OPTIMAL INPUT-OUTPUT STABILIZATION OF INFINITE-DIMENSIONAL DISCRETE TIME-INVARIANT LINEAR SYSTEMS BY OUTPUT INJECTION

MARK R. OPMEER* AND OLOF J. STAFFANS†

Abstract. We study the optimal input-output stabilization of discrete time-invariant linear systems in Hilbert spaces by output injection. We show that a necessary and sufficient condition for this problem to be solvable is that the transfer function has a left factorization over H-infinity. Another equivalent condition is that the filter Riccati equation (of an arbitrary realization) has a solution (in general unbounded and even non densely defined). We further show that after renorming the state space in terms of the smallest solution of the filter Riccati equation, the closed-loop system is not only input-output stable, but also strongly internally $*$-stable.

Key words. Riccati equation, linear quadratic optimal control, input-output stabilization, output injection, infinite-dimensional system, left factorization.

AMS subject classifications. 47N70, 47A48, 47A20, 47A56, 47A62, 93B28, 93C55, 93D15

1. Introduction. This is the second in a series of articles dealing in a novel way with the quadratic cost minimization problem for infinite-dimensional time-invariant linear systems in discrete and continuous time. In the first article [6] we investigated the full information infinite-horizon LQ (Linear Quadratic) problem, and here we will study a deterministic version of the discrete time infinite-horizon Kalman filtering problem.

In [6] we studied the linear dynamical system in discrete *future time* defined by

$$x_{n+1} = Ax_n + Bu_n, \quad y_n = Cx_n + Du_n, \quad n \in \mathbb{Z}^+; \qquad x_0 = z, \tag{1.1}$$

where $A\colon \mathscr{X} \to \mathscr{X}$, $B\colon \mathscr{U} \to \mathscr{X}$, $C\colon \mathscr{X} \to \mathscr{Y}$, and $D\colon \mathscr{U} \to \mathscr{Y}$ are bounded linear operators, $\mathscr{X}$, $\mathscr{U}$ and $\mathscr{Y}$ are Hilbert spaces, and $\mathbb{Z}^+$ is the set of nonnegative integers. Here we shall study the same system in discrete *past time*

$$\begin{aligned} x_{n+1} = Ax_n + Bu_n, \quad y_n = Cx_n + Du_n, \quad n \in \mathbb{Z}^-; \qquad x_0 = z, \\ \exists N \in \mathbb{Z}^+ : x_n = 0 = u_n \quad \forall n \le -N, \end{aligned} \tag{1.2}$$

where $\mathbb{Z}^-$ is the set of negative integers.

A classical problem is to modify the properties of these systems by using either *state feedback* or *output injection*. State feedback was studied in [6], and here we focus on output injection. In the case of state feedback one chooses the control u_n to be given by $u_n = Kx_n + v_n$, where $K\colon \mathscr{X} \to \mathscr{U}$ is bounded linear (state feedback) operator, which in the future time setting results in the closed loop state feedback

*Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom (`m.opmeer@maths.bath.ac.uk`). Mark Opmeer thanks Åbo Akademi for its hospitality and the Academy of Finland and the Magnus Ehrnrooth Foundation for their financial support during his visit to Åbo.
† Department of Mathematics, Åbo Akademi University, FIN-20500 Åbo, Finland (`Olof.Staffans@abo.fi`).

system

$$\begin{aligned} x_{n+1} &= (A+BK)x_n + Bv_n, && n \in \mathbb{Z}^+, \\ y_n &= (C+DK)x_n + Dv_n, && n \in \mathbb{Z}^+, \\ u_n &= Kx_n + v_n, && n \in \mathbb{Z}^+, \\ x_0 &= z. \end{aligned} \tag{1.3}$$

Here the input u of the original system plays the role of one of the two outputs of the closed loop system, and the new input sequence to the closed loop system is v. In the case of output injection we relax the output equation in (1.2) by allowing a nonzero error term $w_n := Cx_n + Du_n - y_n$, and injecting a multiple Hw_n of this term back into the state equation in (1.2), where $H\colon \mathscr{Y} \to \mathscr{X}$. Thus, we now treat y like an input. This leads to the closed loop output injection system

$$\begin{aligned} x_{n+1} &= (A-HC)x_n + (B-HD)u_n + Hy_n, && n \in \mathbb{Z}^-, \\ w_n &= Cx_n + Du_n - y_n, && n \in \mathbb{Z}^-, \\ x_0 &= z, \\ \exists N \in \mathbb{Z}^- &: x_n = 0 = u_n = y_n \quad \forall n \leq -N. \end{aligned} \tag{1.4}$$

One typical goal is to make this closed loop system stable, or at least input-output stable in the sense that the mapping from the two input sequences u and y to the output sequence w is bounded from $\ell^2(\mathbb{Z}^-;\mathscr{U}\times\mathscr{Y})$ to $\ell^2(\mathbb{Z}^-;\mathscr{Y})$. In the optimal version of this problem one does not only require this input-output map to be bounded, but to have the smallest possible norm.

A special solution to this optimal control problem can be found in the following way if we assume $\mathscr{X}$, $\mathscr{U}$, and $\mathscr{Y}$ to be finite-dimensional. For each $z \in \mathscr{X}$ in the reachable subspace we look for the infimum of $\sum_{n=-\infty}^{-1}(\|y_n\|_{\mathscr{Y}}^2 + \|u_n\|_{\mathscr{U}}^2)$ over all input sequence u in (1.2) with finite support for which the final state satisfies $x_0 = z$. In some cases this infimum can be zero even if $z \neq 0$, namely when z is an eigenvector corresponding to an eigenvalue of A which is unstable and controllable but unobservable. To exclude this case we require the system to satisfy the *state coercive past cost condition*: There exists a finite constant M such all solutions of (1.2) satisfy $\|z\|_{\mathscr{X}}^2 \leq M^2 \sum_{n=-\infty}^{-1}(\|y_n\|_{\mathscr{Y}}^2 + \|u_n\|_{\mathscr{U}}^2)$. (We shall later replace this by a weaker condition.) Even when this condition holds the infimum is typically not achieved for a sequence u with finite support, but it is then possible to find unique ℓ^2-sequences u and y and a corresponding sequence x tending to zero as $n \to -\infty$ satisfying the first line of (1.2) which minimizes $\sum_{n=-\infty}^{-1}(\|y_n\|_{\mathscr{Y}}^2 + \|u_n\|_{\mathscr{U}}^2)$ within this class of solutions. The optimal solution is of output injection type, i.e., there exists a bounded linear operator $H\colon \mathscr{Y} \to \mathscr{X}$ such that the first equation in (1.4) holds, and this output injection operator H minimizes the norm of the map from y and u to w in (1.4). The optimal cost of a given final state $z \in \mathscr{X}$ can be written in the form $\langle z, Pz\rangle_{\mathscr{X}}$ for some bounded nonnegative self-adjoint operator P, and the output injection operator H is explicitly given by $H = -(A^*PC^* + BD^*)S_P^{-1}$ where $S_P = I + DD^* + CPC^*$. The optimal cost operator P is the minimal nonnegative self-adjoint solution of the so called filter Riccati equation. The output injection H that we get in this way is optimal even in a stronger sense: if we replace w_n in (1.4) by Ww_n for some bounded linear operator $W\colon \mathscr{Y} \to \mathscr{W}$, then it is still true that the same output injection operator minimizes the ℓ^2 operator norm from the pair $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ to w. The optimal norm of this operator is equal to the norm of $(WPW^*)^{1/2}$. Additionally, the same output injection operator also minimizes the ℓ^2 to ℓ^∞ operator norm from the pair $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ to w, as well as the ℓ^2 to $\mathscr{W}$ norm of the operator from $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ to w_{-1}. In stochastic control theory the system (1.4) with this particular choice of H is known as the Kalman filter, and x_0 in (1.4) is interpreted as the minimal variance estimate of the state at time zero based on past values of $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ of a stochastic version of (1.2). We refer the reader to [3, Section 5.3] for a discussion of the stochastic interpretation of (1.4).

In the above formulation of the input-output stabilization problem there is a hidden assumption which is redundant in the finite-dimensional case, but not in the infinite-dimensional case. Let us denote the different transfer functions $u \mapsto y$, $y \mapsto w$, and $u \mapsto w$ of the systems (1.2) and (1.4) by, respectively,

$$\begin{aligned} G_{u,y}(z) &= zC(I - zA)^{-1}B + D, \\ G_{y,w}(z) &= zC(I - z(A - HC))^{-1}H - I, \\ G_{u,w}(z) &= zC(I - z(A - HC))^{-1}(B - HD) + D. \end{aligned}$$

Then all of these are defined in a neighbourhood of the origin and satisfy $G_{u,y}(z) = G_{y,w}(z)^{-1}G_{u,w}(z)$ in this neighbourhood. The input-output stability of (1.4) implies that both $G_{y,w}$ and $G_{u,w}$ can be extended to H^∞-functions (i.e., bounded analytic functions) in the open unit disc $\mathbb{D}$. Thus, a *necessary condition for the input-output stabilizability of* (1.2) *by output injection is that the transfer function* $G_{u,y}$ *has a left* H^∞ *factorization in the unit disc.* This factorization condition is always satisfied in the finite-dimensional case. Moreover, in the finite-dimensional case every observable realization satisfies the coercive state past cost condition, so that the above outlined procedure can be applied after one factors out the unobservable subspace. In the infinite-dimensional case the situation is considerably more complicated. Obtaining a realization that satisfies the state coercive past cost condition is no longer a matter of simply factoring out the unobservable subspace: one has to choose the realization (and especially the norm in the state space) with care. In the continuous-time case this question is strongly related to choosing the proper function spaces on which to consider a given (formal) partial differential equation, a problem that is well-known to be extremely delicate. The situation is very similar to the situation in [6], where the analogous problem with the finite future cost condition was discussed.

The first main novelty in the present article is the introduction of a condition, which we call *the output coercive past cost condition*, which is weaker than the state coercive past cost condition: there exists a finite constant M such that all solutions of (1.2) satisfy $\|Cz\|^2_{\mathscr{Y}} \le M^2 \sum_{n=-\infty}^{-1} (\|y_n\|^2_{\mathscr{Y}} + \|u_n\|^2_{\mathscr{U}})$. Thus, compared to the state coercive past cost condition we have replaced the arbitrarily chosen finitely reachable vector z in (1.2) by Cz, and replaced the norm in $\mathscr{X}$ by the norm in $\mathscr{Y}$. Since $C\colon \mathscr{X} \to \mathscr{Y}$ is bounded, the output coercive past cost condition is clearly weaker than the state coercive past cost condition. We show (in Theorem 6.10) that every realization of a function that has a left H^∞ factorization satisfies this output coercive past cost condition and that the transfer function of any system that satisfies the output coercive past cost condition has a left H^∞ factorization. Theorem 6.10 also gives a third equivalent condition: the filter Riccati equation has a solution. This solution may not be bounded, or even densely defined. By allowing the optimal cost operator and the optimal output injection operator to be unbounded we are able (in Theorem 7.1) to extend the procedure described above so that *it can always be applied to the given system, as soon as the necessary condition that* $G_{u,y}$ *has a left* H^∞ *factorization holds.* The resulting closed loop system will be input-output stable and have a minimal input-output norm, but is not necessarily internally stable (it may not even be internally well-posed). By changing the norm in the state space (where the new norm is defined in terms of the solution of the input-output stabilization problem) and keeping the same formal operators we construct a new realization that does satisfy the state coercive past cost condition and whose closed-loop system is strongly internally $*$-stable. This change of norm for the open-loop system is considered in Theorem 3.11 and Remark 3.12 and for the closed-loop system in Theorem 7.2.

