

Embedding of Time-Varying Contractive Systems in Lossless Realizations

Alle-Jan van der Veen and Patrick Dewilde

Delft University of Technology, Department of Electrical Engineering
2628 CD Delft, The Netherlands
email: allejan@dutentb.et.tudelft.nl; fax: +31 15 623271

The lossless embedding problem, also known as the Darlington synthesis or unitary extension problem, considers the extension of a given contractive system to become the partial input-output operator of a lossless system. In the paper, the embedding problem is solved for discrete-time time-varying systems with finite but possibly time-varying state dimensions, for the strictly contractive as well as the boundary case. The construction is done in a state space context and gives rise to a time-varying Riccati difference equation which is shown to have a closed-form solution. As a corollary, a discrete-time Bounded Real Lemma is formulated, linking contractiveness of an input-output operator to conditions on its state realization.

Key words: Lossless embedding, linear time-varying systems, Riccati difference equation, bounded real lemma

1. INTRODUCTION

In a Hilbert space setting, a bounded discrete-time linear time-varying system is specified by its input-output mapping: a bounded operator $T : \ell_2^M \rightarrow \ell_2^N$, where ℓ_2^M and ℓ_2^N are certain generalized ℓ_2 sequences. The lossless embedding problem which we study is, given a causal input-output operator T , to find a minimal extension of this system by adding more inputs and outputs to it such that the resulting system Σ ,

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

is lossless: $\Sigma^* \Sigma = I$, $\Sigma \Sigma^* = I$, and has T as its partial input-output operator when the extra inputs are forced to zero: $T = \Sigma_{11}$. The extension should be minimal in the sense that only a minimal number of inputs and outputs are added, and also the degree of the resulting system should not be increased. The solution of the embedding problem inherently involves the (spectral) factorization of $\Sigma_{21}^* \Sigma_{21} = I - T^* T$ and of $\Sigma_{12} \Sigma_{12}^* = I - T T^*$. Hence, a necessary condition for the existence of a lossless embedding is that T is a contractive system: $\|T\| \leq 1$.

We will solve the lossless embedding problem for contractive time-varying systems in a state space context, under the assumption that the number of states of T is finite at any point in time. While it is clear that contractivity is a necessary condition for the existence of an embedding, we will show in the sequel that strict contractivity is also sufficient to construct a solution when T is of locally finite degree. This result has been reported in condensed form in [VD2]. If T is contractive but not strictly contractive, then we need an extra condition to be satisfied in order to construct an embedding: the range of the Hankel operator associated with T should be closed. Not all systems have this property. This is reminiscent of the LTI infinite-dimensional case, where it is known that not all contractive systems have an embedding, see [D2].

⁰This research was supported in part by the commission of the EC under the ESPRIT BRA program 6632 (NANA2). Submitted to Math. Control Signals Systems, 3 July 1992. Resubmitted 24 November 1993, accepted 15 March 1994, revised 26 September 1994. First author's present address (till November 1994): Stanford University, Dept. Computer Science/SCCM, Stanford, Palo Alto CA 94305-2140, tel. (415) 723-4269, fax (415) 723 2411, email allejan@sccm.stanford.edu.

The lossless embedding problem is known under various names in a number of fields: in mathematics as inner dilations or as the unitary extension problem, in control as the Bounded Real Lemma (BRL), and it is largely equivalent to the computation of a spectral factor of $(I - T^*T)$. There are many applications of this problem. In network theory, the classical application is Darlington synthesis, which provides a structured way to construct a realization of a passive system using lossless components only. Once Σ is obtained, it can be factored into various kinds of “ladder” or “lattice” cascade realizations consisting of lossless degree-1 sections [VD3], similar to [DD1, D1] for time-invariant systems. This provides one of the most stable types of realization of transfer functions, with respect to parameter variations and noise sensitivity. Particular applications could be switched multi-rate filter banks, where the switches provide a time-varying state dimension, and implementations of non-uniformly sampled systems. In control, the BRL relates the contractivity of a system to properties of its state space realization [AV, AHD, V1], which has applications in \mathcal{H}_∞ optimal control and sensitivity minimization. There is a close connection between the BRL and certain Riccati equations, and there is growing interest in extensions of results to time-varying systems. Reference material on Riccati equations is contained in the book [BLW].

In the present paper, the classical time-invariant theory is extended to the time-varying context. We describe a causal linear time-varying system by a bounded upper operator

$$T = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots \\ & T_{-1,-1} & T_{-1,0} & T_{-1,1} & \cdots \\ & & \boxed{T_{00}} & T_{01} & \cdots \\ & \mathbf{0} & & T_{11} & \cdots \\ & & & & \ddots \end{bmatrix}$$

mapping input sequences in ℓ_2 , $u = [\cdots u_{-1} u_0 u_1 \cdots]$, to corresponding output sequences y via $y = uT$. The i -th row of T contains the impulse response of the system for an impulse at time i ; causality implies that the impulse response is zero before time i , hence T is upper triangular. The entries T_{ij} of T are matrices; the number of rows of T_{ij} corresponds to the number of inputs of the system at time instant i , while the number of columns of T_{ij} is equal to the number of outputs of the system at time instant j . These numbers are not necessarily constant. In fact, we will show in this paper that the concept of time-varying state dimensions (which is necessary for minimal realizations) forces one to adopt the concept of time-varying input-output dimensions, too. A physical interpretation is that the system contains switches that switch on or off certain inputs, outputs, or states at certain times. Time-varying input-output dimensions also occur in multirate sampled data systems, and in time-varying Hankel-norm model reduction [DV]. With the concept of time-varying input and output dimensions, it is also possible to incorporate finite upper triangular matrices T into the framework of time-varying systems, by choosing the dimensions of T_{ij} to be zero for i, j outside a certain interval. This leads to new computational algorithms for certain types of linear algebra problems [VD3, VD4].

We assume that the upper input-output operators T admit a state space realization of the form

$$\begin{aligned} x_{i+1} &= x_i A_i + u_i B_i \\ y_i &= x_i C_i + u_i D_i \end{aligned} \quad \mathbf{T}_i = \begin{bmatrix} A_i & C_i \\ B_i & D_i \end{bmatrix} \quad (1.1)$$

in which the matrices $\{A_i, B_i, C_i, D_i\}$ are uniformly bounded and have finite (but not necessarily constant) dimensions. For a given system T with realization $\{\mathbf{T}_i\}$, we seek to determine a lossless embedding system Σ with state space realization $\{\Sigma_i\}$ of the form

$$\Sigma_i = \begin{bmatrix} R_i & & \\ & I & \\ & & I \end{bmatrix} \begin{bmatrix} A_i & C_i \\ B_i & D_i \\ \hline B_{2,i} & D_{21,i} & D_{22,i} \end{bmatrix} \begin{bmatrix} C_{2,i} \\ D_{12,i} \\ D_{22,i} \end{bmatrix} \begin{bmatrix} R_{i+1}^{-1} & & \\ & I & \\ & & I \end{bmatrix} \quad (1.2)$$

Σ_i contains the given realization \mathbf{T}_i , suitably state space transformed by some boundedly invertible R_i , so that Σ_{11} is equal to the given input-output operator T . Σ is extended by matrices $B_2, C_2, D_{21}, D_{12}, D_{22}$ corresponding to the

secondary inputs and outputs. It can be shown that a system is lossless if its realization is unitary at all times, and hence we immediately obtain a lossless embedding Σ of T if we require $\Sigma_i \Sigma_i^* = I$, $\Sigma_i^* \Sigma_i = I$. The embedding problem thus reduces to the problem of finding the state transformations R_i and the embedding matrices B_2, C_2, D_{21}, D_{12} in (1.2) such that Σ is unitary. The induced set of orthonormality conditions gives rise to a set of equations whose solution depends at each time instant i on the positivity of a matrix $M_i = R_i^* R_i$, which satisfies the recursive Riccati equation

$$M_{i+1} = A_i^* M_i A_i + B_i^* B_i + [A_i^* M_i C_i + B_i^* D_i] (I - D_i^* D_i - C_i^* M_i C_i)^{-1} [D_i^* B_i + C_i^* M_i A_i]. \quad (1.3)$$

This Riccati equation is similar to that which is obtained in optimal control problems. If T is strictly contractive, then (1.3) has a positive semidefinite solution $\{M_i\}$, which can be specified explicitly in terms of the operator T and the controllability operator of the given realization. The connection is obtained by elaborating on the following idea: partition T as

$$T = \left[\begin{array}{c|cc} \ddots & \vdots & \vdots \\ & T_{i-1,i-1} & T_{i-1,i} \cdots \\ \hline & 0 & T_i \cdots \\ & & \ddots \end{array} \right] = \begin{bmatrix} K_i & H_i \\ 0 & E_i \end{bmatrix} \quad (1.4)$$

T is contractive only if K_i is contractive, for all i in turn. Given the contractiveness of K_i for some i , the conditions for having K_{i+1} contractive can be specified in terms of K_i and the new column of K_{i+1} . When we assume a state realization for T , then it turns out that we can define a finite matrix M_i in terms of K_i and the controllability operator of the realization, such that the contractiveness of K_i (or positivity of $I - K_i^* K_i$) is summarized by having $M_k \geq 0$ for all $k \leq i$. Given the contractivity of K_i , the extra conditions to ensure that K_{i+1} is also contractive is summarized by having $M_{i+1} \geq 0$, where M_{i+1} is obtained by the Riccati recursion (1.3). The explicit solution for M can be used to give simple derivations of properties of the Riccati recursion.

We also consider the (mathematically complicated) boundary case where T is contractive, but not necessarily in the strict sense. In that case, the inverted term in (1.3) is not necessarily bounded and the inverse has to be replaced by a pseudo-inverse. Under the sufficient condition that the realization is uniformly observable, we show that the modified recursion has a hermitian, positive, bounded solution $\{M_i\}$, which we also give in closed form.

Section 2 introduces a convenient diagonal algebra notation for time-varying systems, which is used in section 3 to define a diagonal operator form for K_i and H_i . The recursive equation for M is derived in sections 4 and 5, for the strictly contractive and the boundary case, respectively. These results are used in section 6 to solve the embedding problem.

2. DIAGONAL ALGEBRA NOTATION

Expressions in time-varying state space theory quickly lead to an abundance of time indices. This can be avoided by collecting state space quantities A_i etc. into diagonals. The resulting ‘diagonal algebra’ was introduced in [AD] and refined in [ADD, VD1, DD2, DV]. We adopt the notation from the latter paper.

2.1. Spaces

Our theory will take place in spaces of *non-uniform* ℓ_2 series. The sequence $N = [N_i]_{i \in \mathbf{Z}}$ ($N_i \in \mathbf{N}$) is called an index sequence. Using N , signals live in the space of non-uniform sequences

$$\mathcal{N} = \cdots \oplus \mathcal{N}_{-1} \oplus \boxed{\mathcal{N}_0} \oplus \mathcal{N}_1 \oplus \mathcal{N}_2 \oplus \cdots \in \mathbf{C}^N,$$

where $\mathcal{N}_i \in \mathbf{C}^{N_i}$. (The box denotes the position of the 0-th component.) We write $N = \#(\mathcal{N})$. The space ℓ_2^N imposes an ℓ_2 norm on the series:

$$\ell_2^N = \{x \in \mathcal{N}, \|x\|_2 < \infty\}.$$

We will think of vectors in the \mathcal{N}_j 's to be row vectors, and (row) sequences in \mathcal{N} to have entries in \mathcal{N}_j . Thus, the action of matrices or operators happens at the right of the argument, as in aA , which is the result of the application of the operator A to the sequence a . Finite vectors can be covered in many ways by taking the dimensions of all except a finite number of the \mathcal{N}_j 's equal to zero. The following classes of bounded operators $\ell_2^{\mathcal{M}} \rightarrow \ell_2^{\mathcal{N}}$ are defined:

$\mathcal{X}(\mathcal{M}, \mathcal{N})$: the space of bounded operators $\ell_2^{\mathcal{M}} \rightarrow \ell_2^{\mathcal{N}}$. An operator $T \in \mathcal{X}(\mathcal{M}, \mathcal{N})$ has a matrix representation $T = [T_{ij}]_{-\infty}^{\infty}$, with $T_{ij} \in \mathcal{M}_i \times \mathcal{N}_j$. We usually drop the indexing spaces as relaxation of notation.

