat{\mathfrak{D}}(s) = \langle [s^2 + (-2+T)s - T]^{-1} h, h \rangle.$$

Define

$$s_k := \frac{-\lambda_k + 2 + \sqrt{\lambda_k^2 + 4}}{2}.$$

It is easily seen that

$$(s_k^2 + (-2+T)s_k - T)\varphi_k = 0,$$

and that therefore s_k is a pole of $\widehat{\mathfrak{D}}$. We have $\lim_{k\to\infty} s_k = 1$, so that the transfer function $\widehat{\mathfrak{D}}$ has a non-isolated singularity in the open right-half plane. It follows that $\widehat{\mathfrak{D}}$ is not meromorphic in the right half-plane and in particular that it neither has a right- nor a left-factorization. We conclude that there exists no realization of $\widehat{\mathfrak{D}}$ for which the input finite future cost condition holds nor a realization of $\widehat{\mathfrak{D}}$ for which the output coercive past cost condition holds. In particular, it is impossible to choose a first order representation and a state space for (8.25) that makes the first order representation into an operator node that satisfies the input finite future cost condition (or output coercive past cost condition). This implies that is is impossible that an operator node first order representation exists for which the state finite future cost condition holds.

8.4. A more physical example. We next consider a physically better motivated example. The basic example from the start of Section 8 can be seen as a simplification of this example.

Consider the plant

$$\ddot{w}_p(t) + T_p w_p(t) = u_p(t), \quad y_p(t) = w_p(t).$$

Here $T_p : \mathrm{dom}(T_p) \subset \mathcal{H} \to \mathcal{H}$ is a nonnegative self-adjoint operator with a compact inverse on the infinite-dimensional Hilbert space $\mathcal{H}$ (e.g. minus the Dirichlet Laplacian on L^2 of some bounded domain). Denote the eigenvalues of T by λ_k. The input space $\mathcal{U}$ and output space $\mathcal{Y}$ are both taken to be $\mathcal{H}$. This is an undamped flexible system with force control and position measurement.

Consider a controller of the same form but with damping

$$\ddot{w}_c(t) + D_c \dot{w}_c(t) + T_c w_c(t) = u_c(t), \quad y_c(t) = w_c(t).$$

Interconnect the systems through positive feedback:

$$u_p = y_c + v, \quad u_c = y_p,$$

where v is an additional control to the plant. This is what the theory of negative imaginary systems [26] suggests to do. That theory ensures (at least in the finite dimensional case) stability of the feedback interconnection if the stiffness of the controller T_c is large enough (plus some more minor conditions).

Assume that the parameters of the controller are

$$D_c = 1_{\mathcal{H}}, \quad T_c = \frac{1}{2} T_p^{-1}.$$

This controller stiffness operator violates the condition from negative imaginary theory. The closed loop system with as output the output of the plant and with $w := \left[\begin{smallmatrix} w_p \\ w_c \end{smallmatrix}\right]$ is

$$\ddot{w}(t) + D\dot{w}(t) + Tw(t) = \begin{bmatrix} 1_{\mathcal{H}} \\ 0 \end{bmatrix} v(t), \quad y(t) = \begin{bmatrix} 1_{\mathcal{H}} & 0 \end{bmatrix} w(t),$$

where

$$D = \begin{bmatrix} 0 & 0 \\ 0 & D_c \end{bmatrix}, \quad T = \begin{bmatrix} T_p & -1_{\mathcal{H}} \\ -1_{\mathcal{H}} & T_c \end{bmatrix}.$$

The operators in the first order form with state $\left[\begin{smallmatrix} w \\ \dot{w} \end{smallmatrix}\right]$ are

$$A = \begin{bmatrix} 0 & 1_{\mathcal{H}} \\ -T & -D \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \begin{bmatrix} 1_{\mathcal{H}} \\ 0 \end{bmatrix} \end{bmatrix},$$
$$C = \begin{bmatrix} \begin{bmatrix} 1_{\mathcal{H}} & 0 \end{bmatrix} & 0 \end{bmatrix}.$$

The solutions Q and P of the control and filter Riccati equation on $\mathcal{H}^4$ are now (up to re-ordering) block diagonal with block size 4. Numerical computations indicate that for these blocks we have $\|Q_k\| \approx 2\lambda_k$, $\|P_k\| \approx \lambda_k$ and $\|Q_k P_k\| \approx \lambda_k^2$. It follows that QP is unbounded. We conclude that the transfer function of the closed-loop system has a left and a right factorization (since both Riccati equations have solutions), but no doubly coprime factorization (since the product of the solutions of the Riccati equations is unbounded).

9. Example (significance of the subset Ω). The following example illustrates what happens if one replaces the specifically chosen component Ω of $\rho(A)\cap\mathbb{C}^+$ by some other component. In particular, we show that the normalized coprime factorizations obtained by choosing different components can be genuinely different (i.e. not related by a constant unitary transformation), and that also the optimal costs can be different.