Another major feature of this work is that we extend the theory about the duality of the optimal control and optimal filtering from the standard setting where all possible operators are bounded to a setting which allows for unbounded and not even densely defined state feedback and output injection operators, as well as unbounded and not necessarily densely defined solutions of the control and filter Riccati equations. We prove that the original system satisfies the output coercive past cost condition if and only if the dual system satisfies the future incremental cost condition introduced in [6]. Instead of giving direct proofs (which is possible) of our main results for the optimal output injection problem we have chosen to base the proofs given in this article on the extended duality between state feedback and output injection mentioned above. The main advantage with this approach is that the proofs become simpler and shorter, thanks to the fact that we can make full use of the results proved in [6]. When we convert the results given in [6] for the control Riccati equation to the dual filter Riccati equation we get more or less for free a new and natural setting for the study of the filter Riccati equation, where the filter Riccati operator and the output injection operator are allowed to be unbounded, and not even densely defined.

2. Discrete-time systems. In this section we collect definitions and known results on discrete-time systems that are needed in this article.

We first associate some operators on sequences spaces to the dynamical systems (1.1) and (1.2). In the following definition $\ell^p_c(\mathbb{Z}^-;\mathscr{H})$ with $1 \le p \le \infty$ is the subspace of $\ell^p(\mathbb{Z}^-;\mathscr{H})$ consisting of sequences with compact support and $\mathfrak{s}(\mathbb{Z}^+;\mathscr{H})$ is the space of all sequences $\mathbb{Z}^+ \to \mathscr{H}$.

Definition 2.1.

- *The* input map $\mathcal{B} : \ell^p_c(\mathbb{Z}^-;\mathscr{U}) \to \mathscr{X}$ *is the map that sends* $\{u_n\}_{n\in\mathbb{Z}^-}$ *to* z*:*

$$\mathcal{B}u = \sum_{k=0}^{\infty} A^k B u_{-k-1}.$$

- *The* output map $\mathcal{C} : \mathscr{X} \to \mathfrak{s}(\mathbb{Z}^+;\mathscr{Y})$ *is the map that sends* z *to* $\{y_n\}_{n\in\mathbb{Z}^+}$*:*

$$(\mathcal{C}z)_n = CA^n z.$$

- *The* future input-output map $\mathcal{D} : \mathfrak{s}(\mathbb{Z}^+;\mathscr{U}) \to \mathfrak{s}(\mathbb{Z}^+;\mathscr{Y})$ *is the map that sends* $\{u_n\}_{n\in\mathbb{Z}^+}$ *to* $\{y_n\}_{n\in\mathbb{Z}^+}$*:*

$$(\mathcal{D}u)_n = \sum_{k=0}^{n-1} CA^k B u_k + D u_n.$$

- *The* past input-output map $\mathcal{D}_- : \ell^p_c(\mathbb{Z}^-;\mathscr{U}) \to \ell^p_c(\mathbb{Z}^-;\mathscr{Y})$ *is the map that sends* $\{u_n\}_{n\in\mathbb{Z}^-}$ *to* $\{y_n\}_{n\in\mathbb{Z}^-}$*:*

$$(\mathcal{D}_-u)_n = \sum_{k=-\infty}^{n-1} CA^{n-1-k} B u_k + D u_n.$$

If we equip $\ell^p_c(\mathbb{Z}^-;\mathscr{U})$ with its natural LF topology and $\mathfrak{s}(\mathbb{Z}^+;\mathscr{H})$ with its natural Fréchet topology, then the above maps are continuous.

DEFINITION 2.2. *A state z is called* finite-time reachable *if there exist sequences u, x, y such that* (1.2) *holds. The set of finite-time reachable states is denoted by Ξ_-. If the closure of Ξ_- equals $\mathscr{X}$, then the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is called* controllable. Note that $\Xi_- = R(\mathcal{B})$.

DEFINITION 2.3. *A state z is called* unobservable *if for initial condition z and zero input u the output y of the system* (1.1) *is zero. The set of unobservable states is denoted by $\mathscr{N}$. The node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is called observable if $\mathscr{N} = \{0\}$.* Note that $\mathscr{N} = N(\mathcal{C})$.

DEFINITION 2.4. *The node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is called*
- *exponentially stable if $r(A) < 1$,*
- *strongly stable if $\lim_{k\to\infty} A^k z = 0$ for all $z \in \mathscr{X}$,*
- *strongly $*$-stable if $\lim_{k\to\infty} A^{*k} z^* = 0$ for all $z^* \in \mathscr{X}'$,*
- *input stable if $\mathcal{B}$ extends to a bounded operator $\ell^2(\mathbb{Z}^-;\mathscr{U}) \to \mathscr{X}$,*
- *output stable if $\mathcal{C}$ is a bounded operator $\mathscr{X} \to \ell^2(\mathbb{Z}^+;\mathscr{Y})$,*
- *input-output stable if $\mathcal{D}$ restricts to a bounded operator $\ell^2(\mathbb{Z}^+;\mathscr{U}) \to \ell^2(\mathbb{Z}^+;\mathscr{Y})$,*
- *strongly internally stable if it is strongly stable, input stable, output stable and input-output stable,*
- *strongly internally $*$-stable if it is strongly $*$-stable, input stable, output stable and input-output stable,*

Exponential stability implies strong internal stability and strong internal $*$-stability, but the converse is not true. As announced in the introduction, by output injection and changing the norm in the state space we will be able to make the closed-loop system strongly internally $*$-stable, but it will in general not be exponentially stable.

3. The final state optimal control problem. In this section we investigate the open loop final state optimal control problem, the synthesis of the optimal control as an output injection is considered in Section 7.

We first of all define some spaces and operators that allow us to re-phrase the open loop final state optimal control problem in a form appropriate for the application of a standard optimization technique(the orhogonal projection lemma).

For a finite-time reachable state z define

$$\mathscr{W}_c(z) = \left\{ \begin{bmatrix} y \\ u \end{bmatrix} \in \ell^2_c(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U}) : \exists x \text{ such that (1.2) holds} \right\},$$

the set of compactly supported input-output trajectories with z as final state. Further define

$$\mathscr{G}_c := \left\{ \begin{bmatrix} y \\ u \end{bmatrix} \in \ell^2_c(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U}) : \exists x, z \text{ such that (1.2) holds} \right\},$$

the set of compactly supported input-output trajectories. Note that $\mathscr{G}_c$ is the inverse graph of the past input-output map.

Define the operator

$$\mathcal{J}_c : \mathscr{G}_c \to \mathscr{X}, \qquad \mathcal{J}_c \begin{bmatrix} y \\ u \end{bmatrix} = z,$$

where u, y and z are related by (1.2); i.e. $\mathcal{J}_c$ maps a compactly supported input-output trajectory to the corresponding final state.

Further define the operator

$$\Gamma_p : \mathscr{G}_c \to \mathfrak{s}(\mathbb{Z}^+; \mathscr{Y}), \qquad \Gamma_p = C\mathcal{J}_c,$$

that maps a compactly supported input-output trajectory on $\mathbb{Z}^-$ to the corresponding output on $\mathbb{Z}^+$ when the input is chosen to be zero on $\mathbb{Z}^+$. Finally, define the set of stable past input-output trajectories $\mathscr{G}$ as the closure of $\mathscr{G}_c$ in $\ell^2(\mathbb{Z}^-; \mathscr{Y} \times \mathscr{U})$. Note that $\mathscr{G}$ is the closure of the inverse graph of the past input-output map considered as an unbounded operator $\ell^2(\mathbb{Z}^-; \mathscr{U}) \to \ell^2(\mathbb{Z}^-; \mathscr{Y})$.

To obtain a satisfactory theory for the final state optimal control problem mentioned in the introduction, it is crucial to extend Γ_p to $\mathscr{G}$. For that we make the following assumption.

Definition 3.1. *A node satisfies the* output coercive past cost condition *if there exists a $M > 0$ such that for all $z \in \Xi_-$ and all $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathscr{W}_c(z)$*

$$\|Cz\|_{\mathscr{Y}} \le M \left\| \begin{bmatrix} y \\ u \end{bmatrix} \right\|_{\ell^2(\mathbb{Z}^-; \mathscr{Y} \times \mathscr{U})}.$$

A stronger condition (which ensures that $\mathcal{J}_c$ extends to $\mathscr{G}$) is the following.

Definition 3.2. *A node satisfies the* state coercive past cost condition *if there exists a $M > 0$ such that for all $z \in \Xi_-$ and all $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathscr{W}_c(z)$*

$$\|z\|_{\mathscr{X}} \le M \left\| \begin{bmatrix} y \\ u \end{bmatrix} \right\|_{\ell^2(\mathbb{Z}^-; \mathscr{Y} \times \mathscr{U})}.$$

Remark 3.3. *The output coercive past cost condition is equivalent to $C\mathcal{J}_c : \mathscr{G}_c \to \mathscr{Y}$ extending to a bounded operator $\mathscr{G} \to \mathscr{Y}$ and is equivalent to $\Gamma_p : \mathscr{G}_c \to \mathfrak{s}(\mathbb{Z}^+; \mathscr{Y})$ extending to a bounded operator $\mathscr{G} \to \mathfrak{s}(\mathbb{Z}^+; \mathscr{Y})$, where this latter space is equipped with its natural Fréchet space topology. The state coercive past cost condition is equivalent to $\mathcal{J}_c : \mathscr{G}_c \to \mathscr{X}$ extending to a bounded operator $\mathscr{G} \to \mathscr{X}$.*

Not only $\mathcal{J}_c$ but also its closure will play an important role. This closure will allow us to interpret the notion of final state for some non compactly supported input-output trajectories. In general, $\mathcal{J}_c$ may not be a closable operator. However the closure of the graph of $\mathcal{J}_c$ always defines a closed linear relation (or multi-valued operator) which we will denote by $\mathcal{J}$. The development of a satisfactory theory does not hinge on $\mathcal{J}$ being single-valued. Multi-valuedness of $\mathcal{J}$ relates to ill-posedness of the dynamical system defined on $\mathbb{Z}^-$ when the compact support assumption is not made. We recall some basic definitions regarding multi-valued operators.