$\mathcal{U}(\mathcal{M}, \mathcal{N})$, $\mathcal{L}(\mathcal{M}, \mathcal{N})$, $\mathcal{D}(\mathcal{M}, \mathcal{N})$: the space of bounded upper triangular, lower triangular, and diagonal operators $\ell_2^{\mathcal{M}} \rightarrow \ell_2^{\mathcal{N}}$, respectively.

Our input-output operators T will be operators in \mathcal{X} . Causal input-output operators are in addition upper. Next, shifted spaces and the shift operator are defined:

$\mathcal{N}^{(k)}$: the k -th shift rightwards in the series of spaces, as in $\mathcal{N}^{(1)} = [\cdots \mathcal{N}_{-2} \boxed{\mathcal{N}_{-1}} \mathcal{N}_0, \cdots]$.

Z : the causal bilateral shift operator $Z: \ell_2^{\mathcal{N}} \rightarrow \ell_2^{\mathcal{N}^{(1)}}$, defined by $(xZ)_i = x_{i-1}$.

$A^{(k)}$: the ‘diagonal’ shift of an operator $A \in \mathcal{X}$ in the South-East direction: $A^{(k)} = (Z^k)^* A Z^k$.

$T_{[k]} \in \mathcal{D}(\mathcal{M}^{(k)}, \mathcal{N})$: the k -th diagonal above the main (0-th) diagonal of an operator $T \in \mathcal{U}(\mathcal{M}, \mathcal{N})$. T can be formally decomposed into a sum of shifted diagonal operators as in $T = \sum_{k=0}^{\infty} Z^k T_{[k]}$.

2.2. Realizations

Time-varying state realizations of the type (1.1) can be rewritten in global operator form by assembling the matrices $\{A_i\}$, $\{B_i\}$ etc. into *diagonal operators* $A = \text{diag}(A_i)$, etc., acting on sequences $u = [\cdots u_0 u_1 \cdots] \in \ell_2^{\mathcal{M}}$, $y = [\cdots y_0 y_1 \cdots] \in \ell_2^{\mathcal{N}}$, $x = [\cdots x_0 x_1 \cdots] \in \ell_2^{\mathcal{B}}$. Together, A, B, C, D define a realization \mathbf{T} of T as

$$\begin{aligned} xZ^{-1} &= xA + uB \\ y &= xC + uD \end{aligned} \quad \mathbf{T} = \begin{bmatrix} A & C \\ B & D \end{bmatrix}: (\mathcal{B} \times \mathcal{M}) \rightarrow (\mathcal{B}^{(-1)} \times \mathcal{N}). \quad (2.1)$$

This definition constitutes the same set of time-varying state equations as in (1.1), but now written in an index-free form and acting on sequences. \mathbf{T} is a realization of T if a solution of (2.1) satisfies $uT = y$, that is, if

$$T_{[k]} = \begin{cases} 0, & k < 0 \\ D, & k = 0 \\ B^{(k)} A^{(k-1)} \cdots A^{(1)} C, & k > 0 \end{cases} \quad \text{or} \quad T_{ij} = \begin{cases} 0, & i > j \\ D_i, & i = j \\ B_i A_{i+1} \cdots A_{j-1} C_j, & i < j. \end{cases} \quad (2.2)$$

An important notion in this context is strict stability of a realization. Let ℓ_A be the spectral radius of the operator AZ : $\ell_A = \lim_{n \rightarrow \infty} \|(AZ)^n\|^{1/n}$. We shall say that the realization (2.1) is strictly stable if $\ell_A < 1$.¹ In that case, the operator $(I - AZ)^{-1}$ exists as a bounded operator and by substitution in (2.1) one obtains

$$T = D + BZ(I - AZ)^{-1}C. \quad (2.3)$$

If an upper operator has a state space realization with state space sequences \mathcal{B} where each \mathcal{B}_i has finite dimension, then we shall say that the operator is *locally finite*. We will assume throughout the paper that all input-output operators have this property.

An equivalent realization is found by applying a state transformation $\hat{x} = xR$ on the state sequence x of the system, where $R \in \mathcal{D}(\mathcal{B}, \mathcal{B})$ is a bounded and boundedly invertible diagonal operator. The realization \mathbf{T} is then transformed to

$$\mathbf{T}' = \begin{bmatrix} R & \\ & I \end{bmatrix} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} [R^{(-1)}]^{-1} & \\ & I \end{bmatrix}.$$

¹Since $(AZ)^n = Z^n A^{(n)} \cdots A^{(2)} A^{(1)}$, we have $\ell_A = \lim_{n \rightarrow \infty} \|A^{(n)} \cdots A^{(2)} A^{(1)}\|^{1/n} = \lim_{n \rightarrow \infty} \sup_k \|A_{k-n} \cdots A_{k-2} A_{k-1}\|^{1/n}$, which leads to the more usual definition of exponential stability for time-varying realizations.

It is easy to see that $\ell_{RA[R^{(-1)}]^{-1}} = \ell_A$, hence that strict stability is preserved under the transformation.

In the question whether there exist state transformations such that the resulting realization is in input normal form ($A^*A + B^*B = I$) or output normal form ($AA^* + CC^* = I$), the following Lyapunov equations arise:

$$A^* \Lambda_c A + B^* B = \Lambda_c^{(-1)} \quad \text{resp.} \quad A \Lambda_o^{(-1)} A^* + CC^* = \Lambda_o.$$

For strictly stable realizations ($\ell_A < 1$), bounded solutions exist and are unique, $\Lambda_c \geq 0$, $\Lambda_o \geq 0$, and are given by the uniformly converging sums

$$\begin{aligned} \Lambda_c &= \sum_{k=0}^{\infty} (A^{\{k\}})^* (B^* B)^{(k+1)} A^{\{k\}}, & \text{where } A^{\{k\}} &= A^{(k)} \dots A^{(1)}, A^{\{0\}} = I, \\ \Lambda_o &= \sum_{k=0}^{\infty} (A^{\{k\}})^* (B^* B)^{(k+1)} A^{\{k\}}, & \text{where } A^{\{k\}} &= A \dots A^{(-k+1)}, A^{\{0\}} = I. \end{aligned} \quad (2.4)$$

Λ_c is called the controllability Gramian, Λ_o is called the observability Gramian. A realization is called controllable if $\Lambda_c > 0$ and uniformly controllable if Λ_c is uniformly positive definite, $\Lambda_c \gg 0$, so that it is invertible. Similarly, a realization is called observable if $\Lambda_o > 0$ and uniformly observable if $\Lambda_o \gg 0$. A realization is minimal if it is both controllable and observable. If T admits a locally finite realization, then it is always possible to choose this realization to be either uniformly controllable or uniformly observable, although it may not be possible to have both [VD1, V2]. A uniformly controllable realization can be transformed into a realization in input normal form by a state transformation R given by $\Lambda_c = R^* R$, and a uniformly observable realization can be transformed to output normal form by choosing R a factor of Λ_o^{-1} .

If a strictly stable realization \mathbf{T} is unitary: $\mathbf{T}^* \mathbf{T} = I$, $\mathbf{T} \mathbf{T}^* = I$, then the corresponding input-output operator $T \in \mathcal{U}$ is lossless (or inner): $T^* T = I$, $T T^* = I$. A slightly more general version of this, not using normalized realizations, is given by the following lemma:

Lemma 2.1. *Let $T \in \mathcal{U}$ be an input-output operator with strictly stable state realization \mathbf{T} . Then*

$$\begin{aligned} \exists M \in \mathcal{D}: \quad \mathbf{T}^* \begin{bmatrix} M & \\ & I \end{bmatrix} \mathbf{T} &= \begin{bmatrix} M^{(-1)} & \\ & I \end{bmatrix} & \Rightarrow & T^* T = I \\ \exists M \in \mathcal{D}: \quad \mathbf{T} \begin{bmatrix} M^{(-1)} & \\ & I \end{bmatrix} \mathbf{T}^* &= \begin{bmatrix} M & \\ & I \end{bmatrix} & \Rightarrow & T T^* = I \end{aligned}$$

PROOF In the first relation to prove, the assumption is that there is an $M \in \mathcal{D}$ such that

$$A^* M A + B^* B = M^{(-1)}; \quad A^* M C + B^* D = 0; \quad C^* M C + D^* D = I.$$

Note that the first equation is the Lyapunov equation, so that M is in fact the controllability Gramian Λ_c of \mathbf{T} . Using these equations, we derive that

$$\begin{aligned} I - T^* T &= I - \left[D + BZ(I - AZ)^{-1} C \right]^* \left[D + BZ(I - AZ)^{-1} C \right] \\ &= I - D^* D - C^* (I - Z^* A^*)^{-1} Z^* B^* D - D^* BZ(I - AZ)^{-1} C \\ &\quad - C^* (I - Z^* A^*)^{-1} Z^* B^* BZ(I - AZ)^{-1} C \\ &= C^* M C + C^* (I - Z^* A^*)^{-1} Z^* A^* M C + C^* M A Z(I - AZ)^{-1} C - \\ &\quad - C^* (I - Z^* A^*)^{-1} Z^* (M^{(-1)} - A^* M A) Z(I - AZ)^{-1} C \\ &= C^* (I - Z^* A^*)^{-1} \left\{ (I - Z^* A^*) M (I - AZ) + Z^* A^* M (I - AZ) + \right. \\ &\quad \left. + (I - Z^* A^*) M A Z - Z^* (M^{(-1)} - A^* M A) Z \right\} (I - AZ)^{-1} C \\ &= 0. \end{aligned}$$

The second relation follows likewise. □

If an input-output operator $T \in \mathcal{U}$ is invertible, and the inverse $T^{-1} \in \mathcal{U}$ (T is called *outer*), then D is boundedly invertible and a realization \mathbf{T}^\times of T^{-1} is obtained by rewriting (2.1) as

$$\begin{cases} xZ^{-1} &= x(A - CD^{-1}B) &+ yD^{-1}B \\ u &= -xCD^{-1} &+ yD^{-1} \end{cases} \quad \mathbf{T}^\times = \begin{bmatrix} A - CD^{-1}B & -CD^{-1} \\ D^{-1}B & D^{-1} \end{bmatrix}. \quad (2.5)$$

2.3. Hilbert-Schmidt spaces

In the analysis of time-varying systems, we frequently need to apply input sequences that are zero up to, or after a point i in time, for $i = \dots, -1, 0, 1, \dots$ in turn. To apply an input-output operator to all such sequences in a single expression, it is convenient to act on a *stack* of ℓ_2 -sequences. Thus let

$$\mathcal{X}_2^{\mathcal{M}} = \left\{ u = \begin{bmatrix} \vdots \\ u_0 \\ u_1 \\ \vdots \end{bmatrix}, u_i \in \ell_2^{\mathcal{M}} : \|u\|_{HS}^2 = \sum \|u_i\|_2^2 < \infty \right\}.$$

$\mathcal{X}_2^{\mathcal{M}}$ is a Hilbert-Schmidt space, with respect to the HS inner product

$$\langle A, B \rangle_{HS} = \text{trace } AB^*. \quad (2.6)$$

We will also need Hilbert-Schmidt spaces $\mathcal{U}_2, \mathcal{L}_2, \mathcal{D}_2$ which are subspaces of \mathcal{X}_2 and consist of those elements of $\mathcal{U}, \mathcal{L}, \mathcal{D}$, respectively, for which the HS norm is bounded.