Consider the following function

$$\widehat{\mathfrak{D}}(s) := \begin{cases} \frac{1}{s-2} & |s|<1, \\ 0 & |s|>1. \end{cases}$$

An operator node with this transfer function can be constructed as follows. Let $\mathcal{X}=\ell^2(\mathbb{Z})$ and define the bounded operator A on $\mathcal{X}$ by

$$(Az)_k = z_{k+1}.$$

We note that the spectrum of A equals the unit circle, so that the set $\sigma(A)\cap\mathbb{C}^+$ consists of two components:

$$\Omega_B := \{s\in\mathbb{C}^+ : |s|<1\}, \quad \Omega_U := \{s\in\mathbb{C}^+ : |s|>1\}.$$

Further define $B\in\mathcal{B}(\mathbb{C},\mathcal{X})$ by $(Bv)_k = b_k v$, where $b\in\mathcal{X}$ is defined by

$$b_k = \begin{cases} 2^k & k<0, \\ 0 & k\geq 0, \end{cases}$$

and define $C\in\mathcal{B}(\mathcal{X},\mathbb{C})$ by

$$Cz = z_0.$$

It follows that for $|s|>1$

$$\big((s-A)^{-1}Bv\big)_k = \begin{cases} \sum_{j=k}^{-1} s^j 2^{k-j-1} & k<0, \\ 0 & k\geq 0, \end{cases}$$

and that for $|s|<1$

$$\big((s-A)^{-1}Bv\big)_k = \begin{cases} \dfrac{2^k}{s-2} & k<0, \\ \dfrac{s^k}{s-2} & k\geq 0. \end{cases}$$

We conclude that for $|s|>1$

$$C(s-A)^{-1}B = 0,$$

and that for $|s|<1$

$$C(s-A)^{-1}B = \frac{1}{s-2},$$

so that the given operator node indeed has the given function as its transfer function.

Normalized strongly coprime factorizations (since we deal with a scalar function, there is no difference between left and right) for Ω_B and Ω_U can be easily computed. If we pick Ω_B, then we obtain

$$\mathsf{M}_B(s) = \frac{s-2}{s+\sqrt{5}}, \qquad \mathsf{N}_B(s) = \frac{1}{s+\sqrt{5}},$$
$$\widetilde{\mathsf{X}}_B(s) = \frac{s+2+2\sqrt{5}}{s+\sqrt{5}}, \qquad \widetilde{\mathsf{Y}}_B(s) = \frac{-(2+\sqrt{5})^2}{s+\sqrt{5}}.$$

However, if we pick Ω_U, then we obtain

$$\mathsf{M}_U(s) = 1, \quad \mathsf{N}_U(s) = 0, \quad \widetilde{\mathsf{X}}_U(s) = 1, \quad \widetilde{\mathsf{Y}}_U(s) = 0,$$

which is clearly genuinely different.

We note that for $|s| > 1$ we have

$$C(s-A)^{-1}z = \sum_{j=0}^{\infty} s^{-j-1} z_j,$$

and that for $|s| < 1$ we have

$$C(s-A)^{-1}z = \sum_{j=0}^{\infty} -s^{j} z_{-j-1}.$$

We see that if we pick Ω_U, the output y for initial condition $z \in \mathcal{X}$ and input u is given by

$$y(t) = \sum_{j=0}^{\infty} t^j z_j.$$

Therefore the optimal future control equals zero, the space of finite future cost states is $\ell^2(\mathbb{Z}^-)$ and the optimal cost on this space equals zero. It is easily checked that the control Riccati equation is satisfied for $\mathcal{Z} := \ell^2(\mathbb{Z}^-)$ with $q := 0$, $K\&F\left[\begin{smallmatrix} x_0 \\ u_0 \end{smallmatrix}\right] := -u_0$, $W_\alpha := 1$ and $\mathrm{dom}(K\&F) = \mathcal{Z}$.

For Ω_B we consider the control Riccati equation. It is easily seen that the above mentioned solution for Ω_U is now not a solution: the conditions ii) and iii) from Definition 5.1 are not satisfied. To obtain a solution for the choice Ω_B, we choose $\mathcal{Z} := \ell^2(\mathbb{Z}^+) \oplus \mathrm{span}\{b\}$. It can then be computed that, with $\mathrm{dom}(K\&F)$ as defined in Definition 5.1,

$$\mathrm{dom}(K\&F) = \left\{ \begin{bmatrix} z \\ v \end{bmatrix} \in \begin{bmatrix} \mathcal{Z} \\ \mathbb{C} \end{bmatrix} : z_{-1} = 2^{-1} z_0 \right\}.$$