Definition 3.4. *A multi-valued operator (or relation) $T : \mathscr{H}_1 \to \mathscr{H}_2$ is a subspace $\mathscr{V}_T$ of $\mathscr{H}_1 \times \mathscr{H}_2$. The operator T is called closed when the subspace $\mathscr{V}_T$ is closed. We have for the domain, kernel, range and multi-valued part of T:*

$$\begin{aligned}
D(T) &= \{h_1 \in \mathscr{H}_1 : \exists h_2 \text{ such that } (h_1, h_2) \in \mathscr{V}_T\},\\
N(T) &= \{h_1 \in \mathscr{H}_1 : (h_1, 0) \in \mathscr{V}_T\},\\
R(T) &= \{h_2 \in \mathscr{H}_2 : \exists h_1 \text{ such that } (h_1, h_2) \in \mathscr{V}_T\},\\
M(T) &= \{h_2 \in \mathscr{H}_2 : (0, h_2) \in \mathscr{V}_T\}.
\end{aligned}$$

LEMMA 3.5. *If the node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *satisfies the output coercive past cost condition, then* $M(\mathcal{J}) = N(\mathcal{C})$.

Proof. We denote the closure of Γ_p by $\overline{\Gamma}_p$. Since $\overline{\Gamma}_p = \mathcal{C}\mathcal{J}$, we have $z \in M(\mathcal{J})$ if and only if $\mathcal{C}z \in M(\overline{\Gamma}_p)$. Since under the output coercive past cost condition $\overline{\Gamma}_p$ is single-valued, $M(\overline{\Gamma}_p) = \{0\}$ and it follows that $z \in M(\mathcal{J})$ if and only if $\mathcal{C}z = 0$, i.e. $M(\mathcal{J}) = N(\mathcal{C})$. □

LEMMA 3.6. *If the node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *is observable and satisfies the output coercive past cost condition, then* $\mathcal{J}_c$ *is a closable operator.*

Proof. From Lemma 3.5 and observability it follows that $M(\mathcal{J}) = \{0\}$ so that $\mathcal{J}$ is single-valued. Hence $\mathcal{J}_c$ has a closed extension that is a single valued operator, so it is a closable operator. □

We define $\Xi_p := R(\mathcal{J})$. This space has the interpretation of finite cost reachable elements in the state space $\mathscr{X}$. For $z \in \Xi_p$ we define

$$\mathscr{W}(z) := \left\{ \begin{bmatrix} y \\ u \end{bmatrix} \in D(\mathcal{J}) : \mathcal{J} \begin{bmatrix} y \\ u \end{bmatrix} = z \right\},$$

the set of stable input-output trajectories with z as final state. The final state optimal control problem consists of finding the element of minimal norm in $\mathscr{W}(z)$. To solve that problem we utilize the following well-known orthogonal projection lemma.

LEMMA 3.7. *Let* $\mathscr{H}$ *be a Hilbert space and* $\mathscr{K}$ *a nonempty closed subspace of* $\mathscr{H}$. *Define, for* $h_0 \in \mathscr{H}$, *the affine set*

$$\mathscr{K}(h_0) := \{h \in \mathscr{H} : h = h_0 + k \text{ for some } k \in \mathscr{K}\}.$$

Then there exists a unique $h_{\min} \in \mathscr{K}(h_0)$ *such that*

$$\|h_{\min}\| = \min_{h \in \mathscr{K}(h_0)} \|h\|.$$

The vector $h_{\min}$ *is characterized by the fact that* $\mathscr{K}(h_0) \cap (\mathscr{H} \ominus \mathscr{K}) = \{h_{\min}\}$.

Proof. A proof can be found in many books, e.g. [4, Section 3.2]. □ Applying this orthogonal projection lemma to our problem gives the following.

LEMMA 3.8. *For any* $z \in \Xi_p$, *the space* $\mathscr{W}(z)$ *has a unique element of minimal norm which is characterized by the fact that it is in* $\ell^2(\mathbb{Z}^-; \mathscr{Y} \times \mathscr{U}) \ominus \mathscr{W}(0)$.

Proof. Apply the orthogonal projection lemma (Lemma 3.7) with $\mathscr{H} = \ell^2(\mathbb{Z}^-; \mathscr{Y} \times \mathscr{U})$ and $\mathscr{K} = \mathscr{W}(0)$. Then $\mathscr{W}(z)$ takes the role of $\mathscr{K}(h_0)$. We have that $\mathscr{W}(0) = \mathcal{N}(\mathcal{J})$ is a closed subspace since $\mathcal{J}$ is closed and $\mathscr{W}(z)$ is nonempty since $z \in \Xi_p$. □

We define the set $\mathscr{G}_{\text{opt}} := D(\mathcal{J}) \ominus \mathscr{W}(0)$. This set has the interpretation of the set of all optimal input-output trajectories with a final state in $\mathscr{X}$. We restrict $\mathcal{J}$ to this set to obtain the (possibly multi-valued) operator

$$\mathcal{J}_r : \mathscr{G}_{\text{opt}} \to \mathscr{X}.$$

This operator is clearly closed, injective and has range Ξ_p. We further define

$$\mathcal{I}_p : \Xi_p \subset \mathscr{X} \to \mathscr{G}_{\text{opt}}, \qquad \mathcal{I}_p = \mathcal{J}_r^{-1},$$

the closed operator that maps a final state to the corresponding optimal input-output trajectory. We note that $N(\mathcal{I}_p) = M(\mathcal{J}_r) = M(\mathcal{J})$, so that $\mathcal{I}_p$ is not injective if

$\mathcal{J}$ is multi-valued (i.e. in that case two different final states have the same optimal input-output trajectories). Note that

$$\mathcal{I}_p \mathcal{J} = P_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})\ominus\mathscr{W}(0)}|_{D(\mathcal{J})}, \tag{3.1}$$

the orthogonal projection onto $\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})\ominus\mathscr{W}(0)$, since both equal the map that sends an input-output trajectory to the optimal input-output trajectory with the same final state.

For the final state optimal control problem only the subspace Ξ_p of finite cost final states is of importance, the rest of the state space should be ignored. The norm in the state space is also not relevant in the final state optimal control problem. There is a more natural semi-norm on Ξ_p associated with the final state optimal control problem. On Ξ_p we define the semi-norm

$$\|z\|_p := \|\mathcal{I}_p z\|_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})}.$$

Note that since $\mathcal{I}_p$ is a closed operator, the associated nonnegative symmetric sesquilinear form $\langle \mathcal{I}_p z_1, \mathcal{I}_p z_2\rangle_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})}$ in $\mathscr{X}$ is closed.

The following lemma shows that the space Ξ_p and the semi-norm on it interact well with the node $\left[\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right]$ if the output coercive past cost condition is satisfied.

LEMMA 3.9. *If the node $\left[\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right]$ satisfies the output coercive past cost condition, then it maps $\Xi_p\times\mathscr{U}$ into $\Xi_p\times\mathscr{Y}$ and its restriction to $\Xi_p\times\mathscr{U}$ is bounded with respect to the semi-norm $\|\cdot\|_p$.*

Proof. The operator B obviously maps into Ξ_-, the space of finite-time reachable states. Since $\Xi_- \subset \Xi_p$, certainly B maps into Ξ_p. Since the input u defined by $u_{-1} = v$, $u_k = 0$ for $k < -1$ reaches Bv we have:

$$\|Bv\|_p^2 = \|\mathcal{I}_p Bv\|^2_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})} \le \|v\|^2_{\mathscr{U}} + \|Dv\|^2_{\mathscr{Y}}.$$

So B is bounded with respect to the semi-norm $\|\cdot\|_p$. Note that the output coercive past cost condition was not used for this.

For $z\in\Xi_p$ we have that for all $\left[\begin{smallmatrix} y \\ u\end{smallmatrix}\right]\in\mathscr{W}(z)$

$$\|Cz\|_{\mathscr{Y}} = \|\left(\Gamma_p \begin{bmatrix} y \\ u\end{bmatrix}\right)_0\|_{\mathscr{Y}} \le M\|\begin{bmatrix} y \\ u\end{bmatrix}\|_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})},$$

where we have used that $\Gamma_p : \mathscr{G} \to \mathfrak{s}(\mathbb{Z}^+;\mathscr{Y}\times\mathscr{U})$ is a bounded operator. In particular, the above holds for the element of minimal norm in $\mathscr{W}(z)$: $\left[\begin{smallmatrix} y \\ u\end{smallmatrix}\right] = \mathcal{I}_p z$. This shows that C is bounded with respect to the semi-norm $\|\cdot\|_p$.

If $z = \mathcal{J}\left[\begin{smallmatrix} y \\ u\end{smallmatrix}\right]$, then Az is the image under $\mathcal{J}$ of the trajectory obtained by shifting $\left[\begin{smallmatrix} y \\ u\end{smallmatrix}\right]$ one place to the left and adding $\left[\begin{smallmatrix} Cz \\ 0\end{smallmatrix}\right]$ at the last position. It follows that $Az\in\Xi_p$ if $z\in\Xi_p$ and that

$$\|\mathcal{I}_p Az\|^2_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})} \le \|\mathcal{I}_p z\|^2_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})} + \|Cz\|^2_{\mathscr{Y}},$$

or equivalently that

$$\|Az\|_p^2 \le \|z\|_p^2 + \|Cz\|^2_{\mathscr{Y}}.$$

Using that C is bounded with respect to the semi-norm $\|\cdot\|_p$ it follows that A is bounded with respect to this semi-norm. □

If $\mathcal{I}_p$ is injective, then $\|\cdot\|_p$ is a norm and Lemma 3.9 shows that we can extend the restriction of the node to $\Xi_p\times\mathscr{U}$ to a node whose state space is the completion of Ξ_p under the norm $\|\cdot\|_p$. In general, the situation is slightly more complicated but essentially the same. On the set $P_{\mathscr{X}\ominus\mathscr{N}}\Xi_p$, the seminorm $\|\cdot\|_p$ is in fact a norm since $\mathscr{N} = N(\mathcal{I}_p)$ by Lemma 3.5. Lemma 3.9 shows that, under the output coercive past cost condition, the conditions of [6, Theorem B.14] are satisfied so that the completed $\mathcal{I}_p$-compression of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ exists. This has state space $\mathscr{X}_p$, the completion under the $\|\cdot\|_p$ norm of $P_{\mathscr{X}\ominus\mathscr{N}}\Xi_p$. In case $\mathcal{I}_p$ is injective these two procedures obviously coincide.

Note that $\mathscr{X}_p$ may not be contained in the state space $\mathscr{X}$.

The completed $\mathcal{I}_p$-compression has $\mathcal{J}_c$, $\mathcal{J}$ and $\mathcal{I}_p$ operators which we will denote by $\mathcal{J}_c^{\mathscr{X}_p}$, $\mathcal{J}^{\mathscr{X}_p}$ and $\mathcal{I}_p^{\mathscr{X}_p}$ respectively.