For $u \in \mathcal{X}_2^{\mathcal{M}}, T \in \mathcal{X}(\mathcal{M}, \mathcal{N})$, the expression $y = uT$ is well defined, and gives $y \in \mathcal{X}_2^{\mathcal{N}}$. The i -th row y_i of y satisfies $y_i = u_i T \in \ell_2^{\mathcal{N}}$. The elements of \mathcal{U}_2 contain all ℓ_2 sequences that are zero before point i in time, for all i , and likewise, we will use the space $\mathcal{L}_2 Z^{-1}$ to obtain all sequences that are zero from point i on, for all i . We define $\mathbf{P}_{\mathcal{U}_2}$ as the projection operator of \mathcal{X}_2 onto \mathcal{U}_2 , \mathbf{P}_0 as the projection operator of \mathcal{X}_2 onto \mathcal{D}_2 , and $\mathbf{P}_{\mathcal{L}_2 Z^{-1}}$ as the projection operator of \mathcal{X}_2 onto $\mathcal{L}_2 Z^{-1}$.

Among all operators $[\mathcal{X}_2 \rightarrow \mathcal{X}_2]$, we will only consider those that are *left D -invariant*: $y = uT \Rightarrow Dy = (Du)T$, for all $D \in \mathcal{D}$. Left D -invariant operators essentially treat each row of $u \in \mathcal{X}_2$ independently. Operators in \mathcal{X} are left D -invariant, and so are the above-defined projection operators.

2.4. Diagonal expansions

There is an isomorphism between elements of \mathcal{X}_2 and the space $\ell_2(\mathcal{D})$ of ℓ_2 -sequences of *diagonals*. In particular, when we write $u \in \mathcal{X}_2$ as a sum of its diagonals,

$$u = \dots + u_{[0]} + Zu_{[1]} + Z^2 u_{[2]} + \dots = \dots + u_{[0]} + u_{[1]}^{(-1)} Z + u_{[2]}^{(-2)} Z^2 + \dots,$$

we can associate to u the sequence

$$\tilde{u} = [\dots \quad u_{[0]} \quad u_{[1]}^{(-1)} \quad u_{[2]}^{(-2)} \quad \dots] \in \ell_2(\mathcal{D}).$$

We call \tilde{u} the *diagonal expansion* of $u \in \mathcal{X}_2$. Left D -invariant bounded operators $[\mathcal{X}_2 \rightarrow \mathcal{X}_2]$ have a convenient matrix representation (tensor representation) in terms of the diagonal expansion. For example, for $T \in \mathcal{U}$, we can write $y = uT \Leftrightarrow \tilde{y} = \tilde{u}\tilde{T}$, where

$$\tilde{T} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots \\ & T_{[0]}^{(1)} & T_{[1]} & T_{[2]}^{(-1)} & \dots \\ & & T_{[0]} & T_{[1]}^{(-1)} & \dots \\ & & & T_{[0]}^{(-1)} & \dots \\ & & & & \ddots \end{bmatrix} \quad (2.7)$$

The entries in this matrix representation are themselves diagonals.

3. PRELIMINARY RELATIONS

The analysis of the state space structure needed to represent an input-output operator $T \in \mathcal{U}$ is based on the properties of an operator mapping ‘past’ inputs (inputs in $\mathcal{L}_2\mathcal{Z}^{-1}$) to ‘future’ outputs (the part of outputs in \mathcal{U}_2) and which we shall call the corresponding Hankel operator. Other operators between subspaces of \mathcal{X}_2 will play an important role as well. Using the projection operators defined in the previous section, the action of an input-output mapping $T \in \mathcal{U}$ on an input $u \in \mathcal{L}_2\mathcal{Z}^{-1}$ can be broken down into a few operators on a reduced domain and range. Thus, define the operators H_T , K_T and V_T in the following way:

$$\begin{aligned} H_T &: \mathcal{L}_2\mathcal{Z}^{-1} \rightarrow \mathcal{U}_2, & uH_T &= \mathbf{P}_{\mathcal{U}_2}(uT) \\ K_T &: \mathcal{L}_2\mathcal{Z}^{-1} \rightarrow \mathcal{L}_2\mathcal{Z}^{-1}, & uK_T &= \mathbf{P}_{\mathcal{L}_2\mathcal{Z}^{-1}}(uT) \\ V_T &: \mathcal{L}_2\mathcal{Z}^{-1} \rightarrow \mathcal{D}_2, & uV_T &= \mathbf{P}_0(uT). \end{aligned}$$

For $u \in \mathcal{L}_2\mathcal{Z}^{-1}$ we have that $uT = uK_T + uH_T$. We call H_T the Hankel operator of T : it is the map of inputs in $\mathcal{L}_2\mathcal{Z}^{-1}$ to the part in \mathcal{U}_2 of the corresponding outputs, and plays a crucial role in realization theory [VD1, V2]. K_T can be called a ‘past Toeplitz’ operator associated to T . Note that V_T is a further restriction of H_T .

We will define the one-sided diagonal expansions of signals u in $\mathcal{L}_2\mathcal{Z}^{-1}$ and y in \mathcal{U}_2 as

$$\begin{aligned} \tilde{u}^- &= [u_{[-1]}^{(1)} \quad u_{[-2]}^{(2)} \quad \dots] \in \ell_2^-(\mathcal{D}), \\ \tilde{y}^+ &= [y_{[0]} \quad y_{[1]}^{(-1)} \quad y_{[2]}^{(-2)} \quad \dots] \in \ell_2^+(\mathcal{D}). \end{aligned}$$

Induced by the isomorphism, the definitions

$$\begin{aligned} y = uH_T &\in \mathcal{U}_2 &\Leftrightarrow \tilde{y}^+ &= \tilde{u}^- \tilde{H}_T \in \ell_2^+(\mathcal{D}) \\ y = uK_T &\in \mathcal{L}_2\mathcal{Z}^{-1} &\Leftrightarrow \tilde{y}^- &= \tilde{u}^- \tilde{K}_T \in \ell_2^-(\mathcal{D}) \\ D = uV_T &\in \mathcal{D}_2 &\Leftrightarrow D &= \tilde{u}^- \tilde{V}_T \in \mathcal{D}_2 \end{aligned}$$

lead to diagonal matrix representations of H_T , K_T , and V_T as

$$\tilde{H}_T = \begin{bmatrix} T_{[1]} & T_{[2]}^{(-1)} & T_{[3]}^{(-2)} & \dots \\ T_{[2]} & T_{[3]}^{(-1)} & & \\ T_{[3]} & & \ddots & \\ \vdots & & & \ddots \end{bmatrix}, \quad \tilde{V}_T = \begin{bmatrix} T_{[1]} \\ T_{[2]} \\ T_{[3]} \\ \vdots \end{bmatrix}, \quad \tilde{K}_T = \begin{bmatrix} T_{[0]}^{(1)} & & \mathbf{0} & \\ T_{[1]}^{(1)} & T_{[0]}^{(2)} & & \\ T_{[2]}^{(1)} & T_{[1]}^{(2)} & T_{[0]}^{(3)} & \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}. \quad (3.1)$$

Note that these are (mirrored) submatrices of \tilde{T} in (2.7). Taking the i -th entry of each diagonal gives back the (mirrored) submatrices K_i , H_i of T as defined in (1.4).

Connected to a state realization, we can distinguish controllability and observability operators

$$\mathcal{C} := \begin{bmatrix} B^{(1)} \\ B^{(2)}A^{(1)} \\ B^{(3)}A^{(2)}A^{(1)} \\ \vdots \end{bmatrix} \quad \mathcal{O} := \begin{bmatrix} C & AC^{(-1)} & AA^{(-1)}C^{(-2)} & \dots \end{bmatrix}, \quad (3.2)$$

which play the same role as the corresponding operators in the time-invariant context. If the realization is strictly stable, $\ell_A < 1$, then \mathcal{C}^* and \mathcal{O} are bounded operators $[\mathcal{D}_2 \rightarrow \ell_2^-(\mathcal{D})]$ and $[\mathcal{D}_2 \rightarrow \ell_2^+(\mathcal{D})]$, and in fact, they are diagonal expansions of $[BZ(I-AZ)^{-1}]^*$ and $(I-AZ)^{-1}C$, respectively. From equation (2.4) it is seen that the controllability and observability Gramians are given by $\Lambda_c = \mathcal{C}^*\mathcal{C}$ and $\Lambda_o = \mathcal{O}\mathcal{O}^*$, respectively. It is straightforward to verify using (2.2) that if $\{A, B, C, D\}$ is a realization of T , then \tilde{H}_T admits a decomposition

$$\tilde{H}_T = \mathcal{C}\mathcal{O}.$$

Since \tilde{V}_T is the first column of \tilde{H}_T , we have from equation (3.2) that

$$\tilde{V}_T = \mathcal{C} \cdot \mathcal{C}, \quad (3.3)$$

and from (3.2) and (3.1), it is seen that \mathcal{C} and \tilde{K}_T satisfy the shift-invariance properties

$$\mathcal{C}^{(-1)} = \begin{bmatrix} B \\ \mathcal{C}A \end{bmatrix}, \quad \tilde{K}_T^{(-1)} = \begin{bmatrix} T_{[0]} & 0 \\ \tilde{V}_T & \tilde{K}_T \end{bmatrix}. \quad (3.4)$$

4. STRICTLY CONTRACTIVE SYSTEMS

As indicated in the introduction, a lossless embedding of an input-output operator $T \in \mathcal{U}$ is possible only if T is at least contractive. In this section, we will explore the consequences of assuming the strict contractivity of T , to determine sufficient conditions for an embedding to exist if T is strictly contractive. This is done in two steps. Lemma 4.4 derives a general relation in terms of \tilde{V}_T and \tilde{K}_T which is a direct consequence of the strict contractivity of T . Theorem 4.5 uses this relation to show that some quantity $M \in \mathcal{D}$, defined by $M = \mathcal{C}^*(I - \tilde{K}_T \tilde{K}_T^*)^{-1} \mathcal{C}$, is positive, and will give a recursion for this M in terms of state space quantities of T . This recursion turns out to be the same Riccati recursion as for M in the embedding problem (viz. equation (1.3)), and will prove the essential step in the embedding problem for strictly contractive systems (section 6). The case where T is contractive, but not necessarily strictly contractive, is discussed in section 5.

4.1. Contractivity of an input-output operator

A left D -invariant hermitian operator $A : [\mathcal{X}_2 \rightarrow \mathcal{X}_2]$ is positive semidefinite, $A \geq 0$, if for all $u \in \mathcal{X}_2$, $\langle uA, u \rangle_{HS} \geq 0$. This definition is equivalent to the usual definition of positivity of operators in Hilbert space, but now applies to arguments $u \in \mathcal{X}_2$. A is uniformly positive definite, notation $A \gg 0$, if there exists an $\varepsilon > 0$ such that, for all u in \mathcal{X}_2 , $\langle uA, u \rangle_{HS} \geq \varepsilon \langle u, u \rangle_{HS}$. It is known that a positive definite operator $A \in \mathcal{X}$ is uniformly positive if and only if A is boundedly invertible in \mathcal{X} . We will sometimes use the following form of the above definitions, which are obtained by using the definition of the Hilbert-Schmidt inner product (2.6).

