

Time-Varying System Theory for Computational Networks

Alle-Jan van der Veen and Patrick Dewilde

Many computational schemes in Linear Algebra can be studied from the point of view of Time-Varying Linear Systems theory. This approach not only puts a variety of results in a unified framework, but also generates new and unexpected results such as strong approximations of operators or matrices by computational networks of low complexity, and the embedding of contractive operations in orthogonal computations. In the present paper we develop the required Time-Varying System Theory in a systematic way, and derive a Kronecker or Ho-Kalman type realization method.

1. INTRODUCTION

Consider the computations schematically represented in figure 1. The unfolded or expanded version is shown in fig. 1(a). At each instant of time the computation takes in some input data from an input sequence U and computes new output data which is part of the output sequence Y generated by the processor. To execute the computation, the processor will use some remainder of its past history known as the state, which it had temporarily stored in registers indicated by the symbol Z . We shall limit ourselves to the case where the computations are indeed linear — although it will become clear that the theory cries for extensions.

Characteristic for the computations of a real life system is that it can use only a finite amount of data at any given time in its history. This then easily leads to a recursive model, which is obtained by folding the computation and feeding back the state. This computational model is grasped mathematically by introducing a state sequence X with entries X_i , which are vector quantities (the dimension of each vector equals the number of state variables and need not be constant in time). At time instant i , the mapping from the current input and state to the output

⁰Department of Electrical Engineering, Delft University of Technology, Delft, The Netherlands. In “Algorithms and Parallel VLSI Architectures”, ed. P. Quinton and Y. Robert, vol. II (1991), ch. 11, pp. 103-126.

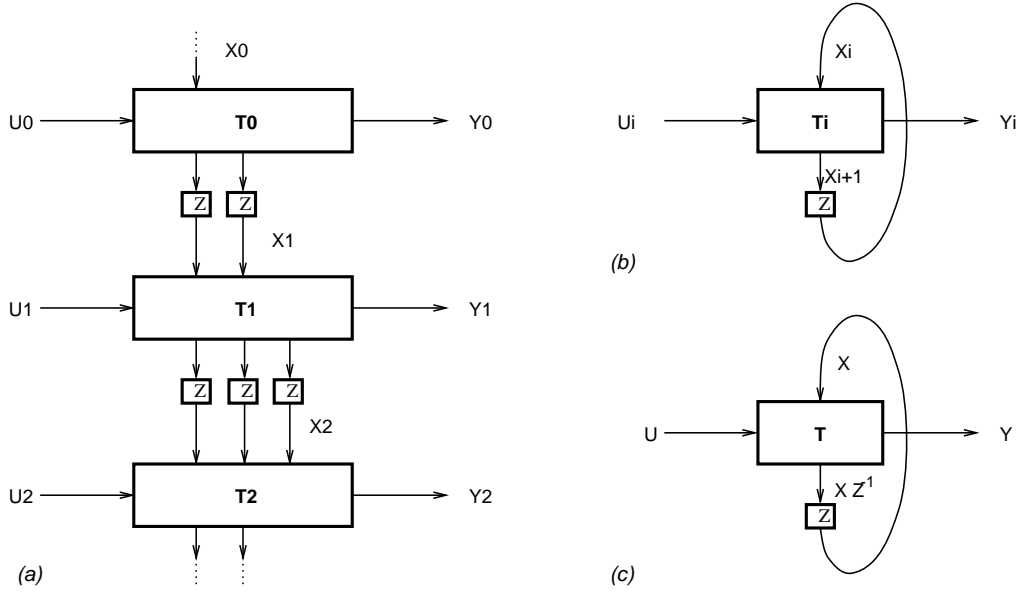


Figure 1. Time-varying state space representations: (a) fully expanded network, (b) time-varying network, (c) compact representation with diagonal operators.

and next state is a linear “memoryless” mapping, \mathbf{T}_i say, such that $[X_{i+1} \ Y_i] = [X_i \ U_i] \mathbf{T}_i$, or written more explicitly,

$$\begin{cases} X_{i+1} = X_i A_i + U_i B_i \\ Y_i = X_i C_i + U_i D_i \end{cases} \quad \mathbf{T}_i = \begin{bmatrix} A_i & C_i \\ B_i & D_i \end{bmatrix}.$$