LEMMA 3.10. *If the node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *satisfies the output coercive past cost condition, then* $\mathcal{J}_c = P_{\mathscr{X}\ominus\mathscr{N}}\mathcal{J}^{\mathscr{X}_p}|_{\mathscr{G}_c}$.

Proof. This follows directly from [6, Theorem B.16]. □

THEOREM 3.11. *The completed* $\mathcal{I}_p$*-compression satisfies the state coercive past cost condition and is observable. The operator* $\mathcal{I}_p^{\mathscr{X}_p}$ *is a unitary map onto its range. The operator* $\mathcal{J}^{\mathscr{X}_p}$ *is a partial isometry with kernel* $\mathscr{W}(0)$.

Proof. Using (3.1) we have the following for $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathscr{G}_c$

$$\begin{aligned}\left\|\mathcal{J}_c^{\mathscr{X}_p}\begin{bmatrix} y \\ u \end{bmatrix}\right\|_p &= \left\|\mathcal{I}_p^{\mathscr{X}_p}\mathcal{J}_c^{\mathscr{X}_p}\begin{bmatrix} y \\ u \end{bmatrix}\right\|_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})} \\ &= \left\|P_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})\ominus N(\mathcal{J}^{\mathscr{X}_p})}\begin{bmatrix} y \\ u \end{bmatrix}\right\|_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})} \\ &\le \left\|\begin{bmatrix} y \\ u \end{bmatrix}\right\|_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})},\end{aligned}$$

which shows that $\mathcal{J}_c^{\mathscr{X}_p} : \mathscr{G}_c \to \mathscr{X}_p$ is bounded. It follows that the state coercive past cost condition is satisfied.

We just showed that $\mathcal{J}_c^{\mathscr{X}_p}$ has a single valued bounded extension to $\mathscr{G}$ and since the state coercive past cost condition is satisfied (which implies the output coercive past cost condition), it follows from Lemma 3.5 that $N(\mathcal{C}^{\mathscr{X}_p}) = \{0\}$, so that the completed $\mathcal{I}_p$-compression is indeed observable.

By definition of norms, $\mathcal{I}_p^{\mathscr{X}_p}$ is an isometry; so it is a unitary map onto its range.

By definition of the norm and (3.1) we have for $g \in D(\mathcal{J}^{\mathscr{X}_p})$

$$\|\mathcal{J}^{\mathscr{X}_p}g\|_p = \|\mathcal{I}_p^{\mathscr{X}_p}\mathcal{J}^{\mathscr{X}_p}g\|_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})} = \|P_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})\ominus\mathscr{W}(0)}g\|_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})},$$

which implies that $\mathcal{J}^{\mathscr{X}_p}$ is a partial isometry with kernel $\mathscr{W}(0)$. □

Remark 3.12. *Define* $\mathcal{I}_{p,-} := \mathcal{I}_p|_{\Xi_-}$. *Similar as for the completed* $\mathcal{I}_p$*-compression, it can be shown that the completed* $\mathcal{I}_{p,-}$*compression is well-defined. It has as state space* $\mathscr{X}_{p,-}$, *the completion of* $P_{\mathscr{X}\ominus\mathscr{N}}\Xi_-$ *under the norm* $\|\cdot\|_p$, *and is controllable, observable and satisfies the state coercive past cost condition.*

4. Recap of the initial state optimal control problem. In this section we review the relevant results from [6], which by the duality theory of the next two sections will lead to synthesis of the optimal control as an output injection in Section 7.

The system under study in this section is the initial state problem (1.1) with the associated cost function $\sum_{n=0}^{\infty}\|u_n\|^2 + \|y_n\|^2$. For $z \in \mathscr{X}$, define

$$\mathscr{V}(z) := \left\{\begin{bmatrix} u \\ y \end{bmatrix} \in \ell^2(\mathbb{Z}^+;\mathscr{U}\times\mathscr{Y}) : \exists x \text{ such that (1.1) holds}\right\},$$

the set of stable input-output trajectories with z as initial state. Further define Ξ_f as the subspace of $\mathscr{X}$ consisting of those z for which $\mathscr{V}(z)$ is non-empty. This is the subspace of finite future cost states (denotes by Ξ_+ in [6]).

The orthogonal projection lemma guarantees that for $z \in \Xi_f$, $\mathscr{V}(z)$ has a unique element of minimal norm. This provides us with a closed operator $\mathcal{I}_f : \Xi_f \subset \mathscr{X} \to \ell^2(\mathbb{Z}^+; \mathscr{U} \times \mathscr{Y}) \ominus \mathscr{V}(0)$, the future minimizing cost operator, which maps a finite cost initial state to the corresponding optimal input-output trajectory.

Definition 4.1. *The finite future incremental cost condition is the condition* $B\mathscr{U} \subset \Xi_f$ *(equivalently:* $\Xi_- \subset \Xi_f$*). The finite future cost condition is the condition* $\Xi_f = \mathscr{X}$.

If the finite future incremental cost condition holds, then the operator

$$\Gamma_f := \mathcal{I}_f \mathcal{B} : \ell^2_c(\mathbb{Z}^-; \mathscr{U}) \to \ell^2(\mathbb{Z}^+; \mathscr{U} \times \mathscr{Y})$$

that maps a compactly supported past input to the corresponding optimal future input-output trajectory is well-defined and bounded. If the finite future cost condition holds, then $\mathcal{I}_f : \mathscr{X} \to \ell^2(\mathbb{Z}^+; \mathscr{U} \times \mathscr{Y})$ is bounded.

On Ξ_f define the semi-norm

$$\|z\|_f := \|\mathcal{I}_f z\|_{\ell^2(\mathbb{Z}^+; \mathscr{U} \times \mathscr{Y})}.$$

Lemma 4.2. *If the node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *satisfies the finite future incremental cost condition, then it maps* $\Xi_f \times \mathscr{U}$ *into* $\Xi_f \times \mathscr{Y}$ *and its restriction to* $\Xi_f \times \mathscr{U}$ *is bounded with respect to the semi-norm* $\|\cdot\|_f$.

Proof. This follows from [6, Lemma 4.8] with $q = q_f$. ☐

Similarly to $\mathscr{X}_p$, the state space $\mathscr{X}_f$ is defined as the completion under the $\|\cdot\|_f$ norm of $P_{\mathscr{X} \ominus \mathscr{N}} \Xi_f$. The restriction of the node mentioned in Lemma 4.2 extends continuously to a node with this state space. That node is called the completed $\mathcal{I}_f$-compression (the completed q_f-compression in [6]) and it satisfies the finite future cost condition and is observable. Similarly, we may define $\mathcal{I}_{f,-} := \mathcal{I}_f|_{\Xi_-}$ and the completed $\mathcal{I}_{f,-}$-compression (the completed q_f^--compression in [6]) which has as state space $\mathscr{X}_{f,-}$, the completion of $P_{\mathscr{X} \ominus \mathscr{N}} \Xi_-$ under the norm $\|\cdot\|_f$, and satisfies the finite future cost condition and is both controllable and observable.

Note that $\mathscr{X}_f$ and $\mathscr{X}_{f,-}$ may not be contained in the state space $\mathscr{X}$.

Remark 4.3. *Denote the* $\mathcal{I}_f$ *operator of the completed* $\mathcal{I}_f$*-compression by* $\mathcal{I}_f^{\mathscr{X}_f}$, *then* $\mathcal{I}_f^{\mathscr{X}_f}$ *is an isometry onto its range (which equals the closure of the range of the* $\mathcal{I}_f$ *operator of the original node that the completed* $\mathcal{I}_f$*-compression was constructed from). The inverse of* $\mathcal{I}_f^{\mathscr{X}_f}$ *(defined on the range of* $\mathcal{I}_f^{\mathscr{X}_f}$*) is a unitary map that sends the optimal input-output trajectory to the initial state.*

The following is the standard control algebraic Riccati equation re-written in a way (using sesquilinear forms) that easily allows for unbounded solutions.

Definition 4.4. *The triple* (q, s, K) *is called a (nonnegative) solution of the* control Riccati equation *of the node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *if*

1. *q is a closed nonnegative symmetric sesquilinear form in $\mathscr{X}$ whose domain satisfies $AD(q) \subset D(q)$, $B\mathscr{U} \subset D(q)$.*
2. *s is a bounded nonnegative symmetric sesquilinear form on $\mathscr{U}$.*
3. *$K : D(q) \to \mathscr{U}$ is a linear operator.*
4. *For all $z \in D(q)$, $u \in \mathscr{U}$ we have*

$$q(Az+Bu, Az+Bu) + \|Cz+Du\|^2_{\mathscr{Y}} + \|u\|^2_{\mathscr{U}} = q(z,z) + s(Kz-u, Kz-u). \quad (4.1)$$

The solution is called classical *when* $D(q) = \mathscr{X}$.

To discuss transfer functions, we use the following notation: H^∞ denotes the Hardy space of uniformly bounded holomorphic functions and $\mathbb{D}$ denotes the unit disc. The transfer function of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is defined in a neighbourhood of zero by $zC(I - zA)^{-1}B + D$. A node is called a *realization* of a holomorphic function defined in a neighbourhood of zero if that function is the transfer function of the node. We note that any holomorphic function defined in a neighbourhood of zero has a realization (in fact, it has infinitely many).

Definition 4.5. *Let* $\mathsf{G} : D(\mathsf{G}) \subset \mathbb{C} \to \mathcal{L}(\mathscr{U}, \mathscr{Y})$ *be holomorphic at the origin. A function* $\left[\begin{smallmatrix} \mathsf{M} \\ \mathsf{N} \end{smallmatrix}\right] \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{U}, \mathscr{U} \times \mathscr{Y}))$ *is called a* right factorization *of* G *if* $\mathsf{M}(z)$ *is invertible for all* z *in a neighbourhood of the origin and* $\mathsf{G}(z) = \mathsf{N}(z)\mathsf{M}(z)^{-1}$ *in a neighbourhood of the origin.*

Theorem 4.6. *Let* $\mathsf{G} : D(\mathsf{G}) \subset \mathbb{C} \to \mathcal{L}(\mathscr{U}, \mathscr{Y})$ *be holomorphic at the origin and let* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *be a realization of* G. *The following are equivalent conditions.*

- $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *satisfies the finite future incremental cost condition.*
- *The control Riccati equation of* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *has a (nonnegative self-adjoint) solution.*
- G *has a right factorization.*

Under these equivalent conditions, the triple (q_f, s_f, K_f) *defined by*

$$\begin{aligned} q_f(z_1, z_2) &:= \langle \mathcal{I}_f z_1, \mathcal{I}_f z_2 \rangle_{\ell^2(\mathbb{Z}^+; \mathscr{U} \times \mathscr{Y})}, \\ s_f(u, v) &:= \langle u, v \rangle_{\mathscr{U}} + \langle Du, Dv \rangle_{\mathscr{Y}} + q_f(Bu, Bv), \\ K_f z &= P_{\mathscr{U}}(\mathcal{I}_f z)_0, \end{aligned}$$

is the smallest nonnegative self-adjoint solution of the control Riccati equation. Here $P_{\mathscr{U}}$ *is the canonical projection* $\mathscr{U} \times \mathscr{Y} \to \mathscr{U}$.