With respect to the given orthogonal decomposition of $\mathcal{Z}$, define

$$q\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right) := (2+\sqrt{5}) y_1 \overline{y}_2,$$

and define $K\&F$ on its domain (given above) by

$$K\&F \begin{bmatrix} z \\ v \end{bmatrix} := (2+\sqrt{5}) z_0 + v.$$

Using that for $\left[\begin{smallmatrix} z \\ v \end{smallmatrix}\right] \in \operatorname{dom}(K\&F)$ the projection of $A\&B\left[\begin{smallmatrix} z \\ v \end{smallmatrix}\right]$ onto $\operatorname{span}\{b\}$ equals $2z_0 + v$, that

$$C\&D \begin{bmatrix} z \\ v \end{bmatrix} = z_0,$$

and the equality

$$(2+\sqrt{5})(2z_0+v)\overline{z_0} + (2+\sqrt{5})(2\overline{z_0}+\overline{v})z_0 + |z_0|^2 + |v|^2 = |(2+\sqrt{5})z_0 + v|^2,$$

it is seen that the control Riccati equation indeed holds with $W_\alpha := 1$.

We remark that there is noting special about the functions zero and $\frac{1}{s-2}$ which appear in the example given in this section: by an appropriate modification of B, other functions can be obtained. In particular, it is possible to construct a transfer function such that if we pick Ω_U, then the function has a coprime factorization and when we pick Ω_B, it doesn't.

Appendix A. Proof of Lemma 3.17.

Let us denote the closed linear relation $\mathcal{X} \to \mathcal{W}$ whose graph is V by T. Then

$$\begin{aligned} Tx &= \{w \in \mathcal{W} \,|\, \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in V\}, \qquad x \in \operatorname{dom}(T) = \mathcal{X}_V, \\ \ker(T) &= \{x \in \mathcal{X} \,|\, \left[\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right] \in V\}, \qquad \operatorname{mul}(T) = \{w \in \mathcal{W} \,|\, \left[\begin{smallmatrix} 0 \\ w \end{smallmatrix}\right] \in V\}, \\ \operatorname{img}(T) &:= \{w \in \mathcal{W} \,|\, \left[\begin{smallmatrix} x \\ w \end{smallmatrix}\right] \in V \text{ for some } x \in \mathcal{X}\}, \end{aligned} \tag{A.1}$$

where $\ker(T)$ and $\operatorname{mul}(T)$ are closed subspaces of $\mathcal{X}$ and $\mathcal{W}$, respectively. The subspace $V^\perp$ is the graph of the relation $-(T^{-1})^* = -(T^*)^{-1}$. We denote $T^{-*} := (T^{-1})^* = (T^*)^{-1}$. Then

$$\begin{aligned} T^{-*}x^\dagger &= \left\{w^\dagger \in \mathcal{W} \,\middle|\, \left[\begin{smallmatrix} x^\dagger \\ -w^\dagger \end{smallmatrix}\right] \in V^\perp\right\}, \qquad x^\dagger \in \operatorname{dom}\left(T^{-*}\right) = \mathcal{X}_{V^\perp}, \\ \ker\left(T^{-*}\right) &= \left\{x^\dagger \in \mathcal{X} \,\middle|\, \left[\begin{smallmatrix} x^\dagger \\ 0 \end{smallmatrix}\right] \in V^\perp\right\}, \qquad \operatorname{mul}\left(T^{-*}\right) = \left\{w^\dagger \in \mathcal{W} \,\middle|\, \left[\begin{smallmatrix} 0 \\ w^\dagger \end{smallmatrix}\right] \in V^\perp\right\}, \\ \operatorname{img}\left(T^{-*}\right) &:= \left\{w^\dagger \in \mathcal{W} \,\middle|\, \left[\begin{smallmatrix} x^\dagger \\ -w^\dagger \end{smallmatrix}\right] \in V^\perp \text{ for some } x^\dagger \in \mathcal{X}\right\}, \end{aligned} \tag{A.2}$$

where $\ker(T^{-*})$ and $\operatorname{mul}(T^{-*})$ are closed subspaces of $\mathcal{X}$ and $\mathcal{W}$, respectively. Moreover, by standard properties of closed relations,

$$\begin{aligned} \operatorname{dom}(T)^\perp &= \operatorname{mul}(T^*) = \ker\left(T^{-*}\right), \\ \operatorname{img}(T)^\perp &= \ker(T^*) = \operatorname{mul}\left(T^{-*}\right), \\ \operatorname{dom}\left(T^{-*}\right)^\perp &= \operatorname{mul}\left(T^{-1}\right) = \ker(T), \\ \operatorname{img}\left(T^{-*}\right)^\perp &= \ker\left(T^{-1}\right) = \operatorname{mul}(T). \end{aligned}$$