It is thus possible to associate with the fully expanded network in figure 1(a) a time-varying network, depicted in fig. 1(b), in which the parameters of the mapping can be changed at each time instant. This is the conventional way in which time-varying networks are treated. We will, instead, pursue what we will call a network description corresponding to fig. 1(c), in which the network operates on the full sequences U and X , and produces the output Y and next state XZ^{-1} , where

$$\begin{aligned} X &= [\cdots \ X_{-1} \ \boxed{X_0} \ X_1 \ X_2 \ \cdots] \\ XZ^{-1} &= [\cdots \ X_0 \ \boxed{X_1} \ X_2 \ X_3 \ \cdots]. \end{aligned}$$

This defines Z as the right-shift operator for sequences. It is thus possible to write the state space description as

$$\begin{cases} XZ^{-1} = XA + UB \\ Y = XC + UD \end{cases} \quad \mathbf{T} = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$$

such that $A = \text{diag}[\cdots \ A_{-1} \ \boxed{A_0} \ A_1 \ \cdots]$ is a diagonal that operates on the sequence X as a direct (“instantaneous”) multiplier. B, C, D are defined likewise. In effect, figure 1(c) provides a compact operator notation for the same computations as in figure 1(a/b), and uses only multiplications by diagonals and the shift operator.

The purpose of the computational scheme defined so far is usually to perform a desired

transformation on the input sequence U with the output sequence Y as the result. Because of causality (we assume that series are represented by rows and that operators act from left to right), such a transformation will be represented by what we call an “upper” operator T . An attractive (and physical) mathematical framework is obtained if the input sequences are constrained to have finite energy, and T to be a bounded operator on such sequences. Hence we focus on bounded upper operators mapping ℓ_2 -sequences U to ℓ_2 -sequences Y via

$$Y = UT.$$

With $U = [\cdots U_{-1} \boxed{U_0} U_1 U_2 \cdots]$, and Y likewise, we will identify T with its (doubly-infinite) matrix representation

$$T = \begin{bmatrix} \ddots & \ddots & & \ddots & & \ddots \\ & T_{-1,-1} & T_{-1,0} & T_{-1,1} & T_{-1,2} & \\ & & \boxed{T_{00}} & T_{01} & T_{02} & \ddots \\ & & & T_{11} & T_{12} & \\ & \mathbf{0} & & & T_{22} & \ddots \\ & & & & & \ddots \end{bmatrix}.$$

(The square identifies the 00-th entry of the matrix.) If T is viewed as the transfer operator of a non-stationary causal linear system with input U and corresponding output Y then the i -th row of T corresponds to the impulse response of the system when excited at time instant i . For time-invariant systems, all elements on the diagonals of T are the same, and T is said to have a Toeplitz structure. In the time-invariant case often more is known about T than just an operator representation, namely a description as rational transfer function or an equivalent state space description. This then allows to perform applications of T with a finite number of operations. It is the purpose of this paper to study these kinds of representations for time-varying systems.

The connection between the I/O operator $T : Y = UT$ and the state space description follows from the expansion

$$Y = UD + UBZC + UB(ZA)ZC + UB(ZA)(ZA)ZC + \cdots$$

and can be written as

$$T = D + B(I - ZA)^{-1} ZC.$$

provided the inverse in the formula is meaningful. As in the time-invariant case, one might wonder conversely whether there exists, for a given transfer operator T , a state space realization \mathbf{T} that realizes the same transfer, yet has this advantage to be finitely computable. This question is known as the identification problem, and will be the topic of this paper. A related problem is the model reduction problem, in which to a given T an approximating state space description of low complexity is pursued. Since the number of state variables need not be constant in time, but can increase and shrink, it is seen that in this respect the time-varying realization theory is much richer, and that the approximation accuracy can be varied in time

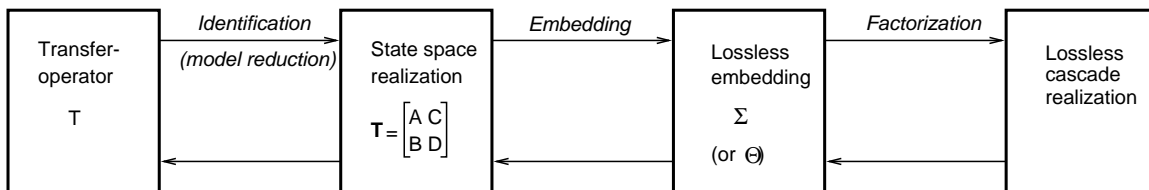


Figure 2. Overview of the cascade synthesis problem

at will. It is also possible to choose the number of state variables to be zero outside a region of interest in time, and to incorporate in this way upper triangular matrices of finite size into the time-varying context.