Proof. This follows from [6, Theorem 6.3] combined with [6, Theorem 3.14]. □

We consider the closed-loop system

$$\begin{aligned} x_{n+1} &= (A + BK)x_n + BEw_n, & n \in \mathbb{Z}^+, \\ y_n &= (C + DK)x_n + DEw_n, & n \in \mathbb{Z}^+, \\ u_n &= Kx_n + Ew_n, & n \in \mathbb{Z}^+, \\ x_0 &= z, \end{aligned} \tag{4.2}$$

where $E \colon \mathscr{W} \to \mathscr{U}$ is a bounded linear operator and $K : D(K) \subset \mathscr{X} \to \mathscr{U}$ is a linear operator with a domain that is A-invariant and that contains the image of B. For such a K, the map from $\{w_n\}_{n \in \mathbb{Z}^+}$ to $\{\left[\begin{smallmatrix} u_n \\ y_n \end{smallmatrix}\right]\}_{n \in \mathbb{Z}^+}$ in (4.2) (with $z = 0$) is well-defined on the sequences with compact support.

Theorem 4.7. *Assume that the finite future incremental cost condition holds. Then* K_f *minimizes both the* $\mathcal{L}(\ell^1(\mathbb{Z}^+, \mathscr{W}), \ell^2(\mathbb{Z}^+, \mathscr{U} \times \mathscr{Y}))$ *and the* $\mathcal{L}(\ell^2(\mathbb{Z}^+, \mathscr{W}), \ell^2(\mathbb{Z}^+, \mathscr{U} \times \mathscr{Y}))$ *norm of the map from* $\{w_n\}_{n \in \mathbb{Z}^+}$ *to* $\{\left[\begin{smallmatrix} u_n \\ y_n \end{smallmatrix}\right]\}_{n \in \mathbb{Z}^+}$ *in* (4.2) *(with* $z = 0$*), where* K *ranges over all linear maps* $D(K) \subset \mathscr{X} \to \mathscr{U}$ *with a domain that is* A*-invariant and that contains the image of* B. *The operator* K_f *also minimizes the* $\mathcal{L}(\mathscr{W}, \ell^2(\mathbb{Z}^+, \mathscr{U} \times \mathscr{Y}))$ *norm of the map* $w_0 \to \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$ *over the same set of feedback operators.*

These minimum norms all equal the square root of $\sup_{\|v\|=1} s_f(Ev, Ev)$.

Proof. The statements on the $\mathcal{L}(\ell^1(\mathbb{Z}^+, \mathscr{W}), \ell^2(\mathbb{Z}^+, \mathscr{U} \times \mathscr{Y}))$ and $\mathcal{L}(\ell^2(\mathbb{Z}^+, \mathscr{W}), \ell^2(\mathbb{Z}^+, \mathscr{U} \times \mathscr{Y}))$ norms follow directly from [6, Theorem 5.1]. The statement about the $\mathcal{L}(\mathscr{W}, \ell^2(\mathbb{Z}^+, \mathscr{U} \times \mathscr{Y}))$ norm follows from slightly adapting the proof of [6, Theorem 5.1] as follows. The lower bound proof remains unchanged and in the upper bound proof v now has to be chosen as $v_0 = Ew_0$ and $v_k = 0$ for $k > 0$. □

Remark 4.8. *The operator underlying the closed-loop system* (4.2) *is generally not a node with state space* $\mathscr{X}$ *(because* K *is generally not bounded on* $\mathscr{X}$*), but it does become a node with state space* $\mathscr{X}_f$ *once we replace* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *with its completed* $\mathcal{I}_f$*-compression. This resulting closed-loop node equals what is called the completed* q_f*-compression of the graph closed-loop node in [6, Theorem 5.3]. Hence, by that theorem, it is strongly internally stable.*

5. Duality of discrete-time systems. In this section we re-consider duality for discrete-time systems in order to investigate the duality between the initial state and final state optimal control problems in the next section.

We consider the adjoint of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ as an operator from $\mathscr{X}' \times \mathscr{Y}'$ to $\mathscr{X}' \times \mathscr{U}'$ where $\mathscr{H}'$ denotes the dual space of the Hilbert space $\mathscr{H}$. A dual space is not identified with the Hilbert space itself unless this is explicitly stated. We denote the duality product between $\mathscr{H}$ and its dual $\mathscr{H}'$ by $\langle \cdot, \cdot \rangle_{\mathscr{H},\mathscr{H}'}$ and consider this to be linear in the $\mathscr{H}$ component and anti-linear in the $\mathscr{H}'$ component.

The dynamical system that we associate to $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$ is the following initial state problem

$$x^*_{n+1} = A^* x^*_n + C^* y^*_n, \quad u^*_n = B^* x^*_n + D^* y^*_n, \quad n \in \mathbb{Z}^+; \qquad x^*_0 = z^*, \tag{5.1}$$

with state space $\mathscr{X}'$, input space $\mathscr{Y}'$ and output space $\mathscr{U}'$. We throughout apply the theory from Section 4 to this dual system. To indicate the distinction between spaces associated to the primal system and the dual system we often use the subscript d, so e.g $\mathscr{V}_d(z^*)$ is the space introduced in the beginning of Section 4 consisting of stable input-output trajectories but now *for the dual system* with initial state z^*.

We define the weighted ℓ^p spaces

$$\ell^p_r(\mathbb{Z}^-; \mathscr{U}) = \{u : \mathbb{Z}^- \to \mathscr{U} : (r^{-n} u_n)_{n \in \mathbb{Z}^-} \in \ell^p(\mathbb{Z}^-; \mathscr{U})\},$$
$$\ell^p_r(\mathbb{Z}^+; \mathscr{Y}) = \{y : \mathbb{Z}^+ \to \mathscr{Y} : (r^{-n} y_n)_{n \in \mathbb{Z}^+} \in \ell^p(\mathbb{Z}^+; \mathscr{Y})\}.$$

Any continuous linear functional on $\ell^2_r(\mathbb{Z}^-; \mathscr{H})$ is of the form

$$\sum_{n=0}^{\infty} \langle h^-_{-n-1}, h^+_n \rangle_{\mathscr{H},\mathscr{H}'} \tag{5.2}$$

for some $h^+ \in \ell^2_r(\mathbb{Z}^+; \mathscr{H}')$ and any such h^+ through the expression (5.2) gives rise to a continuous linear functional on $\ell^2_r(\mathbb{Z}^-; \mathscr{H})$. Similarly, any continuous anti-linear functional on $\ell^2_r(\mathbb{Z}^+; \mathscr{H}')$ is of the form (5.2) for some $h^- \in \ell^2_r(\mathbb{Z}^-; \mathscr{H})$ and any such h^- gives rise to a continuous anti-linear functional. So we may treat $\ell^2_r(\mathbb{Z}^-; \mathscr{H})$ and $\ell^2_r(\mathbb{Z}^+; \mathscr{H}')$ as each others duals. With some abuse of notation we denote the duality product by

$$\langle h^-, h^+ \rangle_{\ell^2(\mathscr{H})} := \sum_{n=0}^{\infty} \langle h^-_{-n-1}, h^+_n \rangle_{\mathscr{H},\mathscr{H}'}. \tag{5.3}$$

The dual of $\ell^1_r(\mathbb{Z}^+; \mathscr{H})$ can similarly be identified with $\ell^\infty_r(\mathbb{Z}^-; \mathscr{H}')$ through the expression (5.2). The dual of the subspace $\ell^\infty_{r,0}(\mathbb{Z}^-; \mathscr{H})$ of $\ell^\infty_r(\mathbb{Z}^-; \mathscr{H})$ consisting of those sequences h such that $\lim_{k \to -\infty} r^{-k} h_k = 0$ can be identified with $\ell^1_r(\mathbb{Z}^+; \mathscr{H}')$ through (5.2).

If $r > r(A)$, the spectral radius of A, then the input map extends to a bounded operator $\ell^2_r(\mathbb{Z}^-;\mathscr{U}) \to \mathscr{X}$ and the output map is a bounded operator $\mathscr{X} \to \ell^2_r(\mathbb{Z}^+;\mathscr{Y})$. The input map $\mathcal{B}$ of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is adjoint to the output map $\mathcal{C}_d$ of the dual node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$ in the sense that

$$\langle \mathcal{B}u, z^*\rangle_{\mathscr{X},\mathscr{X}'} = \langle u, \mathcal{C}_d z^*\rangle_{\ell^2(\mathscr{U})}.$$

The past input-output map of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ extends to a bounded operator $\ell^2_r(\mathbb{Z}^-;\mathscr{U}) \to \ell^2_r(\mathbb{Z}^-;\mathscr{Y})$ and the future input-output map of the dual node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$ restricts to a bounded operator $\ell^2_r(\mathbb{Z}^+;\mathscr{Y}') \to \ell^2_r(\mathbb{Z}^+;\mathscr{U}')$. With the above identification of dual spaces, these operators are adjoints. Similarly, the restriction of the future input-output map of the adjoint node to a bounded operator $\ell^1_r(\mathbb{Z}^+;\mathscr{Y}') \to \ell^2_r(\mathbb{Z}^+;\mathscr{U}')$ and the extension of the past input-output map to a bounded operator $\ell^2_r(\mathbb{Z}^-;\mathscr{U}) \to \ell^\infty_{r,0}(\mathbb{Z}^-;\mathscr{Y})$ are adjoint operators with the above identification of dual spaces.

The following lemma characterizes duality in terms of trajectories without explicit reference to the node. It can be derived from [1, Lemma 4.6], but for the convenience of the reader we include a direct proof.