Lemma 4.1. *Let $A \in [\mathcal{X}_2 \rightarrow \mathcal{X}_2]$ be a left D -invariant hermitian operator. Then*

$$\begin{aligned} A \geq 0 &\Leftrightarrow \mathbf{P}_0(uAu^*) \geq 0, \text{ for all } u \in \mathcal{X}_2, \\ A \gg 0 &\Leftrightarrow \exists \varepsilon > 0 : \mathbf{P}_0(uAu^*) \geq \varepsilon \mathbf{P}_0(uu^*), \text{ for all } u \in \mathcal{X}_2. \end{aligned}$$

PROOF $\langle uA, u \rangle_{HS} = \text{trace } \mathbf{P}_0(uAu^*)$. Because of left D -invariance, $\text{trace } \mathbf{P}_0(uAu^*) \geq 0$ for all $u \in \mathcal{X}_2$ implies that $D\mathbf{P}_0(uAu^*)D^* \geq 0$ for all $D \in \mathcal{D}$: in particular, all individual entries of the diagonal $\mathbf{P}_0(uAu^*)$ must be positive semidefinite, so that $\mathbf{P}_0(uAu^*) \geq 0$. The reverse is obvious. \square

Let T be an input-output operator in \mathcal{U} . We define T to be contractive, respectively strictly contractive, if

$$I - TT^* \geq 0, \quad \text{resp.} \quad I - TT^* \gg 0.$$

In the latter case, $I - TT^*$ is boundedly invertible. In this section, we will from now on focus on the case that T is strictly contractive. The more general case is treated in section 5. Because of the identity $I + T^*(I - TT^*)^{-1}T = (I - T^*T)^{-1}$ it is clear that $I - TT^* \gg 0$ implies that $I - T^*T \gg 0$ also.

Lemma 4.2. *If $T \in \mathcal{X}$ is strictly contractive, then K_T and \tilde{K}_T are strictly contractive.*

PROOF Let $u \in \mathcal{L}_2\mathcal{Z}^{-1}$, and $y = uK_T$. Since T is strictly contractive, we have from the above definition that

$$\begin{aligned} \mathbf{P}_0[u(I - K_T K_T^*)u^*] &= \mathbf{P}_0(uu^*) - \mathbf{P}_0(yy^*) \\ &\geq \mathbf{P}_0[u(I - TT^*)u^*] \\ &\geq \varepsilon \mathbf{P}_0(uu^*) \quad (\text{some } \varepsilon > 0). \end{aligned}$$

Hence K_T is strictly contractive. A similar derivation holds for \tilde{K}_T , which is isometrically isomorphic to K_T . \square

4.2. Strict contractivity in terms of a state space realization

The following lemma is standard. It will be used to derive a recursive relation that describes the contractivity of $\tilde{K}_T^{(-1)}$ in terms of that of \tilde{K}_T .

Lemma 4.3. (Schur Complements and Inversion Formula) *With \mathcal{H}_1 and \mathcal{H}_2 Hilbert spaces, let $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$, $B : \mathcal{H}_1 \rightarrow \mathcal{H}_2$, $C : \mathcal{H}_2 \rightarrow \mathcal{H}_2$ be bounded operators, and let A and C be self-adjoint. Then*

$$X := \begin{bmatrix} A & B^* \\ B & C \end{bmatrix} \gg 0 \quad \Leftrightarrow \quad \begin{cases} (1) & C \gg 0 \\ (2) & A - B^*C^{-1}B \gg 0. \end{cases}$$

If $X \gg 0$, then

$$\begin{bmatrix} A & B^* \\ B & C \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & C^{-1} \end{bmatrix} + \begin{bmatrix} I \\ -C^{-1}B \end{bmatrix} (A - B^*C^{-1}B)^{-1} \begin{bmatrix} I & -B^*C^{-1} \end{bmatrix}.$$

PROOF $X \gg 0$ implies that $C \gg 0$, so that C^{-1} exists. The result is immediate, from the factorization

$$\begin{bmatrix} A & B^* \\ B & C \end{bmatrix} = \begin{bmatrix} I & B^*C^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A - B^*C^{-1}B & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I & 0 \\ C^{-1}B & I \end{bmatrix}.$$

□

Lemma 4.4. *Let $T \in \mathcal{U}$ be an input-output operator. If T is strictly contractive, then*

$$I - T_{[0]}^* T_{[0]} - \tilde{V}_T^* (I - \tilde{K}_T \tilde{K}_T^*)^{-1} \tilde{V}_T \gg 0.$$

PROOF Since T is strictly contractive, lemma 4.2 ensures that \tilde{K}_T and $\tilde{K}_T^{(-1)}$ are also strictly contractive. Using equation (3.4), we have that

$$I - \tilde{K}_T^{(-1)*} \tilde{K}_T^{(-1)} = \begin{bmatrix} I - T_{[0]}^* T_{[0]} - \tilde{V}_T^* \tilde{V}_T & -\tilde{V}_T^* \tilde{K}_T \\ -\tilde{K}_T^* \tilde{V}_T & I - \tilde{K}_T^* \tilde{K}_T \end{bmatrix}. \quad (4.1)$$

With lemma 4.3, it is seen that this expression is uniformly positive if and only if

$$\begin{cases} (1) & I - \tilde{K}_T^* \tilde{K}_T \gg 0 \\ (2) & I - T_{[0]}^* T_{[0]} - \tilde{V}_T^* \tilde{V}_T - \tilde{V}_T^* \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* \tilde{V}_T \gg 0. \end{cases}$$

The first condition is satisfied because T is strictly contractive. The second condition is equal to the result, because of the equality $I + \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* = (I - \tilde{K}_T \tilde{K}_T^*)^{-1}$. □

Theorem 4.5. *Let $T \in \mathcal{U}$ be a locally finite input-output operator with state space realization $\{A, B, C, D\}$, where $A \in \mathcal{D}(\mathcal{B}, \mathcal{B}^{(-1)})$ is strictly stable ($\ell_A < 1$). If T is strictly contractive, then $M \in \mathcal{D}(\mathcal{B}, \mathcal{B})$, defined by*

$$M = C^* (I - \tilde{K}_T \tilde{K}_T^*)^{-1} C, \quad (4.2)$$

satisfies the relations $M \geq 0$, $I - D^* D - C^* M C \gg 0$, and

$$M^{(-1)} = A^* M A + B^* B + [A^* M C + B^* D] (I - D^* D - C^* M C)^{-1} [D^* B + C^* M A]. \quad (4.3)$$

If in addition the state space realization is uniformly controllable, then $M \gg 0$.

PROOF If T is strictly contractive, then M is well-defined and $M \geq 0$. With the definition of M and using the fact that $D = T_{[0]}$ and $\tilde{V}_T = C \cdot C$ (equation (3.3)), the uniform strict positivity of $I - D^* D - C^* M C$ follows directly from

lemma 4.4. The recursive relation for M is obtained by an application of Schur's inversion formula (lemma 4.3) to equation (4.1), which gives

$$\left[I - \tilde{K}_T^{(-1)*} \tilde{K}_T^{(-1)} \right]^{-1} = \begin{bmatrix} 0 & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \end{bmatrix} + \begin{bmatrix} I & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* \tilde{V}_T \end{bmatrix} \Phi^{-2} \begin{bmatrix} I & \tilde{V}_T^* \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \end{bmatrix} \quad (4.4)$$

with

$$\begin{aligned} \Phi^2 &= I - T_{[0]}^* T_{[0]} - \tilde{V}_T^* \tilde{V}_T - \tilde{V}_T^* \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* \tilde{V}_T \\ &= I - D^* D - C^* M C. \end{aligned}$$

The invertibility of Φ^2 was already shown. Inserting this expression into the definition of $M^{(-1)}$, and using the expression for $\mathcal{C}^{(-1)}$ in (3.4), $M^{(-1)}$ is obtained as

$$\begin{aligned} M^{(-1)} &= \mathcal{C}^{(-1)*} \left[I - \tilde{K}_T^{(-1)} \tilde{K}_T^{(-1)*} \right]^{-1} \mathcal{C}^{(-1)} = \\ &= \mathcal{C}^{(-1)*} \left[I + \tilde{K}_T^{(-1)} \left(I - \tilde{K}_T^{(-1)*} \tilde{K}_T^{(-1)} \right)^{-1} \tilde{K}_T^{(-1)*} \right] \mathcal{C}^{(-1)} \\ &= \begin{bmatrix} B^* & A^* \mathcal{C}^* \end{bmatrix} \begin{bmatrix} B \\ \mathcal{C} A \end{bmatrix} + \begin{bmatrix} B^* & A^* \mathcal{C}^* \end{bmatrix} \begin{bmatrix} T_{[0]} & 0 \\ \tilde{V}_T & \tilde{K}_T \end{bmatrix} \\ &\quad \cdot \left(\begin{bmatrix} 0 & 0 \\ 0 & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \end{bmatrix} + \begin{bmatrix} I & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* \tilde{V}_T \end{bmatrix} \Phi^{-2} \begin{bmatrix} I & \tilde{V}_T^* \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \end{bmatrix} \right) \cdot \begin{bmatrix} T_{[0]} & 0 \\ \tilde{V}_T & \tilde{K}_T \end{bmatrix}^* \begin{bmatrix} B \\ \mathcal{C} A \end{bmatrix} \\ &= B^* B + A^* \mathcal{C}^* \mathcal{C} A + A^* \mathcal{C}^* \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* \mathcal{C} A + \\ &\quad + \left(B^* D + A^* \mathcal{C}^* \left[I + \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* \right] \mathcal{C} \right) \cdot \Phi^{-2} \cdot \left(D^* B + C^* \mathcal{C}^* \left[I + \tilde{K}_T^* (I - \tilde{K}_T \tilde{K}_T^*)^{-1} \tilde{K}_T \right] \mathcal{C} A \right) \\ &= B^* B + A^* M A + (A^* M C + B^* D) \Phi^{-2} (D^* B + C^* M A). \end{aligned}$$

□

The equation (4.3) for M is actually a recursive equation, which becomes apparent if we write $M = \text{diag}[M_i]$ and take the i -th entry of every diagonal in the equation: this produces the Riccati recursion (1.3). Theorem 4.5 claims that for a strictly contractive system, the Riccati recursion has a positive solution M , which is given in explicit form. This M plays a crucial role in the construction of a lossless embedding, later in section 6. It also furnishes part of the proof of the Bounded Real Lemma.

5. CONTRACTIVE SYSTEMS: THE BOUNDARY CASE

We will now derive an equivalent of theorem 4.5 for the case where T is contractive but not necessarily strictly contractive: $I - T T^* \geq 0$. While the mathematical derivation is more complicated now, the resulting theorem is only slightly altered. It will turn out that K_T is not strictly contractive, and that, instead of $(I - \tilde{K}_T \tilde{K}_T^*)^{-1}$, we will have to use the pseudo-inverse of $(I - \tilde{K}_T^* \tilde{K}_T)$. Mathematical complications arise because the range of $(I - \tilde{K}_T^* \tilde{K}_T)$ is not necessarily closed, so that its pseudo-inverse can be unbounded.

5.1. Schur inversion formulas for positive semi-definite operators

Let be given some operator A on a Hilbert space \mathcal{H} . The range of A is $\mathcal{R}(A) = \{Ax : x \in \mathcal{H}\}$, its closure is $\overline{\mathcal{R}(A)}$, and its nullspace is denoted by $\mathcal{N}(A) = \{x : Ax = 0\}$, which is a closed subspace. An orthogonal complement is denoted by \perp . The operator pseudo-inverse of A is defined as follows (following Beutler and Root [BR]).