Proof of (i). In terms of the relation T we have $\|x\|_V = \inf\{\|w\| \mid w \in Tx\}$. It follows that with $T_o := P_{\operatorname{mul}(T)^\perp} T$ the operator part of T, $\|x\|_V = \|T_o x\|_{\mathcal{W}}$. This implies that $\|\cdot\|_V^2$ is a closed quadratic form on $\mathcal{X}$ with domain $\mathcal{X}_V = \operatorname{dom}(T_o) = \operatorname{dom}(T)$. An analogous argument shows that $\|x^\dagger\|_{V^\perp}^2 = \|T_o^{-*} x^\dagger\|_{\mathcal{W}}^2$, and that $\|\cdot\|_{V^\perp}^2$ is a closed quadratic form on $\mathcal{X}$ with domain $\mathcal{X}_{V^\perp} = \operatorname{dom}(T^{-*}) = \operatorname{dom}(T_o^{-*})$, where $T_o^{-*} := P_{\operatorname{mul}(T^{-*})^\perp} T^{-*}$ is the operator part of T^{-*}.

Proof of (ii). In the proof of (ii) we need to further factor out $\ker(T)$ from $\operatorname{dom}(T)$ in order to make the operator $T_i := T_o|_{\operatorname{dom}(T) \cap \ker(T)^\perp}$ injective. As explained

in, e.g., [25, Appendix A], T_i is a closed injective densely defined operator with dense range when regarded as an operator from $\overline{\operatorname{dom}(T)} \cap \ker(T)^\perp = \overline{\operatorname{dom}(T)} \cap \overline{\operatorname{dom}(T^{-*})}$ to $\overline{\operatorname{img}(T)} \cap \operatorname{mul}(T)^\perp = \overline{\operatorname{img}(T)} \cap \overline{\operatorname{img}(T^{-*})}$. Furthermore, if we decompose the spaces $\mathcal{X}$ and $\mathcal{W}$ as

$$\begin{aligned} \mathcal{X} &= \ker(T) \oplus \left(\overline{\operatorname{dom}(T)} \cap \ker(T)^\perp\right) \oplus \operatorname{dom}(T)^\perp, \\ &= \operatorname{dom}\left(T^{-*}\right)^\perp \oplus \left(\ker\left(T^{-*}\right)^\perp \cap \overline{\operatorname{dom}\left(T^{-*}\right)}\right) \oplus \ker\left(T^{-*}\right) \\ \mathcal{W} &= \operatorname{img}(T)^\perp \oplus \left(\operatorname{mul}(T)^\perp \cap \overline{\operatorname{img}(T)}\right) \oplus \operatorname{mul}(T) \\ &= \operatorname{mul}\left(T^{-*}\right) \oplus \left(\overline{\operatorname{img}\left(T^{-*}\right)} \cap \operatorname{mul}\left(T^{-*}\right)^\perp\right) \oplus \operatorname{img}\left(T^{-*}\right)^\perp, \end{aligned} \tag{A.3}$$

then with respect to these decompositions (in this order) the (graphs of the) multi-valued operators T and T^{-1} and their adjoints are given by

$$\begin{aligned} T &= \left\{ \left(\begin{bmatrix} x_0 \\ x_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ T_i x_1 \\ w_0 \end{bmatrix} \right) \middle| \begin{bmatrix} x_0 \\ x_1 \\ w_0 \end{bmatrix} \in \begin{bmatrix} \ker(T) \\ \operatorname{dom}(T_i) \\ \operatorname{mul}(T) \end{bmatrix} \right\}, \\ T^{-1} &= \left\{ \left(\begin{bmatrix} 0 \\ w_1 \\ w_0 \end{bmatrix}, \begin{bmatrix} x_0 \\ T_i^{-1} w_1 \\ 0 \end{bmatrix} \right) \middle| \begin{bmatrix} x_0 \\ w_1 \\ w_0 \end{bmatrix} \in \begin{bmatrix} \ker(T) \\ \operatorname{img}(T_i) \\ \operatorname{mul}(T) \end{bmatrix} \right\}, \\ T^* &= \left\{ \left(\begin{bmatrix} w_0 \\ w_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ T_i^* w_1 \\ x_0 \end{bmatrix} \right) \middle| \begin{bmatrix} w_0 \\ w_1 \\ x_0 \end{bmatrix} \in \begin{bmatrix} \operatorname{mul}\left(T^{-*}\right) \\ \operatorname{img}\left(T_i^{-*}\right) \\ \ker\left(T^{-*}\right) \end{bmatrix} \right\}, \\ T^{-*} &= \left\{ \left(\begin{bmatrix} 0 \\ x_1 \\ x_0 \end{bmatrix}, \begin{bmatrix} w_0 \\ T_i^{-*} x_1 \\ 0 \end{bmatrix} \right) \middle| \begin{bmatrix} w_0 \\ x_1 \\ x_0 \end{bmatrix} \in \begin{bmatrix} \operatorname{mul}\left(T^{-*}\right) \\ \operatorname{dom}\left(T_i^{-*}\right) \\ \ker\left(T^{-*}\right) \end{bmatrix} \right\}. \end{aligned} \tag{A.4}$$