LEMMA 5.1. *If* $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathscr{G}_c$, $z = \mathcal{J}_c \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ *and* $\begin{bmatrix} y^* \\ u^* \end{bmatrix}$ *is a trajectory of* (5.1) *with initial condition* z^**, then*

$$\langle z, z^*\rangle_{\mathscr{X},\mathscr{X}'} = \left\langle \mathcal{R}\begin{bmatrix} y \\ u \end{bmatrix}, \begin{bmatrix} y^* \\ u^* \end{bmatrix}\right\rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})}, \tag{5.4}$$

where the operator $\mathcal{R}$ *is defined by*

$$\mathcal{R} : \ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U}) \to \ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U}), \qquad \mathcal{R}\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} -y \\ u \end{bmatrix},$$

and we have used the duality (5.3)*. Conversely, if* $\begin{bmatrix} y^* \\ u^* \end{bmatrix} \in \mathfrak{s}(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}')$ *and* $z^* \in \mathscr{X}'$ *satisfy* (5.4) *for all* $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathscr{G}_c$ *with* $z = \mathcal{J}_c\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$*, then* $\begin{bmatrix} y^* \\ u^* \end{bmatrix}$ *is a trajectory of* (5.1) *with initial condition* z^**.*

Proof. The first part of the lemma simply follows by substitution. The converse follows by iteratively applying the assumption as follows. Apply (5.4) with $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ the sequence that is zero everywhere except at position -1 where it equals $\left[\begin{smallmatrix} Dv \\ v \end{smallmatrix}\right]$ with $v \in \mathscr{U}$ arbitrary. It follows that

$$\langle v, u_0^* - D^*y_0^*\rangle_{\mathscr{U},\mathscr{U}'} = \left\langle \begin{bmatrix} -Dv \\ v \end{bmatrix}, \begin{bmatrix} y_0^* \\ u_0^* \end{bmatrix}\right\rangle = \langle Bv, z^*\rangle = \langle v, B^*z^*\rangle,$$

which since v was arbitrary implies $u_0^* = B^*z^* + D^*y_0^*$. Taking the element of $\mathscr{G}_c$ whose second component is zero everywhere except at position -2 where it equals v and whose first component is the corresponding output (i.e. Dv at position -2, CBv at position -1 and zero elsewhere) gives $u_1^* = B^*A^*z^* + D^*y_1^* + B^*C^*y_0^*$. Continuing in this fashion shows that $\begin{bmatrix} y^* \\ u^* \end{bmatrix}$ is an input-output trajectory of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$ with initial condition z^* as desired. ∎

We will also use the adjoint of $\mathcal{R}$:

$$\mathcal{R}^* : \ell^2(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}') \to \ell^2(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}'), \qquad \mathcal{R}^*\begin{bmatrix} y^* \\ u^* \end{bmatrix} = \begin{bmatrix} -y^* \\ u^* \end{bmatrix}.$$

Since we do not identify the dual of a Hilbert space with itself, a subspace $\mathscr{V}$ of a Hilbert space $\mathscr{H}$ has two orthogonal subspaces: the subspace $\mathscr{V}^\perp \subset \mathscr{H}'$ of continuous linear functionals on $\mathscr{H}$ that are zero on the subspace $\mathscr{V}$ and the subspace $\mathscr{H} \ominus \mathscr{V}$ that is the orthogonal complement of $\mathscr{V}$ in the sense that $\mathscr{V} \oplus (\mathscr{H} \ominus \mathscr{V}) = \mathscr{H}$. In this article we use both of these notions of orthogonal subspace and use the notations $\perp$ and $\ominus$ as above to distinguish these two notions.

Remark 5.2. *We identify the dual G' of a closed subspace G of H by the corresponding subspace of H', so that $\langle g, g' \rangle_{G,G'} = \langle g, g' \rangle_{H,H'}$ for all $g \in G$ and $g' \in G'$. Under this identification $G' = H' \ominus G^\perp$ and $G^\perp = (H \ominus G)'$. In particular, we have $\mathscr{G}' = \ell^2(\mathbb{Z}^+; \mathscr{Y}' \times \mathscr{U}') \ominus \mathscr{G}^\perp$.*

Remark 5.3. *The above duality set-up is somewhat non-standard. We would in principle lose nothing by using the standard duality set-up (i.e. identifying the dual of a Hilbert space with the Hilbert space itself and sequences spaces on $\mathbb{Z}^+$ with sequence spaces on $\mathbb{Z}^+$) which has dual systems running backwards in time. However, in [5] we will consider an optimal control problem on $\mathbb{Z}$ where the state has to pass through a target state x_0 at $n = 0$ and this problem naturally breaks down into a final state problem on $\mathbb{Z}^-$ and an initial state problem on $\mathbb{Z}^+$. This is our main reason for wanting to study final state systems defined on $\mathbb{Z}^-$ and therefore to identify the dual of an initial state system defined on $\mathbb{Z}^+$ to be a final state system defined on $\mathbb{Z}^-$. The Kalman filter is also most naturally posed on $\mathbb{Z}^-$, which is another reason for treating the equivalent optimal output injection problem on $\mathbb{Z}^-$ as well. Not identifying the state space $\mathscr{X}$ with its dual is appropriate since we want to identify the 'natural state space' $\mathscr{X}_p$ on which to consider the final state problem as the dual of the 'natural state space' $\mathscr{X}_{f,d}$ on which to consider the initial state problem for the dual system (Lemma 6.6). The only reason for not identifying $\mathscr{U}$ and $\mathscr{Y}$ with their respective duals is consistency.*

6. Duality between the optimal control problems. We next lemma relates the spaces $\mathscr{G}$ and $\mathscr{V}_d(0)$ of stable past and future input-output trajectories.

LEMMA 6.1. *We have $\mathcal{R}\mathscr{G} = \mathscr{V}_d(0)^\perp$.*

Proof. It immediately follows from (5.4) with $z^* = 0$ that $\mathcal{R}\mathscr{G}_c \perp \mathscr{V}_d(0)$. By continuity of $\mathcal{R}$ we conclude that $\mathcal{R}\mathscr{G} \perp \mathscr{V}_d(0)$, so that $\mathcal{R}\mathscr{G} \subset \mathscr{V}_d(0)^\perp$.

We now prove that $(\mathcal{R}\mathscr{G}_c)^\perp \subset \mathscr{V}_d(0)$, which through $\mathscr{V}_d(0)^\perp \subset (\mathcal{R}\mathscr{G}_c)^{\perp\perp} = \overline{\mathcal{R}\mathscr{G}_c} = \mathcal{R}\overline{\mathscr{G}_c} = \mathcal{R}\mathscr{G}$ gives the desired other inclusion. So assume that $\begin{bmatrix} y^* \\ u^* \end{bmatrix}$ is orthogonal to $\mathcal{R}\mathscr{G}_c$. Then $\begin{bmatrix} y^* \\ u^* \end{bmatrix}$ satisfies (5.4) for all $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathscr{G}_c$ with $z^* = 0$ and it follows from Lemma 5.1 that $\begin{bmatrix} y^* \\ u^* \end{bmatrix}$ is a trajectory of the dual system with initial condition zero. Since by assumption $\begin{bmatrix} y^* \\ u^* \end{bmatrix} \in \ell^2(\mathbb{Z}^+; \mathscr{Y}' \times \mathscr{U}')$ we have that it is an element of $\mathscr{V}_d(0)$ as desired. ∎

LEMMA 6.2. *We have $\mathcal{R}^* \left(\ell^2(\mathbb{Z}^+; \mathscr{Y}' \times \mathscr{U}') \ominus \mathscr{V}_d(0)\right) = \mathscr{G}'$.*

Proof. Denote the identification map implicit in (5.3) by $\mathfrak{I} : \ell^2(\mathbb{Z}^-; \mathscr{Y} \times \mathscr{U}) \to \ell^2(\mathbb{Z}^+; \mathscr{Y}' \times \mathscr{U}')$ (i.e. if we would identify $\mathscr{Y}$ and $\mathscr{U}$ with their respective duals then it is simpy the reflection). Then it is easily seen that $\mathfrak{I}^*\mathcal{R}^*\mathfrak{I}\mathcal{R} = I$. From Lemma 6.1 it follows that $\mathcal{R}\mathscr{G} = \mathscr{V}_d(0)^\perp$ so that $\mathfrak{I}\mathcal{R}\mathscr{G} = \ell^2(\mathbb{Z}^+; \mathscr{Y}' \times \mathscr{U}') \ominus \mathscr{V}_d(0)$. We conclude that $\mathfrak{I}^*\mathcal{R}^* \left(\ell^2(\mathbb{Z}^+; \mathscr{Y}' \times \mathscr{U}') \ominus \mathscr{V}_d(0)\right) = \mathscr{G}$ from which the result follows using that $\mathfrak{I}^*$ is an isomorphism from $\mathscr{G}'$ onto $\mathscr{G}$. ∎

LEMMA 6.3. *The input-output trajectory to final state map* $\mathcal{J}_c$ *for the node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *and the future minimizing operator* $\mathcal{I}_{f,d}$ *of its dual node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$ *are related by*

$$\mathcal{J}_c^* = \mathcal{R}^* \mathcal{I}_{f,d} \tag{6.1}$$

as unbounded operators $\mathscr{X}' \to \mathscr{G}'$.

Proof. The basic duality relationship (5.4) with $z^* \in \Xi_{f,d}$ and $\left[\begin{smallmatrix} y^* \\ u^* \end{smallmatrix}\right] = \mathcal{I}_{f,d} z^*$ gives for $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] \in \mathscr{G}_c$:

$$\left\langle \mathcal{J}_c \begin{bmatrix} y \\ u \end{bmatrix}, z^* \right\rangle_{\mathscr{X},\mathscr{X}'} = \left\langle \mathcal{R} \begin{bmatrix} y \\ u \end{bmatrix}, \mathcal{I}_{f,d} z^* \right\rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})}.$$

By Lemma 6.2 we have $\mathcal{R}^* \mathcal{I}_{f,d} z^* \in \mathscr{G}'$ so the above can be re-written as

$$\left\langle \mathcal{J}_c \begin{bmatrix} y \\ u \end{bmatrix}, z^* \right\rangle_{\mathscr{X},\mathscr{X}'} = \left\langle \begin{bmatrix} y \\ u \end{bmatrix}, \mathcal{R}^* \mathcal{I}_{f,d} z^* \right\rangle_{\mathscr{G},\mathscr{G}'},$$

which shows that $\mathcal{J}_c$ and $\mathcal{R}^* \mathcal{I}_{f,d}$ are adjoint to each other. So we only still need to show $D(\mathcal{I}_{f,d}) \supset D(\mathcal{J}_c^*)$. By definition

$$D(\mathcal{J}_c^*) = \left\{ z^* \in \mathscr{X}' : \exists \begin{bmatrix} y^\$ \\ u^\$ \end{bmatrix} \in \mathscr{G}' \text{ such that } \forall \begin{bmatrix} y \\ u \end{bmatrix} \in \mathscr{G}_c \right.$$
$$\left. \left\langle \mathcal{J}_c \begin{bmatrix} y \\ u \end{bmatrix}, z^* \right\rangle_{\mathscr{X},\mathscr{X}'} = \left\langle \begin{bmatrix} y \\ u \end{bmatrix}, \begin{bmatrix} y^\$ \\ u^\$ \end{bmatrix} \right\rangle_{\mathscr{G},\mathscr{G}'} \right\}.$$

Using Lemma 6.2 it follows that

$$D(\mathcal{J}_c^*) = \left\{ z^* \in \mathscr{X}' : \exists \begin{bmatrix} y^\# \\ u^\# \end{bmatrix} \in \ell^2(\mathbb{Z}^+; \mathscr{Y}' \times \mathscr{U}') \ominus \mathscr{V}_d(0) \text{ such that } \forall \begin{bmatrix} y \\ u \end{bmatrix} \in \mathscr{G}_c \right.$$
$$\left. \left\langle \mathcal{J}_c \begin{bmatrix} y \\ u \end{bmatrix}, z^* \right\rangle_{\mathscr{X},\mathscr{X}'} = \left\langle \mathcal{R} \begin{bmatrix} y \\ u \end{bmatrix}, \begin{bmatrix} y^\# \\ u^\# \end{bmatrix} \right\rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})} \right\}.$$

Using Lemma 5.1 it follows from the equality in the domain definition that $\left[\begin{smallmatrix} y^\# \\ u^\# \end{smallmatrix}\right]$ is a trajectory of the dual node for initial condition z^*. So $z^* \in D(\mathcal{J}_c^*)$ has finite cost and so $z^* \in D(\mathcal{I}_{f,d})$. Hence $\mathcal{J}_c^* = \mathcal{R}^* \mathcal{I}_{f,d}$. ∎

THEOREM 6.4. *The node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *satisfies the output coercive past cost condition if and only if its dual node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$ *satisfies the finite future incremental cost condition.*

Proof. The output coercive past cost condition implies that $C\mathcal{J}_c : \mathscr{G}_c \to \mathscr{Y}$ extends to a bounded operator $\mathscr{G} \to \mathscr{Y}$. By [7, Theorem 13.2] its adjoint equals $\mathcal{J}_c^* C^*$. It follows that $\mathcal{J}_c^* C^*$ is a bounded operator $\mathscr{Y}' \to \mathscr{G}'$. In particular the range of C^* is contained in the domain of $\mathcal{J}_c^*$, which by Lemma 6.3 equals $\Xi_{f,d}$. This is exactly the finite future incremental cost condition for the dual node.