Definition 5.1. Let \mathcal{H} be a Hilbert space, and A be a bounded linear operator defined on \mathcal{H} . The linear operator $A^\dagger : \mathcal{H} \rightarrow \mathcal{H}$ is a pseudo-inverse of A if and only if it is defined on $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ (which is dense in \mathcal{H}) and satisfies the following conditions:

$$\begin{aligned} (1) \quad \mathcal{N}(A^\dagger) &= \mathcal{R}(A)^\perp \\ (2) \quad \overline{\mathcal{R}(A^\dagger)} &= \mathcal{N}(A)^\perp \quad (= \overline{\mathcal{R}(A^*)}) \\ (3) \quad AA^\dagger x &= x \quad \text{for all } x \in \mathcal{R}(A). \end{aligned}$$

It is proven in [BR] that $(A^\dagger)^\dagger = A$, $(A^\dagger)^* = (A^*)^\dagger$, $(A^*A)^\dagger = A^\dagger A^{*\dagger}$, and that A^\dagger is bounded if and only if $\mathcal{R}(A)$ is closed. We will also apply a result of Douglas [D3] on majorization of operators on Hilbert spaces:

Theorem 5.2. Let A and B be bounded operators on a Hilbert space \mathcal{H} . The following are equivalent:

$$\begin{aligned} (1) \quad AA^* &\leq \lambda^2 BB^* \quad (\text{some } \lambda > 0), \\ (2) \quad \mathcal{R}(A) &\subset \mathcal{R}(B), \\ (3) \quad A &= BC \quad \text{for some bounded operator } C \text{ on } \mathcal{H}. \end{aligned}$$

If (1)-(3) are valid, then a unique operator C exists such that

$$\begin{aligned} (a) \quad \|C\| &= \inf\{\mu : AA^* \leq \mu BB^*\}, \\ (b) \quad \mathcal{N}(A) &= \mathcal{N}(C), \\ (c) \quad \mathcal{R}(C) &\subset \overline{\mathcal{R}(B^*)}. \end{aligned}$$

The ‘unique operator C ’ in this theorem is in fact $C = B^\dagger A$, since also B^\dagger is uniquely defined and $B^\dagger A$ qualifies for C . Consequently, if $AA^* \leq BB^*$, then this C satisfies $\|C\| < 1$.

Using pseudo-inverses, the Schur inversion formula (lemma 4.3) can be extended to the case where X is not uniformly positive.

Lemma 5.3. With \mathcal{H}_1 and \mathcal{H}_2 Hilbert spaces, let $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$, $B : \mathcal{H}_1 \rightarrow \mathcal{H}_2$, $C : \mathcal{H}_2 \rightarrow \mathcal{H}_2$ be bounded operators, and let A and C be self-adjoint. Then

$$X := \begin{bmatrix} A & B^* \\ B & C \end{bmatrix} \geq 0 \quad \Leftrightarrow \quad \begin{cases} (1) & C \geq 0, \\ (2) & \mathcal{R}(B) \subset \mathcal{R}(C^{1/2}); \quad \text{i.e., } B_1 = C^{\dagger/2}B \text{ is bounded,} \\ (3) & A - B_1^* B_1 \geq 0. \end{cases}$$

Lemma 5.4. Let A, B, C, X be as in lemma 5.3. Let $X \geq 0$ and write $B_1 = C^{\dagger/2}B$. Define the operator W^\ddagger :

$$W^\ddagger = \begin{bmatrix} (A - B_1^* B_1)^{\dagger/2} & \\ & I \end{bmatrix} \begin{bmatrix} I & -B_1^* \\ & I \end{bmatrix} \begin{bmatrix} I & \\ & C^{\dagger/2} \end{bmatrix}.$$

Then W^\ddagger is well-defined and bounded on $\mathcal{R}(X^{1/2})$. If v is some bounded operator with range in $\mathcal{R}(X^{1/2})$, and if

$$v_1 = X^{\dagger/2}v, \quad v_2 = W^\ddagger v$$

then v_1 and v_2 are bounded, and $v_1^* v_1 = v_2^* v_2$.

The proof of both lemmas appears as an appendix. Note that $W^\ddagger \neq X^{\dagger/2}$, but rather $W^\ddagger = UX^{\dagger/2}$ on $\mathcal{R}(X^{1/2})$, where U is some Hilbert space isometry such that $U^*U = I$. The point is that W^\ddagger is specified in terms of A, B, C , whereas it is hard to do so for $X^{\dagger/2}$.

5.2. Contractivity in terms of a state space realization

We are now ready to derive a solution to the embedding problem along the lines of section 4 for the case where T is contractive, but not necessarily strictly contractive. Recall the definition of H_T and K_T of section 3.

Lemma 5.5. *Let T be an input-output operator in \mathcal{U} . If T is contractive, then*

$$I - K_T K_T^* \geq H_T H_T^* \geq 0, \quad (5.1)$$

and hence K_T and \tilde{K}_T are contractive.

PROOF Let $u \in \mathcal{L}_2 Z^{-1}$, and put $y = uT = uK_T + uH_T$. The contractivity of T implies

$$\begin{aligned} & \mathbf{P}_0(uu^*) - \mathbf{P}_0(yy^*) \geq 0 \\ \Leftrightarrow & \mathbf{P}_0(u[I - TT^*]u^*) \geq 0 \\ \Leftrightarrow & \mathbf{P}_0(u[I - K_T K_T^* - H_T H_T^*]u^*) \geq 0 \\ \Leftrightarrow & \mathbf{P}_0(u[I - K_T K_T^*]u^*) \geq \mathbf{P}_0(uH_T H_T^* u^*) \geq 0. \end{aligned}$$

Hence $I - K_T K_T^* \geq 0$ on $\mathcal{L}_2 Z^{-1}$. \tilde{K}_T is isometrically isomorphic to K_T and is also contractive. \square

Corollary 5.6. *If \mathbf{T} is a uniformly observable realization of T , then $\mathcal{R}(\tilde{K}_T^* \mathcal{C}) \subset \mathcal{R}(I - \tilde{K}_T^* \tilde{K}_T)^{1/2}$ and hence \mathcal{C}_1 defined by*

$$\mathcal{C}_1 = (I - \tilde{K}_T^* \tilde{K}_T)^{\dagger/2} \tilde{K}_T^* \mathcal{C} \quad (5.2)$$

is bounded.

PROOF Apply theorem 5.2 to (5.1). From $I - K_T K_T^* \geq H_T H_T^*$ it follows that $H_T = (I - K_T K_T^*)^{1/2} N$, for some operator N with $\|N\| \leq 1$. Taking diagonal expansions, we have that $\tilde{H}_T = (I - \tilde{K}_T \tilde{K}_T^*)^{1/2} \tilde{N}$, and with $\tilde{H}_T = \mathcal{C} \mathcal{O}$ such that $\mathcal{O} \mathcal{O}^* \gg 0$, we obtain

$$\begin{aligned} \tilde{K}_T^* \mathcal{C} &= \tilde{K}_T^* \mathcal{C} \mathcal{O} \mathcal{O}^* (\mathcal{O} \mathcal{O}^*)^{-1} \\ &= \tilde{K}_T^* \tilde{H}_T \mathcal{O}^* (\mathcal{O} \mathcal{O}^*)^{-1} \\ &= \tilde{K}_T^* (I - \tilde{K}_T \tilde{K}_T^*)^{1/2} \tilde{N} \mathcal{O}^* (\mathcal{O} \mathcal{O}^*)^{-1} \\ &= (I - \tilde{K}_T^* \tilde{K}_T)^{1/2} \mathcal{C}_1 \end{aligned}$$

where $\mathcal{C}_1 = \tilde{K}_T^* \tilde{N} \cdot \mathcal{O}^* (\mathcal{O} \mathcal{O}^*)^{-1}$ is bounded. \square

For \mathcal{C}_1 defined in (5.2), define the operator $M \in \mathcal{D}$ by

$$M = \mathcal{C}^* \mathcal{C} + \mathcal{C}_1^* \mathcal{C}_1. \quad (5.3)$$

M is bounded, and $M \gg 0$ if $\mathcal{C}^* \mathcal{C} \gg 0$, i.e., if the realization is uniformly controllable. This definition of M is compatible with the definition of M in (4.2) if T is strictly contractive, viz. $M = \mathcal{C}^* (I - \tilde{K}_T \tilde{K}_T^*)^{-1} \mathcal{C}$, because then $\mathcal{C}_1^* \mathcal{C}_1 = \mathcal{C}^* \tilde{K}_T (I - \tilde{K}_T \tilde{K}_T^*)^{-1} \tilde{K}_T^* \mathcal{C}$, and $I + \tilde{K}_T (I - \tilde{K}_T \tilde{K}_T^*)^{-1} \tilde{K}_T^* = (I - \tilde{K}_T \tilde{K}_T^*)^{-1}$. The latter relation is however not necessarily valid if a pseudo-inverse is used.

The following theorem subsumes theorem 4.5.

Theorem 5.7. *Let $T \in \mathcal{U}$ be an input-output operator with a strictly stable state space realization $\{A, B, C, D\}$. If T is contractive and the realization is uniformly observable, then M defined by (5.2) and (5.3) is bounded, $M \geq 0$, and*

$$M^{(-1)} = A^* M A + B^* B + ([A^* M C + B^* D] \Phi^\dagger) \cdot (\Phi^\dagger [D^* B + C^* M A]) \quad (5.4)$$

with $\Phi = (I - D^* D - C^* M C)^{1/2}$ and $I - D^* D - C^* M C \geq 0$. If, in addition, the state space realization is [uniformly] controllable then $M > 0$ [$M \gg 0$].

PROOF The proof uses the expressions for \tilde{V}_T , \tilde{K}_T and \mathcal{C} as given by equations (3.3) and (3.4). To find an expression for $M^{(-1)}$, put

$$X = (I - \tilde{K}_T^* \tilde{K}_T)^{(-1)} = \begin{bmatrix} I - T_{[0]}^* T_{[0]} - \tilde{V}_T^* \tilde{V}_T & -\tilde{V}_T^* \tilde{K}_T \\ -\tilde{K}_T^* \tilde{V}_T & I - \tilde{K}_T^* \tilde{K}_T \end{bmatrix}.$$

According to lemma 5.5, $X \geq 0$. Lemma 5.3 then implies that $\mathcal{R}(\tilde{K}_T^* \tilde{V}_T) \subset \mathcal{R}(I - \tilde{K}_T^* \tilde{K}_T)^{1/2}$ so that $(I - \tilde{K}_T^* \tilde{K}_T)^{\dagger/2} \tilde{K}_T^* \tilde{V}_T = \mathcal{C}_1 C$ is bounded. (This result would also follow from corollary 5.6 because $\mathcal{R}(\tilde{K}_T^* \tilde{V}_T) = \mathcal{R}(\tilde{K}_T^* \mathcal{C} C) \subset \mathcal{R}(\tilde{K}_T^* \mathcal{C})$.) Let

$$\begin{aligned} \Phi &= \left[I - T_{[0]}^* T_{[0]} - \tilde{V}_T^* \tilde{V}_T - C^* \mathcal{C}_1^* \mathcal{C}_1 C \right]^{1/2} \\ &= \left[I - D^* D - C^* (C^* \mathcal{C} + \mathcal{C}_1^* \mathcal{C}_1) C \right]^{1/2} \\ &= (I - D^* D - C^* M C)^{1/2}. \end{aligned}$$

The third item of lemma 5.3 implies that $I - D^* D - C^* M C \geq 0$. Put

$$\begin{aligned} W^\dagger &= \begin{bmatrix} \Phi^\dagger & \\ & I \end{bmatrix} \begin{bmatrix} I & C^* \mathcal{C}_1^* \\ & I \end{bmatrix} \begin{bmatrix} I & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{\dagger/2} \end{bmatrix} \\ v &= [\tilde{K}_T^* \mathcal{C}]^{(-1)} = \tilde{K}_T^{*(-1)} \begin{bmatrix} B \\ \mathcal{C} A \end{bmatrix} = \begin{bmatrix} D^* B + C^* \mathcal{C}^* \mathcal{C} A \\ \tilde{K}_T^* \mathcal{C} A \end{bmatrix}. \end{aligned}$$