From here we can, among others, obtain the corresponding decompositions of the operator parts of T and T^{-*} by restricting them to the closures of their domains and dropping the multi-valued part of the range space, and they are given by

$$\begin{aligned} T_o &= \begin{bmatrix} 0 & 0 \\ 0 & T_i \end{bmatrix}, & \operatorname{dom}(T_o) &= \begin{bmatrix} \ker(T) \\ \operatorname{dom}(T_i) \end{bmatrix}, \\ T_o^{-*} &= \begin{bmatrix} 0 & 0 \\ 0 & T_i^{-*} \end{bmatrix}, & \operatorname{dom}\left(T_o^{-*}\right) &= \begin{bmatrix} \ker\left(T^{-*}\right) \\ \operatorname{dom}\left(T_i^{-*}\right) \end{bmatrix}, \end{aligned} \tag{A.5}$$

where $\operatorname{dom}(T_i) = \operatorname{dom}(T) \cap \ker(T)^\perp$ and $\operatorname{dom}\left(T_i^{-*}\right) = \operatorname{dom}\left(T^{-*}\right) \cap \ker\left(T^{-*}\right)^\perp$ have the same closure, equal to the middle component in the decomposition of $\mathcal{X}$ given in (A.3). Here the injective operator part T_i^{-*} of T^{-*} is equal to $(T_i^*)^{-1} = (T_i^{-1})^*$. See [25, Appendix A] for details.

After this digression we now return to the proof of Part (ii) of Lemma 3.17. As we saw in the proof of part (i), $\|x_0\|_V^2 = \|T_o x_0\|_{\mathcal{X}}^2$ and $\|x_0^\dagger\|_{V^\perp}^2 = \|T_o^{-*} x_0^\dagger\|_{\mathcal{X}}^2$. If we interpret T_o as a densely defined operator on $\overline{\operatorname{dom}(T_o)}$, then T_o^* is a (single-valued) operator, and $\|x_0\|_V^2 = \langle x_0, T_o^* T_o x_0 \rangle_{\mathcal{X}}$ for all $x_0 \in \operatorname{dom}(T_o^* T_o)$. However, if we instead interpret T_o as an operator acting in $\mathcal{X}$, then T_o^* is a relation with $\operatorname{mul}(T_o^*) = \operatorname{dom}(T_o)^\perp$. The self-adjoint relation Q_V defined in (ii) is equal to $Q_V =$

$T_o^*T_o$ where T_o^* is interpreted in the latter sense. This implies that

$$\begin{aligned}\operatorname{mul}(Q_V) &= \operatorname{mul}(T_o^*) = \operatorname{dom}(T_o)^\perp = \operatorname{dom}(T)^\perp = \ker\left(T^{-*}\right),\\ \ker(Q_V) &= \ker(T_o) = \ker(T) = \operatorname{dom}\left(T^{-*}\right)^\perp,\\ \operatorname{dom}(Q_V) &= \ker(T) + \operatorname{dom}(T_o^*T_o),\end{aligned}$$

where $\operatorname{dom}(T_o^*T_o)$ refers to the domain of $T_o^*T_o$ with T_o^* the single-valued operator as above, and that the graph of Q_V can be decomposed with respect to the decomposition of $\mathcal{X}$ in (A.3) as

$$Q_V = \left\{ \left(\begin{bmatrix} x_0 \\ x_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ T_i^*T_ix_1 \\ x_2 \end{bmatrix} \right) \middle| \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \in \begin{bmatrix} \ker(T) \\ \operatorname{dom}(T_i^*T_i) \\ \operatorname{dom}(T)^\perp \end{bmatrix} \right\}.$$

The same argument with T replaced by T^{-*} gives an analogous decomposition of $Q_{V^\perp}$, namely

$$Q_{V^\perp} = \left\{ \left(\begin{bmatrix} 0 \\ x_1^\dagger \\ x_0^\dagger \end{bmatrix}, \begin{bmatrix} x_2^\dagger \\ T_i^{-1}T_i^{-*}x_1^\dagger \\ 0 \end{bmatrix} \right) \middle| \begin{bmatrix} x_0^\dagger \\ x_1^\dagger \\ x_2^\dagger \end{bmatrix} \in \begin{bmatrix} \operatorname{dom}(T^{-*})^\perp \\ \operatorname{dom}\left(T_i^{-1}T_i^{-*}\right) \\ \ker(T^{-*}) \end{bmatrix} \right\}.$$

Here $\ker(T) = \operatorname{dom}(T^{-*})^\perp$, $\operatorname{dom}(T)^\perp = \ker(T^{-*})$, and $T_i^{-1}T_i^{-*} = (T_i^*T_i)^{-1}$. Thus, $Q_{V^\perp} = Q_V^{-1}$.