Conversely, assume that the range of C^* is contained in $\Xi_{f,d}$. The operator $\mathcal{J}_c^* C^*$ is closed as it is the adjoint of the densely defined operator $C\mathcal{J}_c$ (we again use that $(C\mathcal{J}_c)^* = \mathcal{J}_c^* C^*$ by [7, Theorem 13.2]). By the range assumption $\mathcal{J}_c^* C^*$ is defined on all of $\mathscr{Y}'$, so that by the closed graph theorem it is a bounded operator $\mathscr{Y}' \to \mathscr{G}'$. Since $(C\mathcal{J}_c)^*$ is a bounded (and everywhere defined) operator, the operator $C\mathcal{J}_c$ is closable, and its closure is the bounded operator $((C\mathcal{J}_c)^*)^*$. Thus, $C\mathcal{J}_c$ extends to a bounded operator $\mathscr{G} \to \mathscr{Y}$. So the output coercive past cost condition holds. ∎

In the next proposition, we again use the weighted ℓ^2 space duality from (5.2).

LEMMA 6.5. *The map $\Gamma_p : \mathscr{G}_c \to \mathfrak{s}(\mathbb{Z}^+;\mathscr{Y})$ for the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and the map $\Gamma_{f,d} : \ell^2_c(\mathbb{Z}^-;\mathscr{Y}') \to \mathscr{V}_d(0)$ for the dual node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$ are related by*

$$\Gamma_p^* = \mathcal{R}^*\Gamma_{f,d}.$$

Proof. Using Lemma 6.3 we have

$$\Gamma_p^* = (\mathcal{C}\mathcal{J}_c)^* = \mathcal{J}_c^*\mathcal{C}^* = \mathcal{R}^*\mathcal{I}_{f,d}\mathcal{B}_d = \mathcal{R}^*\Gamma_{f,d}.$$

□

As the next lemma shows, the dual space of the state space of the completed $\mathcal{I}_p$-compression can be identified with the state space of the completed $\mathcal{I}_f$-compression of the dual node of the completed $\mathcal{I}_p$-compression.

LEMMA 6.6. *Any bounded linear functional on $\mathscr{X}_p$ can be identified with an element z^* of $(\mathscr{X}_p)_{f,d}$ through*

$$\langle z, z^*\rangle_{\mathscr{X}_p,\mathscr{X}_p'} = \left\langle \mathcal{I}_p^{\mathscr{X}_p} z, \mathcal{R}^*\mathcal{I}_{f,d}^{(\mathscr{X}_p)_{f,d}} z^*\right\rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})}. \tag{6.2}$$

This duality is with respect to the pivot space $\mathscr{X}$ in the sense that

$$\langle z, z^*\rangle_{\mathscr{X}_p,\mathscr{X}_p'} = \langle z, z^*\rangle_{\mathscr{X},\mathscr{X}'}$$

if $z \in \mathscr{X}_p \cap \mathscr{X}$ and $z^* \in (\mathscr{X}_p)_{f,d} \cap \mathscr{X}'$. Moreover, this duality is norm-preserving in the sense that $\|z^*\|_{(\mathscr{X}_p)_{f,d}}$ equals the $\mathscr{X}_p'$ norm of the corresponding functional.

Proof. It is easily seen that for a given $z^* \in (\mathscr{X}_p)_{f,d}$ the expression (6.2) defines a bounded linear functional on $\mathscr{X}_p$, so it remains the prove the converse. By the duality between $\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})$ and $\ell^2(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}')$ and the identification of $\mathscr{X}_p$ with $R(\mathcal{I}_p^{\mathscr{X}_p})$ it follows that any linear functional on $\mathscr{X}_p$ must be of the form

$$\left\langle \mathcal{I}_p^{\mathscr{X}_p} z, v\right\rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})}$$

for some $v \in \ell^2(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}')$. We have $R(\mathcal{I}_p^{\mathscr{X}_p}) = D(\mathcal{J}_r^{\mathscr{X}_p}) = D(\mathcal{J}^{\mathscr{X}_p}) \ominus N(\mathcal{J}^{\mathscr{X}_p})$ so we may assume that $v \in N(\mathcal{J}^{\mathscr{X}_p})^\perp$. So $v \in R(\mathcal{J}^{\mathscr{X}_p *})$ since $\mathcal{J}^{\mathscr{X}_p}$ as a partial isometry has closed range. Applying Lemma 6.3 to the completed $\mathcal{I}_p$-compression (and using that $\mathcal{J}_c^* = \mathcal{J}^*$) then gives the result.

That the duality is with respect to $\mathscr{X}$ follows from (5.4) with $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right] = \mathcal{I}_p^{\mathscr{X}_p} z$ and $\left[\begin{smallmatrix} y^* \\ u^* \end{smallmatrix}\right] = \mathcal{I}_{f,d}^{(\mathscr{X}_p)_{f,d}} z^*$. The norm preservation follows from the fact that $\mathcal{I}_{f,d}^{(\mathscr{X}_p)_{f,d}}$, $\mathcal{R}$ and $\mathcal{I}_p^{\mathscr{X}_p}$ are isometries. □

LEMMA 6.7. *The operator $\mathcal{J}^{\mathscr{X}_p}$ is a co-isometry.*

Proof. According to Lemma 6.3 this is equivalent to showing that $\mathcal{I}_{f,d} : \mathscr{X}_p' \to \ell^2(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}')$ is an isometry. By the identification of the dual space in Lemma 6.6 this in turn is equivalent to $\mathcal{I}_{f,d} : (\mathscr{X}_p)_{f,d} \to \ell^2(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}')$ being an isometry, which is true by definition of the norm in $(\mathscr{X}_p)_{f,d}$. □

The filter Riccati equation of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is simply the control Riccati equation of the dual node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]^*$.

DEFINITION 6.8. *The triple (p, r, T) is called a (nonnegative) solution of the* filter Riccati equation *of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ if*

1. *p is a closed nonnegative symmetric sesquilinear form in $\mathscr{X}$ whose domain satisfies $A^*D(p) \subset D(p)$, $C^*\mathscr{Y} \subset D(p)$.*
2. *r is a bounded nonnegative symmetric sesquilinear form on $\mathscr{Y}$.*
3. *$T : D(p) \to \mathscr{Y}$ is a linear operator.*
4. *For all $z \in D(p)$, $y \in \mathscr{Y}$ we have*

$$p(A^*z+C^*y, A^*z+C^*y)+\|B^*z+D^*y\|^2_{\mathscr{U}}+\|y\|^2_{\mathscr{Y}} = p(z,z)+r(Tz-y,Tz-y). \tag{6.3}$$

The solution is called classical *when $D(p) = \mathscr{X}$.*

DEFINITION 6.9. *Let $\mathsf{G} : D(\mathsf{G}) \subset \mathbb{C} \to \mathcal{L}(\mathscr{U},\mathscr{Y})$ be holomorphic at the origin. A function $[\tilde{\mathsf{M}}, \tilde{\mathsf{N}}] \in H^\infty(\mathbb{D}, \mathcal{L}(\mathscr{Y} \times \mathscr{U}, \mathscr{Y}))$ is called a* left factorization *of G if $\tilde{\mathsf{M}}(z)$ is invertible for all z in a neighbourhood of the origin and $\mathsf{G}(z) = \tilde{\mathsf{M}}(z)^{-1}\tilde{\mathsf{N}}(z)$ in a neighbourhood of the origin.*

THEOREM 6.10. *Let $\mathsf{G} : D(\mathsf{G}) \subset \mathbb{C} \to \mathcal{L}(\mathscr{U},\mathscr{Y})$ be holomorphic at the origin and let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ be realization of G. The following are equivalent conditions.*

- *$\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ satisfies the output coercive past cost condition.*
- *The filter Riccati equation of $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ has a (nonnegative self-adjoint) solution.*
- *G has a left factorization.*

Under these equivalent conditions, the filter Riccati equation of $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ has a smallest (nonnegative self-adjoint) solution (p_p, T_p, r_p) with domain $\Xi_{f,d}$.

Proof. According to Theorem 6.4 the output coercive past cost condition implies that the dual node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ satisfies the finite future incremental cost condition. Theorem 4.6 then shows that the dual node has a solution to its control Riccati equation. This implies that the original node has a solution to its filter Riccati equation.

If the node has a solution to its filter Riccati equation, then the dual node has a solution to its control Riccati equation. It follows from Theorem 4.6 that the transfer function of the dual node has a right factorization: $\mathsf{G}_d(z) = \mathsf{N}(z)\mathsf{M}(z)^{-1}$. Realizing that the transfer function of the original node G and that of the dual node G_d are related by $\mathsf{G}_d(z) = \mathsf{G}(\bar{z})^*$ we obtain $\mathsf{G}(z) = \tilde{\mathsf{M}}(z)^{-1}\tilde{\mathsf{N}}(z)$ with $\tilde{\mathsf{M}}(z) := \mathsf{M}(\bar{z})^*$ and $\tilde{\mathsf{N}}(z) := \mathsf{N}(\bar{z})^*$. So the transfer function of the original node has a left factorization.

Assuming that the transfer function of the original node has a left factorization, it is easily seen as above that the transfer function of the dual node has a right factorization. It follows from Theorem 4.6 that the dual node satisfies the finite future incremental cost condition. Theorem 6.4 then shows that the original node satisfies the output coercive past cost condition.