Then lemma 5.4 yields that the operator $v_1 = X^{\dagger/2} v = \mathcal{C}_1^{(-1)}$ is bounded, and $v_2 = W^\dagger v$ is such that $v_1^* v_1 = v_2^* v_2$. Evaluation of v_2 gives

$$\begin{aligned} v_2 = W^\dagger v &= \begin{bmatrix} \Phi^\dagger & \\ & I \end{bmatrix} \begin{bmatrix} I & C^* \mathcal{C}_1^* \\ & I \end{bmatrix} \begin{bmatrix} I & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{\dagger/2} \end{bmatrix} \begin{bmatrix} D^* B + C^* \mathcal{C}^* \mathcal{C} A \\ \tilde{K}_T^* \mathcal{C} A \end{bmatrix} \\ &= \begin{bmatrix} \Phi^\dagger & \\ & I \end{bmatrix} \begin{bmatrix} I & C^* \mathcal{C}_1^* \\ & I \end{bmatrix} \begin{bmatrix} D^* B + C^* \mathcal{C}^* \mathcal{C} A \\ \mathcal{C}_1 A \end{bmatrix} \\ &= \begin{bmatrix} \Phi^\dagger (D^* B + C^* M A) \\ \mathcal{C}_1 A \end{bmatrix}. \end{aligned}$$

Hence

$$\begin{aligned} [\mathcal{C}_1^* \mathcal{C}_1]^{(-1)} &= v_1^* v_1 = v_2^* v_2 \\ &= A^* \mathcal{C}_1^* \mathcal{C}_1 A + ([B^* D + A^* M C] \Phi^\dagger) \cdot (\Phi^\dagger [D^* B + C^* M A]) \end{aligned}$$

and with $\mathcal{C}^{(-1)} = \begin{bmatrix} B \\ \mathcal{C} A \end{bmatrix}$ we finally obtain

$$\begin{aligned} M^{(-1)} &= [\mathcal{C}^* \mathcal{C}]^{(-1)} + [\mathcal{C}_1^* \mathcal{C}_1]^{(-1)} \\ &= B^* B + A^* \mathcal{C}^* \mathcal{C} A + A^* \mathcal{C}_1^* \mathcal{C}_1 A + ([B^* D + A^* M C] \Phi^\dagger) \cdot (\Phi^\dagger [D^* B + C^* M A]) \\ &= A^* M A + B^* B + ([B^* D + A^* M C] \Phi^\dagger) \cdot (\Phi^\dagger [D^* B + C^* M A]). \end{aligned}$$

□

The result of this section is thus a relatively simple extension of theorem 4.5, although we need the given realization to be uniformly observable. This condition is too strong: we only need “observability at the boundary”, but this is hard to express. The recursion for M is very close to (and encompasses) the expression that we have obtained before in the strictly contractive case. The above theorem will allow the embedding theorems in the next section to include contractive systems that need not be strictly contractive. It also gives part of the proof of the Bounded Real Lemma.

6. LOSSLESS EMBEDDING

In this section, we will solve the lossless embedding problem as defined in the introduction. We start with an intermediate result.

Theorem 6.1. (Isometric embedding) *Let $T \in \mathcal{U}(\mathcal{M}, \mathcal{N})$ be a locally finite input-output operator with strictly stable state realization $\mathbf{T} = \{A, B, C, D\}$. If $I - T^* T \gg 0$, or $I - T^* T \geq 0$ and T is uniformly observable, then T has an extension $\Sigma_a \in \mathcal{U}(\mathcal{M} \times \mathcal{N}, \mathcal{N})$,*

$$\Sigma_a = \begin{bmatrix} T \\ \Sigma_{21} \end{bmatrix}$$

such that $\Sigma_a^* \Sigma_a = I$ and $A_{\Sigma_a} = A$. A realization for Σ_{21} is

$$\Sigma_{21} = \begin{bmatrix} A & C \\ B_2 & D_{21} \end{bmatrix} = \begin{bmatrix} A & C \\ -\Phi^\dagger(D^*B + C^*MA) & \Phi \end{bmatrix} \quad (6.1)$$

where $\Phi = (I - D^*D - C^*MC)^{1/2}$ and M is as defined in (5.3).

PROOF Let Σ_a be of the form

$$\Sigma_a = \begin{bmatrix} A & C \\ B & D \\ \hline B_2 & D_{21} \end{bmatrix} \quad (6.2)$$

in which B_2 and D_{21} are to be determined such that $\Sigma_a^* \Sigma_a = I$. Using lemma 2.1, this is the case if there is an $M \geq 0$ such that

$$\begin{cases} A^*MA + B^*B + B_2^*B_2 = M^{(-1)} \\ A^*MC + B^*D + B_2^*D_{21} = 0 \\ C^*MC + D^*D + D_{21}^*D_{21} = I \end{cases} \quad (6.3)$$

We will show that M given by equation (5.3) is a positive semidefinite solution to these equations. Indeed, under the conditions imposed on T , theorem 4.5 [theorem 5.7] ensures that this M satisfies $M \geq 0$, $I - D^*D - C^*MC \gg 0$ [$I - D^*D - C^*MC \geq 0$], and

$$M^{(-1)} = A^*MA + B^*B + ([A^*MC + B^*D]\Phi^\dagger) \cdot (\Phi^\dagger[D^*B + C^*MA]), \quad (6.4)$$

where $\Phi = (I - D^*D - C^*MC)^{1/2}$. With B_2 and D_{21} as in (6.1), it immediately follows that equations (6.3) are satisfied. \square

In the above theorem, M can be interpreted as the controllability Gramian of Σ , and since $M = C^*C + C_1^*C_1$ with C_1 as in (5.2), it is seen that $C_1^*C_1$ is the controllability Gramian of Σ_{21} . (A more detailed analysis shows that $-C_1$ is its controllability operator.)

Suppose that $\|T\| < 1$ so that $I - T^*T$ is invertible. A result of Arveson [A], which is applicable in the present context, claims that there is a factor Σ_{21} of $I - T^*T$ which is *outer*, i.e., such that $\Sigma_{21}^{-1} \in \mathcal{U}$. We will show that our choice for Σ_{21} , as defined by the realization Σ_{21} in (6.1), is in fact outer. To this end, we will look at a possible realization for Σ_{21}^{-1} , viz. (2.5),

$$\Sigma_{21}^\times = \begin{bmatrix} A^\times & C^\times \\ B^\times & D^\times \end{bmatrix} = \begin{bmatrix} A - CD_{21}^{-1}B_2 & -CD_{21}^{-1} \\ D_{21}^{-1}B_2 & D_{21}^{-1} \end{bmatrix} \quad (6.5)$$

and will show that this realization is strictly stable: $\ell_{A^\times} < 1$. In that case, we can conclude that $\Sigma_{21}^{-1} \in \mathcal{U}$.

Proposition 6.2. *Suppose $\|T\| < 1$. Define Σ_{21}^\times as in (6.5) and theorem 6.1. Then $\ell_{A^\times} < 1$, and Σ_{21} is outer.*

PROOF We first assert that the controllability operator of Σ_{21}^\times is given by $C^\times = -(I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* C$. It is sufficient to show that the given formula of C^\times satisfies the recursion $C^{\times(-1)} = \begin{bmatrix} B^\times \\ C^\times A^\times \end{bmatrix}$. Indeed, with equations (3.3), (3.4), (4.4),

$$\begin{aligned} C^{\times(-1)} &= -(I - \tilde{K}_T^* \tilde{K}_T)^{-(-1)} \tilde{K}_T^{*(-1)} C^{(-1)} = \\ &= - \left(\begin{bmatrix} 0 & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \end{bmatrix} + \begin{bmatrix} I & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* \tilde{V}_T \end{bmatrix} \Phi^{-2} \begin{bmatrix} I & \tilde{V}_T^* \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \end{bmatrix} \right) \begin{bmatrix} D^* & \tilde{V}_T^* \\ 0 & \tilde{K}_T^* \end{bmatrix} \begin{bmatrix} B \\ CA \end{bmatrix} \\ &= - \begin{bmatrix} 0 & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* CA \end{bmatrix} - \begin{bmatrix} I & \\ & (I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* CC \end{bmatrix} \Phi^{-2} (D^*B + C^*MA) \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} -\Phi^{-2}(D^*B + C^*MA) \\ -(I - \tilde{K}_T^* \tilde{K}_T)^{-1} \tilde{K}_T^* C[A + C\Phi^{-2}(D^*B + C^*MA)] \end{bmatrix} \\
&= \begin{bmatrix} D_{21}^{-1}B_2 \\ C^\times(A - CD_{21}^{-1}B_2) \end{bmatrix}.
\end{aligned}$$

The controllability Gramian of Σ_{21}^\times is $\Lambda^\times = C^* \tilde{K}_T (I - \tilde{K}_T^* \tilde{K}_T)^{-2} \tilde{K}_T^* C$, which is bounded because the inverse is bounded and C^*C is bounded. According to a result of Anderson and Moore [AM, thm. 4.3] (see also [N]), if Λ^\times is bounded and $\ell_A < 1$,² then $\ell_{A^\times} < 1$. It follows that $\Sigma_{21}^{-1} \in \mathcal{U}$, so that Σ_{21} is outer. \square

Using theorem 6.1, it is straightforward to solve the lossless embedding problem.

Theorem 6.3. (Orthogonal embedding) *Let $T \in \mathcal{U}(\mathcal{M}_1, \mathcal{N}_1)$ be a locally finite input-output operator with strictly stable state realization $\mathbf{T} = \{A, B, C, D\}$. If $I - T^*T \gg 0$, or $I - T^*T \geq 0$ and \mathbf{T} is uniformly observable, and if the realization \mathbf{T} is uniformly controllable, then the lossless embedding problem has a solution $\Sigma \in \mathcal{U}(\mathcal{M}_1 \times \mathcal{N}_1, \mathcal{N}_1 \times \mathcal{N}_2)$ such that Σ is inner, $\Sigma_{11} = T$, Σ_{21} is outer, and Σ has a unitary realization Σ where A_Σ is state equivalent to A . If $A \in \mathcal{D}(\mathcal{B}, \mathcal{B}^{(-1)})$, then \mathcal{N}_2 is specified by $\#(\mathcal{N}_2) = \#(\mathcal{B}) - \#(\mathcal{B}^{(-1)}) + \#(\mathcal{M}_1)$.*

PROOF The proof is by construction. Let Σ be of the form

$$\begin{aligned}
[\Sigma_a \ \Sigma_b] &= \begin{bmatrix} A & C & C_2 \\ B & D & D_{12} \\ B_2 & D_{21} & D_{22} \end{bmatrix} \\
\Sigma &= \begin{bmatrix} R & & \\ & I & \\ & & I \end{bmatrix} [\Sigma_a \ \Sigma_b] \begin{bmatrix} [R^{(-1)}]^{-1} & & \\ & I & \\ & & I \end{bmatrix} = [\Sigma'_a \ \Sigma'_b],
\end{aligned} \tag{6.6}$$

in which $R \in \mathcal{D}(\mathcal{B}, \mathcal{B})$ is a boundedly invertible state transformation. $R, B_2, D_{12}, D_{21}, D_{22}$ are to be determined such that Σ is unitary, in which case Σ is inner (lemma 2.1).