Proof of (iii). Let $Q_V^{1/2}$ and $Q_{V^\perp}^{1/2}$ be the nonnegative self-adjoint square roots of the relations Q_V and $Q_{V^\perp}$ whose graphs are given by

$$\begin{aligned} Q_V^{1/2} &= \left\{ \left(\begin{bmatrix} x_0 \\ x_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ Q_i^{1/2}x_1 \\ x_2 \end{bmatrix} \right) \middle| \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \in \begin{bmatrix} \ker(\|\cdot\|_V) \\ \operatorname{dom}\big(Q_i^{1/2}\big) \\ \ker(\|\cdot\|_{V^\perp}) \end{bmatrix} \right\},\\ Q_{V^\perp}^{1/2} &= \left\{ \left(\begin{bmatrix} 0 \\ x_1^\dagger \\ x_0^\dagger \end{bmatrix}, \begin{bmatrix} x_2^\dagger \\ Q_i^{-1/2}x_1^\dagger \\ 0 \end{bmatrix} \right) \middle| \begin{bmatrix} x_0^\dagger \\ x_1^\dagger \\ x_2^\dagger \end{bmatrix} \in \begin{bmatrix} \ker(\|\cdot\|_V) \\ \operatorname{img}\big(Q_i^{1/2}\big) \\ \ker(\|\cdot\|_{V^\perp}) \end{bmatrix} \right\}, \end{aligned}$$

where $Q_i = T_i^*T_i$ is the injective operator part of Q_V and

$$\begin{aligned}\mathcal{X}_V &= \ker(T) + \operatorname{dom}\big(Q_i^{1/2}\big) = \ker(\|\cdot\|_V) + \operatorname{dom}\big(Q_i^{1/2}\big),\\ \mathcal{X}_{V^\perp} &= \operatorname{img}\big(Q_i^{1/2}\big) + \ker\left(T^{-*}\right) = \operatorname{img}\big(Q_i^{1/2}\big) + \ker(\|\cdot\|_{V^\perp}).\end{aligned}$$

Let $x \in \mathcal{X}_V$ and $x^\dagger \in \mathcal{X}_{V^\perp}$, and decompose x and $x^\dagger$ according to the decomposition of $\mathcal{X}$ above into $x = x_0 + x_1$ and $x^\dagger = x_1^\dagger + x_0^\dagger$ with $x_0 \in \ker(\|\cdot\|_V)$, $x_1 \in \operatorname{dom}\big(Q_i^{1/2}\big)$, $x_0^\dagger \in \ker(\|\cdot\|_{V^\perp})$, and $x_1^\dagger \in \operatorname{img}\big(Q_i^{1/2}\big)$. Then $\|x\|_V = \|Q_i^{1/2}x_1\|_{\mathcal{X}}$, $\|x^\dagger\|_{V^\perp} = \|Q_i^{-1/2}x_1^\dagger\|_{\mathcal{X}}$, and

$$\langle x, x^\dagger\rangle_{\mathcal{X}} = \langle x_1, x_1^\dagger\rangle_{\mathcal{X}} = \langle Q_i^{-1/2}Q_i^{1/2}x_1, x_1^\dagger\rangle_{\mathcal{X}} = \langle Q_i^{1/2}x_1, Q_i^{-1/2}x_1^\dagger\rangle_{\mathcal{X}}.$$

Therefore, since $\overline{\operatorname{dom}\big(Q_i^{1/2}\big)} = \overline{\operatorname{img}\big(Q_i^{1/2}\big)}$,

$$\begin{aligned}\|x\|_V = \|Q_i^{1/2}x_1\|_{\mathcal{X}} &= \sup_{z_1^\dagger\in\operatorname{dom}(Q_i^{1/2}),\|z_1^\dagger\|_{\mathcal{X}}\le 1} |\langle Q_i^{1/2}x_1, z_1^\dagger\rangle_{\mathcal{X}}| \\ &= \sup_{x_1^\dagger\in\operatorname{img}(Q_i^{1/2}),\|Q_i^{-1/2}x_1^\dagger\|_{\mathcal{X}}\le 1} |\langle Q_i^{1/2}x_1, Q_i^{-1/2}x_1^\dagger\rangle_{\mathcal{X}}| \\ &= \sup_{x_1^\dagger\in\operatorname{img}(Q_i^{1/2}),\|x_1^\dagger\|_{V^\perp}\le 1} |\langle x_1, x_1^\dagger\rangle_{\mathcal{X}}| \\ &= \sup_{x^\dagger\in\mathcal{X}_{V^\perp},\|x^\dagger\|_{V^\perp}\le 1} |\langle x, x^\dagger\rangle_{\mathcal{X}}|,\end{aligned}$$

and thus $x \in \mathcal{X}_V$. This proves the first half (3.12). The proof of the second half of (3.12) is analogous.

It follows from (3.12) that if $x \in \mathcal{X}_V$ then the first supremum in (3.11) is finite, and if $x^\dagger \in \mathcal{X}_{V^\perp}$ then the second supremum in (3.11) is finite.