Existence of the smallest solution follows from the existence of the smallest solution (q_f, K_f, s_f) of the control Riccati equation of the dual node. □

7. The optimal output injection problem. We consider the closed-loop system

$$\begin{aligned}
x_{n+1} &= (A - HC)x_n + (B - HD)u_n + Hy_n, && n \in \mathbb{Z}^-,\\
w_n &= WCx_n + WDu_n - Wy_n, && n \in \mathbb{Z}^-,\\
x_0 &= z,\\
\exists N \in \mathbb{Z}^- &: x_n = 0 = u_n \quad \forall n \le -N.
\end{aligned} \tag{7.1}$$

where $W : \mathscr{Y} \to \mathscr{H}$ is a given bounded linear operator and $H : \mathscr{Y} \to \mathscr{X}_e$ is a bounded linear operator. Here $\mathscr{H}$ is a Hilbert space and $\mathscr{X}_e$ is a Hilbert space that contains $P_{\mathscr{X} \ominus \mathscr{N}}\Xi_p$ as a dense subspace and is such that the restriction of the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ to $P_{\mathscr{X} \ominus \mathscr{N}}\Xi_p \times \mathscr{U}$ extends continuously to a bounded linear operator from $\mathscr{X}_e \times \mathscr{U}$ to $\mathscr{X}_e \times \mathscr{Y}$.

THEOREM 7.1. *Assume that* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *satisfies the output coercive past cost condition. Define* $H_p : \mathscr{Y} \to \mathscr{X}_p$ *by*

$$H_p y = \mathcal{J}^{\mathscr{X}_p} P_{\mathscr{G}} g, \tag{7.2}$$

where $g_{-1} = \left[\begin{smallmatrix} y \\ 0 \end{smallmatrix}\right]$ *and* $g_n = 0$ *for* $n < -1$ *and* $P_{\mathscr{G}}$ *is the orthogonal projection* $\ell^2(\mathbb{Z}^-; \mathscr{Y} \times \mathscr{U}) \to \mathscr{G}$.

Then H_p *minimizes both the* $\mathcal{L}(\ell^2(\mathbb{Z}^-, \mathscr{Y} \times \mathscr{U}), \ell^\infty(\mathbb{Z}^-, \mathscr{H}))$ *and the* $\mathcal{L}(\ell^2(\mathbb{Z}^-, \mathscr{Y} \times \mathscr{U}), \ell^2(\mathbb{Z}^-, \mathscr{H}))$ *norm of the map from* $\{\left[\begin{smallmatrix} y_n \\ u_n \end{smallmatrix}\right]\}_{n\in\mathbb{Z}^-}$ *to* $\{w_n\}_{n\in\mathbb{Z}^-}$ *in* (7.1)*, where* H *ranges over all linear maps* $\mathscr{Y} \to \mathscr{X}_e$ *with* $\mathscr{X}_e$ *a Hilbert space that contains* $P_{\mathscr{X}\ominus\mathscr{N}}\Xi_p$ *as a dense subspace and is such the restriction of the node* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *to* $P_{\mathscr{X}\ominus\mathscr{N}}\Xi_p \times \mathscr{U}$ *extends continuously to a node with* $\mathscr{X}_e$ *as state space. The operator* H_p *also minimizes the* $\mathcal{L}(\ell^2(\mathbb{Z}^-, \mathscr{Y} \times \mathscr{U}), \mathscr{H})$ *norm of the map* $\{\left[\begin{smallmatrix} y_n \\ u_n \end{smallmatrix}\right]\}_{n\in\mathbb{Z}^-} \mapsto w_{-1}$.

Denote the smallest nonnegative self-adjoint solution of the filter Riccati equation of $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *by* (p_p, T_p, r_p). *Then the minimum norms mentioned above all equal the square root of* $\sup_{\|h\|=1} r_p(W^*h, W^*h)$.

Proof. The assumption that $P_{\mathscr{X}\ominus\mathscr{N}}\Xi_p \subset \mathscr{X}_e$ implies that $P_{\mathscr{X}\ominus\mathscr{N}}\Xi_- \subset \mathscr{X}_e$ and so ensures that the extended node has the same transfer function as the original node. In particular, it too satisfies the output coercive past cost condition (this follows from Theorem 6.10). So the dual of the extended node satisfies the finite future incremental cost condition.

We have that $K := -H^*$ is a bounded operator from $\mathscr{X}_e'$ to $\mathscr{Y}'$ and $E := W^*$ is a bounded operator from $\mathscr{H}'$ to $\mathscr{Y}'$. Obviously, the domain of K (which equals the whole state space $\mathscr{X}_e'$) is A^* invariant and contains the range of C^*. The adjoint of the closed-loop system (7.1) considered as a dynamical system on $\mathbb{Z}^+$ then is

$$\begin{aligned} x^d_{n+1} &= (A^* + C^*K)x^d_n + C^*Ew^*_n, && n \in \mathbb{Z}^+, \\ u^*_n &= (B^* + D^*K)x^d_n + D^*Ew^*_n, && n \in \mathbb{Z}^+, \\ -y^*_n &= Kx^d_n + Ew^*_n, && n \in \mathbb{Z}^+, \\ x^d_0 &= z^*, \end{aligned} \tag{7.3}$$

i.e. it is the closed-loop system (4.2) of the adjoint node of the extended node with u replaced by $-u$.

From the above it follows that the search over H in the optimal output injection problem translates to the search over K in the optimal feedback problem.

By the discussion in Section 5, the input-output maps of (7.1) and (7.3) are adjoints when both are considered $\ell^2 \to \ell^2$ and also when considered $\ell^1 \to \ell^2$ and $\ell^2 \to \ell^\infty$ respectively. Also the maps $\{\left[\begin{smallmatrix} y_n \\ u_n \end{smallmatrix}\right]\}_{n\in\mathbb{Z}^-} \mapsto w_{-1}$ and $w^*_0 \mapsto \left\{\left[\begin{smallmatrix} y^*_n \\ u^*_n \end{smallmatrix}\right]\right\}_{n\in\mathbb{Z}^+}$ are adjoints.

From the fact that the operator norm of an operator equals that of its adjoint and Theorem 4.7, it follows that the square root of $\sup_{\|h\|=1} r_p(W^*h, W^*h)$ is a lower-bound for all three operator norms considered and that this lower-bound is reached for $H = -K^*_{f,d}$. The formula for H_p follows once we show that $H_p = -K^*_{f,d}$ as operators $\mathscr{Y} \to \mathscr{X}_p$, which we now obtain.

We have, for $y \in \mathscr{Y}$ and $z^* \in (\mathscr{X}_p)_{f,d} = \mathscr{X}_p'$,

$$\langle H_p y, z^* \rangle_{\mathscr{X}_p, \mathscr{X}_p'} = \langle \mathcal{J}^{\mathscr{X}_p} P_{\mathscr{G}} g, z^* \rangle_{\mathscr{X}_p, \mathscr{X}_p'},$$

which by Lemma 6.6 equals

$$\left\langle \mathcal{I}_p^{\mathscr{X}_p} \mathcal{J}^{\mathscr{X}_p} P_{\mathscr{G}} g, \mathcal{R}^* \mathcal{I}_{f,d}^{(\mathscr{X}_p)_{f,d}} z^* \right\rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})}.$$

By (3.1) the above equals

$$\left\langle P_{\ell^2(\mathbb{Z}^-;\mathscr{Y}\times\mathscr{U})\ominus N(\mathcal{J}^{\mathscr{X}_p})} P_{\mathscr{G}} g, \mathcal{R}^* \mathcal{I}_{f,d}^{(\mathscr{X}_p)_{f,d}} z^* \right\rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})}.$$

By Lemma 6.3 we can omit the projection onto the orthogonal complement of the kernel in this last formula. By Lemma 6.2 and Remark 5.2 we have $\mathcal{R}^*\mathcal{I}_{f,d}^{(\mathscr{X}_p)_{f,d}} z^* \in \ell^2(\mathbb{Z}^+;\mathscr{Y}'\times\mathscr{U}')\ominus\mathscr{G}^\perp$ so that the projection onto $\mathscr{G}$ may also be omitted. Hence

$$\langle H_p y, z^* \rangle_{\mathscr{X}_p,\mathscr{X}_p'} = \langle \mathcal{R} g, \mathcal{I}_{f,d}^{(\mathscr{X}_p)_{f,d}} z^* \rangle_{\ell^2(\mathscr{Y}\times\mathscr{U})} = -\langle y, K_{f,d} z^* \rangle_{\mathscr{Y},\mathscr{Y}'},$$

where the last equality holds by definition of $K_{f,d}$ and $\mathcal{R}$. This proves H_p given by (7.2) is indeed the optimal output injection. □

The formula (7.2) shows that the optimal output injection $H_p y$ for the node $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is the final state of the dynamical system associated to that node for some input defined on $\mathbb{Z}^-$ in terms of y.

THEOREM 7.2. *Assume that* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ *satisfies the output coercive past cost condition. Then the closed-loop system of the completed* $\mathcal{I}_p$*-compression with the optimal output injection is strongly internally* $*$*-stable.*

Proof. The proof of Theorem 7.1 shows that the dual of the closed-loop system of the completed $\mathcal{I}_p$-compression with the optimal output injection is the closed-loop system of a completed $\mathcal{I}_f$-compression with the optimal state feedback. From Remark 4.8 it follows that this latter system is strongly internally stable. The result immediately follows. □

Remark 7.3. *Many of the operators defined here are closely related to analogous operators introduced in [2]. More precisely, the completed* $\mathcal{I}_p$*-compression is a passive observable and backward conservative i/s/o system if we equip its output space* $\mathscr{Y}$ *with the (equivalent) inner product induced by the quadratic form* r *in Definition 6.8 and use* $\mathscr{U}\times\mathscr{Y}$ *as the input space. With regard to this system, the co-isometry* $\mathcal{J}^{\mathscr{X}_p}$ *coincides with the input map* $\mathfrak{B}_{\Sigma_{i/s/o}}$ *in [2, Section 10], the isometry* $\mathcal{I}_{f,d}$ *is the output map of the adjoint system, and the two Hankel operators* Γ_p *and* $\Gamma_{f,d}$ *can be interpreted as compressions of the past/future map* Γ_Σ *in [2] and its adjoint.*

REFERENCES

[1] Damir Z. Arov and Olof J. Staffans. State/signal linear time-invariant systems theory: passive discrete time systems. *Internat. J. Robust Nonlinear Control*, 17(5-6):497–548, 2007.

[2] Damir Z. Arov and Olof J. Staffans. Two canonical passive state/signal shift realizations of passive discrete time behaviors. *J. Funct. Anal.*, to appear, 2009.

[3] Michael Green and David J.N. Limebeer. *Linear robust control*. Prentice Hall, Englewood Cliffs, New Jersey, 1995.

[4] Erwin Kreyszig. *Introductory functional analysis with applications.* John Wiley & Sons, New York-London-Sydney, 1978.

[5] Mark R. Opmeer and Olof J. Staffans. Optimal dynamic stabilization of infinite-dimensional discrete time-invariant linear systems. *in preparation*.

[6] 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.

[7] Walter Rudin. *Functional analysis.* McGraw-Hill Book Co., New York, 1973. McGraw-Hill Series in Higher Mathematics.