First, determine M, B_2, D_{12} and hence Σ_a as in theorem 6.1. Because \mathbf{T} is uniformly controllable, $M \gg 0$. If we define the state transformation R by $M = R^*R$, then R is invertible, and Σ'_a is an isometry ($\Sigma_a'^* \Sigma'_a = I$). The extension of a rectangular isometric matrix to a square unitary matrix by adding columns is a standard linear algebra procedure that always has a solution. The same holds for diagonals of matrices. Hence, we can extend Σ'_a to a unitary matrix Σ , which is the realization of an inner system Σ . The resulting dimension sequence of Σ is given by $[\#(\mathcal{B}) + \#(\mathcal{M}_1) + \#(\mathcal{N}_1)]$, and the number of columns of each diagonal entry of Σ'_a is the sequence $[\#(\mathcal{B}^{(-1)}) + \#(\mathcal{N}_1)]$, hence the number of columns to be added is equal to $\#(\mathcal{N}_2) = \#(\mathcal{B}) - \#(\mathcal{B}^{(-1)}) + \#(\mathcal{M}_1)$. This number is non-negative because the columns of Σ'_a are linearly independent. \square

One difference with the time-invariant situation is that the solution of the embedding problem gives rise to a time-varying number of added extra outputs if the number of states of T is time-varying ($\mathcal{B} \neq \mathcal{B}^{(-1)}$), even if the number of inputs and outputs of T is fixed. Another difference is that, for the boundary case, we need both uniform controllability and uniform observability in order to construct an embedding. It is known that not every time-varying system admits such a realization, not even if it has a finite state dimension; the condition is that the range of H_T must be closed. See [VD1, V2].

The construction used in the proof of the theorem is computational: it can be used to recursively compute the realization of a lossless embedding from a realization of T . The recursion runs forward in time: from M_k and the realization of T at time instant k , we can compute both M_{k+1} (using recursion (1.3)) and Σ_k , the realization of Σ at instant k . An exact initial point for the recursion can be obtained for the special cases where T has zero states before a certain point n in time (take M_n to be a 0×0 matrix), or when T is time-invariant or periodic before point n in time,

²The actual condition in [AM] is that $(A - CD_{21}^{-1}B_2, D_{21}^{-1}B_2)$ is *uniformly stabilizable*, but it is also shown that this is the case if and only if $(A, D_{21}^{-1}B_2)$ is uniformly stabilizable. For this, it is sufficient that $\ell_A < 1$.

in which case M_n is the solution of the resulting algebraic Riccati equation. In other, more general cases, we can take $M_n = 0$ as an approximate initial value. It can be shown that the Riccati recursion with this initial value does not break down (because it is the exact initial value of a related system, which has zero states before time instant n and is the same as the original system after time n), and converges to the exact solution if the system is strictly stable [V2]. A proof of this is omitted for brevity.

A reformulation of theorem 6.1 and proposition 6.2 leads to the Bounded Real Lemma which appears in system and control theory.

Theorem 6.4. (Bounded Real Lemma) *Let $T \in \mathcal{U}(\mathcal{M}, \mathcal{N})$ be a bounded causal locally finite input-output operator, with strictly stable state realization $\mathbf{T} = \{A, B, C, D\}$, and $A \in \mathcal{D}(\mathcal{B}, \mathcal{B}^{(-1)})$.*

- $\|T\| < 1$ if and only if there exists $M \in \mathcal{D}(\mathcal{B}, \mathcal{B})$, $B_2 \in \mathcal{D}(\mathcal{N}, \mathcal{B}^{(-1)})$, $D_{21} \in \mathcal{D}(\mathcal{N}, \mathcal{N})$ solving

$$\begin{cases} A^*MA + B^*B + B_2^*B_2 = M^{(-1)} \\ C^*MC + D^*D + D_{21}^*D_{21} = I \\ A^*MC + B^*D + B_2^*D_{21} = 0 \end{cases} \quad (6.7)$$

with $M \geq 0$, $I - D^*D - C^*MC \gg 0$ and $\ell_{A - CD_{21}^{-1}B_2} < 1$.

- If \mathbf{T} is uniformly observable, then $\|T\| \leq 1$ if and only if (6.7) has a solution M, B_2, D_{21} such that $M \geq 0$.

PROOF The ‘only if’ part is directly derived from theorem 6.1 and proposition 6.2. The ‘if’ part is a corollary of theorem 6.1: given such M , it follows that there exists an isometric embedding Σ_a such that $\Sigma_a^*\Sigma_a = T^*T + \Sigma_{21}^*\Sigma_{21} = I$, so that $\Sigma_{21}^*\Sigma_{21} = I - T^*T \geq 0$. If in addition D_{21} is invertible and $\ell_{A - CD_{21}^{-1}B_2} < 1$, then by proposition 6.2 we can conclude that Σ_{21} is invertible, so that $I - T^*T \gg 0$, i.e., $\|T\| < 1$. \square

7. CONCLUDING REMARKS

Many control applications give rise to the Riccati equation (1.3). Usually, the existence of a stabilizing solution is of importance. In the context of our embedding problem, this would be a solution for which $A - CD_{21}^{-1}B_2$ is strictly stable, or Σ_{21} is outer. The uniqueness of such a solution is a standard result which is straightforward to prove.

While this paper was in review, more has become known on time-varying Riccati equations. We mention in particular the papers [N], in which detailed attention is paid to the convergence of the recursion to maximal/minimal solutions, and [HI], where the solution of a Kalman-Szegö-Popov-Yakubovich (KSPY) system of equations is presented. The equations (6.7) can be viewed as a particular instance of these equations. Although [HI] gives solutions to a more general class of problems, the boundary case is not considered. A major difference with [HI] is in the *proofs* of the results: whereas [HI] relies heavily on insights gained in optimal control theory, the approach taken in this paper is more based on first principles: $I - T^*T = \Sigma_{21}^*\Sigma_{21} \Leftrightarrow I - \tilde{K}_T^*\tilde{K}_T = \tilde{K}_{\Sigma_{21}}^*\tilde{K}_{\Sigma_{21}}$. The analysis of the latter equation directly leads to a recursion in which the given expressions for M, D_{21}, B_2 turn up, along with an explicit expression for the controllability operator of the realization for Σ_{21} . Similar analysis of $\tilde{K}_{\Sigma_{21}^{-1}} = (\tilde{K}_{\Sigma_{21}})^{-1}$ leads to the realization of the inverse, the given expression for the controllability operator, and the fact that our choice for Σ_{21} is outer.

A. APPENDIX: DERIVATION OF LEMMAS 5.3 AND 5.4

The contents of lemmas 5.3 and 5.4 is well known for finite matrices (see e.g., [CHM, BCHM]) for generalized inverse formulas involving Schur complements). The matrix case is readily extended to operators if the operators are assumed to have closed range. Without this condition, complications arise because the pseudo-inverses that are involved are unbounded operators.

We will repeatedly use theorem 5.2 in the following form. Let $X \geq 0$ be a bounded operator on a Hilbert space \mathcal{H} . If v is a bounded operator whose range is in $\mathcal{R}(X)$, then $v = Xv_1$, for some bounded $v_1 \in \overline{\mathcal{R}(X^*)}$ for which we can take $v_1 = X^\dagger v$. A second fact that is used in the proof of lemma 5.4 is that $X^\dagger X = \mathbf{P}_{X^*}$: the orthogonal projector onto $\overline{\mathcal{R}(X^*)}$, with domain \mathcal{H} [BR].

A.1. Proof of lemma 5.3

Suppose first that $X \geq 0$; we show that (1), (2), (3) hold. It is immediate that $A \geq 0$, $C \geq 0$. The fact that $\mathcal{R}(B) \subset \mathcal{R}(C^{1/2})$ is proven by showing that there exists λ such that $BB^* \leq \lambda C$; Douglas' theorem then implies the result. The proof is by contradiction. Suppose that there is not such a λ . Then there exists a sequence $\{x_n : n \in \mathbf{N}\}$ such that

$$(BB^*x_n, x_n) \geq n(Cx_n, x_n) > 0. \quad (\text{A.1})$$

In particular, $\|B^*x_n\| > 0$ (all n). For any u_n , $X \geq 0$ implies

$$\left(\begin{bmatrix} A & B^* \\ B & C \end{bmatrix} \begin{bmatrix} u_n \\ x_n \end{bmatrix}, \begin{bmatrix} u_n \\ x_n \end{bmatrix} \right) \geq 0,$$

i.e., $(Au_n, u_n) + (B^*x_n, u_n) + (Bu_n, x_n) + (Cx_n, x_n) \geq 0$. Choose $u_n = -\frac{1}{\sqrt{n}}B^*x_n$. Using (A.1), we obtain

$$\left(B \left\{ \frac{A}{n} - \frac{2}{\sqrt{n}} + \frac{I}{n} \right\} B^*x_n, x_n \right) \geq 0.$$

But if $n > \|I+A\|^2$, the term in braces is smaller than $-1/\sqrt{n}$, which gives a contradiction. Hence $\mathcal{R}(B) \subset \mathcal{R}(C^{1/2})$.

Define $L = C^{1/2}$ (although $L = L^*$, we will not use this), and let $B_1 = L^\dagger B$. Then B_1 is bounded, and $B = LB_1$ with $\mathcal{R}(B_1) \subset \overline{\mathcal{R}(L^*)}$, which implies

$$\mathcal{N}(B_1^*) \supset \mathcal{N}(L). \quad (\text{A.2})$$

To prove $A - B_1^*B_1 \geq 0$, we will show that

$$X = \begin{bmatrix} A & B_1^*L^* \\ LB_1 & LL^* \end{bmatrix} \geq 0 \quad \Rightarrow \quad \begin{bmatrix} A & B_1^* \\ B_1 & I \end{bmatrix} \geq 0, \quad (\text{A.3})$$

from which $A - B_1^*B_1 \geq 0$ follows directly by applying vectors of the form $\begin{bmatrix} I \\ -B_1 \end{bmatrix} a$. Thus for $x \in \mathcal{H}_1 \oplus \mathcal{H}_2$, take x of the form

$$x = \begin{bmatrix} u \\ x_1 + x_2 \end{bmatrix} \in \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{N}(L) \oplus \mathcal{R}(L^*) \end{bmatrix}$$

where $x_1 \in \mathcal{N}(L)$ and $x_2 \in \mathcal{R}(L^*)$. Note that $\mathcal{N}(L) \oplus \mathcal{R}(L^*)$ is dense in \mathcal{H}_2 . Then $\mathcal{N}(B_1^*) \supset \mathcal{N}(L)$ implies $B_1^*x_1 = 0$, while $x_2 \in \mathcal{R}(L^*)$ implies that $x_2 = L^*x'_2$, for some bounded x'_2 . Using these observations, it follows that

$$\begin{aligned} & \left(\begin{bmatrix} A & B_1^* \\ B_1 & I \end{bmatrix} \begin{bmatrix} u \\ x_1 + x_2 \end{bmatrix}, \begin{bmatrix} u \\ x_1 + x_2 \end{bmatrix} \right) \\ &= (Au, u) + (B_1^*x_1, u) + (B_1u, x_1) + (x_1, x_1) + (B_1^*x_2, u) + (B_1u, x_2) + (x_2, x_2) \\ &\geq (Au, u) + (B_1^*x_2, u) + (B_1u, x_2) + (x_2, x_2) \\ &= (Au, u) + (B^*x'_2, u) + (B_1u, x'_2) + (x'_2, x'_2) \\ &= \left(X \begin{bmatrix} u \\ x'_2 \end{bmatrix}, \begin{bmatrix} u \\ x'_2 \end{bmatrix} \right) \geq 0. \end{aligned}$$

Hence relation (A.3) holds on a dense subset of $\mathcal{H}_1 \oplus \mathcal{H}_2$. By continuity, it holds everywhere, and consequently $A - B_1^*B_1 \geq 0$.