Conversely, suppose that the supremum in the first half of (3.11) is finite, and decompose x into $x = \left[\begin{smallmatrix} x_0 \\ x_1 \\ x_2 \end{smallmatrix}\right]$ in accordance with the decomposition in (A.3). Then $x \in \mathcal{X}_V$ if and only if $x_1 \in \operatorname{dom}\big(Q_i^{1/2}\big)$ and $x_2 = 0$.

For each $x^\dagger = \left[\begin{smallmatrix} 0 \\ x_1^\dagger \\ x_0^\dagger \end{smallmatrix}\right] \in \mathcal{X}_{V^\perp}$ with $x_1^\dagger \in \operatorname{img}\big(Q_i^{1/2}\big)$ and $x_0^\dagger \in \ker\left(\|\cdot\|_{V^\perp}\right)$ we have $\langle x, x^\dagger\rangle_{\mathcal{X}} = \langle x_1, x_1^\dagger\rangle_{\mathcal{X}} + \langle x_2, x_0^\dagger\rangle_{\mathcal{X}}$. Here $x_0^\dagger$ can be an arbitrary vector in $\ker\left(\|\cdot\|_{V^\dagger}\right)$, so the finiteness of the supremum in the first half of (3.11) implies that $x_2 = 0$. It remains to show that $x_1 \in \operatorname{dom}\big(Q_i^{1/2}\big)$.

Each $x_1^\dagger \in \operatorname{img}\big(Q_i^{1/2}\big)$ can be rewritten in the form $x_1^\dagger = Q_i^{1/2}z_1^\dagger$ where $z_1^\dagger \in \operatorname{dom}\big(Q_i^{1/2}\big)$ and $\|x_1^\dagger\|_{V^\perp} = \|Q_i^{-1/2}x_1^\dagger\|_{\mathcal{X}} = \|z_1^\dagger\|_{\mathcal{X}}$. Thus, the first half of (3.11) can be rewritten in the form

$$\sup_{z_1^\dagger\in\operatorname{dom}(Q_i^{1/2}),\|z_1^\dagger\|_{\mathcal{X}}\le 1} |\langle x_1, Q_i^{1/2}z_1^\dagger\rangle_{\mathcal{X}}| = \sup_{x^\dagger\in\mathcal{X}_{V^\perp},\|x^\dagger\|_{V^\perp}\le 1} |\langle x, x^\dagger\rangle_{\mathcal{X}}| < \infty.$$

This implies that $x_1 \in \operatorname{dom}\big((Q_i^{1/2})^*\big) = \operatorname{dom}\big(Q_i^{1/2}\big) = \mathcal{X}_V$, and completes the proof of the converse part of the first half of (3.11). The proof of the converse part of the second half of (3.11) is analogous.

REFERENCES

[1] D. Z. Arov, M. A. Kaashoek, and D. R. Pik. The Kalman-Yakubovich-Popov inequality for discrete time systems of infinite dimension. *J. Operator Theory*, 55(2):393–438, 2006.

[2] D. Z. Arov and M. A. Nudelman. Passive linear stationary dynamical scattering systems with continuous time. *Integral Equations Operator Theory*, 24(1):1–45, 1996.

[3] Damir Z. Arov and Olof J. Staffans. The infinite-dimensional continuous time Kalman-Yakubovich-Popov inequality. In *The extended field of operator theory*, volume 171 of *Oper. Theory Adv. Appl.*, pages 37–72. Birkhäuser, Basel, 2007.

[4] Ruth F. Curtain and Mark R. Opmeer. Normalized doubly coprime factorizations for infinite-dimensional linear systems. *Math. Control Signals Systems*, 18(1):1–31, 2006.

[5] Ruth F. Curtain and Hans Zwart. *An introduction to infinite-dimensional linear systems theory*, volume 21 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1995.

[6] Giuseppe Da Prato and Michel Delfour. Stabilization and unbounded solutions of the Riccati equation. In *Proc. CDC*, pages 352–357, 1988.

[7] Giuseppe Da Prato and Michel Delfour. Linear quadratic control problem without stabilizability. In *Stabilization of flexible structures (Montpellier, 1989)*, volume 147 of *Lecture Notes in Control and Inform. Sci.*, pages 126–147. Springer, Berlin, 1990.

[8] Giuseppe Da Prato and Michel Delfour. Unbounded solutions to the linear quadratic control problem. *SIAM J. Control Optim.*, 30(1):31–48, 1992.

[9] John C. Doyle, Keith Glover, Pramod P. Khargonekar, and Bruce A. Francis. State-space solutions to standard H_2 and H_∞ control problems. *IEEE Trans. Automat. Control*, 34(8):831–847, 1989.