It remains to prove the reverse implication: $X \geq 0$ if the three conditions are satisfied. Because $C \geq 0$ a decomposition of C as $C = LL^*$ is defined. Using this decomposition and $B = LB_1$,

$$X = \begin{bmatrix} A & B_1^* L^* \\ LB_1 & LL^* \end{bmatrix} = \begin{bmatrix} I & B_1^* \\ & L \end{bmatrix} \begin{bmatrix} A - B_1^* B_1 & \\ & I \end{bmatrix} \begin{bmatrix} I & \\ B_1 & L^* \end{bmatrix}.$$

Under the stated conditions, the operator

$$W = \begin{bmatrix} I & \\ & L \end{bmatrix} \begin{bmatrix} I & B_1^* \\ & I \end{bmatrix} \begin{bmatrix} (A - B_1^* B_1)^{1/2} & \\ & I \end{bmatrix} \quad (\text{A.4})$$

is well defined, and is a factor of X such that $X = WW^*$. Hence $X \geq 0$. \square

A.2. Proof of lemma 5.4

Let $X \geq 0$ have a factorization $X = WW^*$, then $\mathcal{R}(X^{1/2}) = \mathcal{R}(W)$ (by theorem 5.2). It can be inferred from Beutler and Root [BR] that $X^\dagger = W^{*\dagger}W^\dagger = X^{\dagger/2}X^{\dagger/2}$, hence if $\mathcal{R}(v) \subset \mathcal{R}(X^{1/2}) = \mathcal{R}(W)$, then v_1 and v_2 defined by

$$\begin{aligned} v_1 &= X^{\dagger/2}v, & \mathcal{R}(v_1) &\subset \overline{\mathcal{R}(X^{1/2})} \\ v_2 &= W^\dagger v, & \mathcal{R}(v_2) &\subset \overline{\mathcal{R}(W^*)} \end{aligned}$$

are bounded, and³ $v_1^*v_1 = v_2^*v_2$.

Let $L = C^{1/2}$, $B_1 = L^\dagger B$ and put W as in (A.4), so that $X = WW^*$. Define the operator W^\ddagger by

$$W^\ddagger = \begin{bmatrix} (A - B_1^* B_1)^{\dagger/2} & \\ & I \end{bmatrix} \begin{bmatrix} I & -B_1^* \\ & I \end{bmatrix} \begin{bmatrix} I & \\ & L^\dagger \end{bmatrix}.$$

We will prove that $W^\ddagger = W^\dagger$ on $\mathcal{R}(W)$. The result will be, for a bounded operator v with $\mathcal{R}(v) \subset \mathcal{R}(X^{1/2}) = \mathcal{R}(W)$, that $W^\dagger v = W^\ddagger v$, so that $v_1 := X^{\dagger/2}v$ and $v_2 := W^\dagger v$ are bounded and satisfy $v_1^*v_1 = v_2^*v_2$. Indeed, for any v with range in $\mathcal{R}(W)$ we have that the operator $v_1 = W^\dagger v$ is bounded and such that $v = Wv_1$. Hence $W^\ddagger v = W^\ddagger Wv_1 = W^\dagger Wv_1 = W^\dagger v$, so that $W^\ddagger = W^\dagger$ on $\mathcal{R}(W)$ if and only if

$$W^\ddagger W = W^\dagger W \quad \text{on } \overline{\mathcal{R}(W^*)}.$$

To analyze $W^\ddagger W$, we first prove that $B_1^* - B_1^* L^\dagger L = 0$. Indeed, if $x \in \mathcal{N}(L)$ then $x \in \mathcal{N}(B_1^*)$ (by equation (A.2)), and hence both $B_1^* x = 0$ and $Lx = 0$. If, on the other hand, $x \in \mathcal{N}(L)^\perp$, then $L^\dagger Lx = x$ since $L^\dagger L$ is the projector onto $\mathcal{N}(L)^\perp$, and hence $B_1^* L^\dagger Lx = B_1^* x$.

With the definition of W^\ddagger and the above result,

$$\begin{aligned} W^\ddagger W &= \begin{bmatrix} (A - B_1^* B_1)^{\dagger/2} & \\ & I \end{bmatrix} \begin{bmatrix} I & -B_1^* \\ & I \end{bmatrix} \begin{bmatrix} I & \\ & L^\dagger \end{bmatrix} \begin{bmatrix} I & \\ & L \end{bmatrix} \begin{bmatrix} I & B_1^* \\ & I \end{bmatrix} \begin{bmatrix} (A - B_1^* B_1)^{1/2} & \\ & I \end{bmatrix} \\ &= \begin{bmatrix} (A - B_1^* B_1)^{\dagger/2} & \\ & I \end{bmatrix} \begin{bmatrix} I & B_1^* - B_1^* L^\dagger L \\ & L^\dagger L \end{bmatrix} \begin{bmatrix} (A - B_1^* B_1)^{1/2} & \\ & I \end{bmatrix} \\ &= \begin{bmatrix} (A - B_1^* B_1)^{\dagger/2} (A - B_1^* B_1)^{1/2} & \\ & L^\dagger L \end{bmatrix} =: \begin{bmatrix} \mathbf{P}_1 & \\ & \mathbf{P}_2 \end{bmatrix}. \end{aligned}$$

\mathbf{P}_1 and \mathbf{P}_2 are projectors onto $\overline{\mathcal{R}(A - B_1^* B_1)^{1/2}}$ and $\overline{\mathcal{R}(L^*)}$, respectively. Now, using

$$W^* = \begin{bmatrix} (A - B_1^* B_1)^{1/2} & \\ & I \end{bmatrix} \begin{bmatrix} I & \\ B_1 & L^* \end{bmatrix}$$

³We are careful here not to write $X^\dagger v$. Although $\overline{\mathcal{R}(X)} = \overline{\mathcal{R}(X^{1/2})}$, we only have that $\mathcal{R}(X) \subset \mathcal{R}(X^{1/2})$, and hence $X^\dagger v$ can be unbounded with $\mathcal{R}(v) \in \mathcal{R}(X^{1/2})$.

and $\mathcal{R}(B_1) \subset \overline{\mathcal{R}(L^*)}$, we have that

$$\overline{\mathcal{R}(W^*)} \subset \overline{\mathcal{R}} \left[\begin{array}{c} (A - B_1^* B_1)^{1/2} \\ L^* \end{array} \right].$$

Since $W^\dagger W$ is the projector onto $\overline{\mathcal{R}(W^*)}$, and $W^\ddagger W$ is the projector onto the range at the right hand side of the expression, this proves that $W^\ddagger W = W^\dagger W$ on $\overline{\mathcal{R}(W^*)}$, as required. Hence $W^\ddagger = W^\dagger$ on $\mathcal{R}(W)$, which also implies that W^\ddagger is well-defined on $\mathcal{R}(W)$. \square

REFERENCES

- [A] W. Arveson, Interpolation problems in nest algebras, *J. Functional Anal.*, **20** (1975), 208–233.
- [AD] D. Alpay and P. Dewilde, Time-varying signal approximation and estimation, *Signal Processing, Scattering and Operator Theory, and Numerical Methods* (M.A. Kaashoek, J.H. van Schuppen, and A.C.M. Ran, eds.), Proc. Int. Symp. MTNS-89, vol. III, pp. 1–22. Birkhäuser Verlag, 1990.
- [ADD] D. Alpay, P. Dewilde, and H. Dym, Lossless Inverse Scattering and reproducing kernels for upper triangular operators, *Extension and Interpolation of Linear Operators and Matrix Functions* (I. Gohberg, ed.), Operator Theory, Advances and Applications, vol. 47, pp. 61–135. Birkhäuser Verlag, 1990.
- [AHD] B.D.O. Anderson, K.L. Hitz, and N.D. Diem, Recursive algorithm for spectral factorization, *IEEE Trans. Circuits Syst.*, **21** (1974), 742–750.
- [AM] B.D.O. Anderson and J.B. Moore, Detectability and stabilizability of time-varying discrete-time linear systems, *SIAM J. Control and Optimization*, **19** (1981), 20–32. Comments in *IEEE Trans. Automat. Control*, vol. 37, no. 3, 1992, pp. 409–410.
- [AV] B.D.O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis*, Prentice Hall, 1973.
- [BCHM] F. Burns, D. Carlson, E. Haynsworth, and T. Markham, Generalized inverse formulas using Schur complements, *SIAM J. Applied Math.*, **26** (1974), 254–259.
- [BLW] S. Bittanti, A.J. Laub, and J.C. Willems, eds. *The Riccati Equation*, Comm. Control Eng. Series. Springer Verlag, 1991.
- [BR] F.J. Beutler and W.L. Root, The operator pseudo-inverse in control and systems identification, *Generalized Inverses and Applications* (M. Zuhair Nashed, ed.), pp. 397–494. Academic Press, 1976.
- [CHM] D. Carlson, E. Haynsworth, and T. Markham, A generalization of the Schur complement by means of the Moore-Penrose inverse, *SIAM J. Applied Math.*, **26** (1974), 169–175.
- [D1] U.B. Desai, A state-space approach to orthogonal digital filters, *IEEE Trans. Circuits Syst.*, **38** (1991), 160–169.
- [D2] P. Dewilde, Input-output description of roomy systems, *SIAM J. Control Optim.*, **14** (1976), 712–736.
- [D3] R.G. Douglas, On majorization, factorization and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.*, **17** (1966), 413–415.
- [DD1] E. Deprettere and P. Dewilde, Orthogonal cascade realization of real multiport digital filters, *Circuit Theory and Appl.*, **8** (1980), 245–272.
- [DD2] P. Dewilde and H. Dym, Interpolation for upper triangular operators, *Time-Variant Systems and Interpolation* (I. Gohberg, ed.), Operator Theory: Advances and Applications, vol. 56, pp. 153–260. Birkhäuser Verlag, 1992.
- [DV] P.M. Dewilde and A.J. van der Veen, On the Hankel-norm approximation of upper-triangular operators and matrices, *Integral Eq. Operator Th.*, **17** (1993), 1–45.

- [HI] A. Halanay and V. Ionescu, Generalized discrete-time Popov-Yakubovich theory, *Systems & Control Letters*, **20** (1993), 1–6.
- [N] G. De Nicolao, On the time-varying Riccati difference equation of optimal filtering, *SIAM J. Control and Optimization*, **30** (1992), 1251–1269.
- [V1] P.P. Vaidyanathan, The discrete-time Bounded-Real Lemma in digital filtering, *IEEE Trans. Circuits Syst.*, **32** (1985), 918–924.
- [V2] A.J. van der Veen, *Time-Varying System Theory and Computational Modeling: Realization, Approximation, and Factorization*, PhD thesis, Delft University of Technology, Delft, The Netherlands, 1993.
- [VD1] A.J. van der Veen and P.M. Dewilde, Time-varying system theory for computational networks, *Algorithms and Parallel VLSI Architectures, II* (P. Quinton and Y. Robert, eds.), pp. 103–127. Elsevier, 1991.
- [VD2] A.J. van der Veen and P.M. Dewilde, Orthogonal embedding theory for contractive time-varying systems, *Recent Advances in Mathematical Theory of Systems, Control, Networks and Signal Processing II (Proc. Int. Symp. MTNS-91)* (H. Kimura and S. Kodama, eds.), pp. 513–518. MITA Press, Japan, 1992.
- [VD3] A.J. van der Veen and P.M. Dewilde, Time-varying computational networks: Realization, orthogonal embedding and structural factorization, *Integration, the VLSI journal*, **16** (1993), 267–291.
- [VD4] A.J. van der Veen and P.M. Dewilde, On low-complexity approximation of matrices, *Linear Algebra and its Applications*, **205/206** (1994), 1145–1202.