[10] F. Flandoli, I. Lasiecka, and R. Triggiani. Algebraic Riccati equations with nonsmoothing observation arising in hyperbolic and Euler-Bernoulli boundary control problems. *Ann. Mat. Pura Appl. (4)*, 153:307–382 (1989), 1988.

[11] R. E. Kalman. Contributions to the theory of optimal control. *Bol. Soc. Mat. Mexicana (2)*, 5:102–119, 1960.

[12] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME Ser. D. J. Basic Engrg.*, 82:35–45, 1960.

[13] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Trans. ASME Ser. D. J. Basic Engrg.*, 83:95–108, 1961.

[14] Pramod P. Khargonekar and Eduardo D. Sontag. On the relation between stable matrix fraction factorizations and regulable realizations of linear systems over rings. *IEEE Trans. Automat. Control*, 27(3):627–638, 1982.

[15] I. Lasiecka and R. Triggiani. *Differential and algebraic Riccati equations with application to boundary/point control problems: continuous theory and approximation theory*, volume 164 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1991.

[16] Jarmo Malinen. Discussion on: "Boundary control of a class of hyperbolic systems". *Eur. J. Control*, 9(6):605–607, 2003.

[17] Jarmo Malinen, Olof J. Staffans, and George Weiss. When is a linear system conservative? *Quart. Appl. Math.*, 64(1):61–91, 2006.

[18] David G. Meyer and Gene F. Franklin. A connection between normalized coprime factorizations and linear quadratic regulator theory. *IEEE Trans. Automat. Control*, 32(3):227–228, 1987.

[19] Kalle M. Mikkola. Coprime factorization and dynamic stabilization of transfer functions. *SIAM J. Control Optim.*, 45(6):1988–2010, 2007.

[20] Kalle M. Mikkola. Weakly coprime factorization and continuous-time systems. *IMA J. Math. Control Inform.*, 25(4):515–546, 2008.

[21] C. N. Nett, C. A. Jacobson, and M. J. Balas. A connection between state-space and doubly coprime fractional representations. *IEEE Trans. Automat. Control*, 29(9):831–832, 1984.

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

[23] 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.

[24] Mark R. Opmeer and Olof J. Staffans. Optimal input-output stabilization of infinite-dimensional discrete time-invariant linear systems by output injection. *SIAM J. Control Optim.*, 48(8):5084–5107, 2010.

[25] Mark R. Opmeer and Olof J. Staffans. Coprime factorization and optimal control on the doubly infinite discrete time axis. *SIAM J. Control Optim.*, 50(1):266–285, 2012.

[26] Ian R. Petersen and Alexander Lanzon. Feedback control of negative-imaginary systems: flexible structures with colocated actuators and sensors. *IEEE Control Syst. Mag.*, 30(5):54–72, 2010.

[27] Dietmar Salamon. Realization theory in Hilbert space. *Math. Systems Theory*, 21(3):147–164, 1989.

[28] Yurii L. Šmuljan. Invariant subspaces of semigroups and the Lax–Phillips scheme. Deposited in VINITI, No. 8009-B86, Odessa, 49 pages, 1986.

[29] Olof J. Staffans. Quadratic optimal control of stable systems through spectral factorization. *Math. Control Signals Systems*, 8(2):167–197, 1995.

[30] Olof J. Staffans. Quadratic optimal control of stable well-posed linear systems. *Trans. Amer. Math. Soc.*, 349(9):3679–3715, 1997.

[31] Olof J. Staffans. Coprime factorizations and well-posed linear systems. *SIAM J. Control Optim.*, 36(4):1268–1292, 1998.

[32] Olof J. Staffans. Quadratic optimal control of well-posed linear systems. *SIAM J. Control Optim.*, 37(1):131–164, 1999.

[33] Olof J. Staffans. *Well-posed linear systems*, volume 103 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.

[34] Olof J. Staffans and George Weiss. Transfer functions of regular linear systems. II. The system operator and the Lax-Phillips semigroup. *Trans. Amer. Math. Soc.*, 354(8):3229–3262, 2002.

[35] M. Vidyasagar. *Control system synthesis: A factorization approach.* MIT Press Series in Signal Processing, Optimization, and Control, 7. MIT Press, Cambridge, MA, 1985.

[36] George Weiss. Weak L^p-stability of a linear semigroup on a Hilbert space implies exponential stability. *J. Differential Equations*, 76(2):269–285, 1988.

[37] George Weiss. Regular linear systems with feedback. *Math. Control Signals Systems*, 7(1):23–57, 1994.

[38] George Weiss. Transfer functions of regular linear systems. I. Characterizations of regularity. *Trans. Amer. Math. Soc.*, 342(2):827–854, 1994.

[39] George Weiss and Hans Zwart. An example in linear quadratic optimal control. *Systems Control Lett.*, 33(5):339–349, 1998.

[40] Martin Weiss and George Weiss. Optimal control of stable weakly regular linear systems. *Math. Control Signals Systems*, 10(4):287–330, 1997.