Context

The present paper is based on the research published in [1, 2, 3, 4], in which the theory of a generalization of the z -transform for upper non-commutative operators, called the W -transform, was developed and the interpolating properties of lossless time-varying (or non-stationary) systems represented by these operators was investigated. Most notation in the present paper is adopted from these papers, although it is generalized to furnish time-varying state dimensions.

Starting in the 1950's (or even earlier), time-varying network and state space theory and extensions of important system theoretic notions to the time varying case have been discussed by many authors. While most of the early work is on time-continuous linear systems and differential equations with time-varying coefficients (see *e.g.*, [5] for a 1960 survey), time-discrete systems have gradually come into favor. Some important more recent approaches that parallel the state space realization part of the present paper are the monograph by Feintuch/Saeks [6], in which a Hilbert resolution space setting is taken, and recent work by Kamen *et al.* [7, 8], where time varying systems are put into an algebraic framework of polynomial rings. However, many results in particular on controllability, detectability, stabilizability *etc.* have been discussed by many authors without using these specialized mathematical means (see *e.g.*, Anderson/Moore [9] and references therein), by time-indexing the state space matrices $\{A, B, C, D\}$ and deriving expressions (iterations) in terms of these matrices. There is usually a one-to-one correspondence between these expressions and their equivalent in our notation.

A Tour of the Results

The results obtained so far are depicted in Fig. 2 and summarized below. The present paper deals with item 1. Item 2 has been published in [10], while item 3 will be the subject of a separate treatment still to be published.

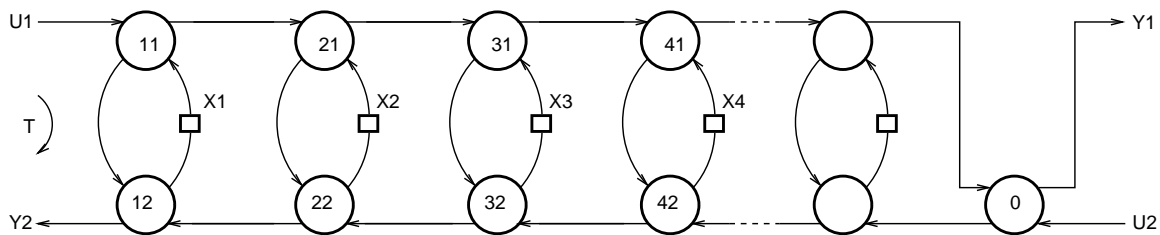


Figure 3. Cascade realization of the lossless embedding.

1. *Identification.* Proper definitions of shift-invariant input and output state spaces of the system T are possible. By selecting a (strong) basis in either of these spaces, minimal $\{A, B, C, D\}$ -realizations can be computed. In addition, these spaces define a Hankel operator that maps the input state space to the output state space. We shall prove a Kronecker [11] or Ho-Kalman [12] type theorem which shows that the system order is equal to the rank of the Hankel operator. Moreover, a diagonal expansion of the Hankel operator reveals its relation to the given data in T , which will in turn lead to an identification scheme that has a close resemblance to subspace identification techniques for time-invariant systems, and that can also be used to find solutions to model reduction problems at a later stage.
2. *Embedding.* If T corresponds to a system that is inner (respect. J -inner), then selecting a (J)-orthogonal basis in either the input or output state space will yield an orthogonal (or lossless) realization. If, on the other hand, the given T is not inner but contractive, we show that a realization of T can be extended (by adding an extra input and output, and supplementing states where needed) to yield an orthogonal realization that “embeds” the given system in the sense that T will be the transfer operator from one input to one output if the other inputs are put to zero.
3. *Factorization.* Finally, it is possible to factor an orthogonal multiport realization matrix into a minimal number of elementary (2×2) orthogonal operations. Corresponding to this factorization is a network structure that consists of a cascade of elementary lossless sections, as in figure 3. In this figure, the embedded transfer operator T is the transfer from input U_1 to output Y_2 , when the extra input U_2 is zero.

2. NOTATION, SETTING AND MATHEMATICAL PRELIMINARIES

Spaces

We consider a generalization of ℓ_2 sequences

$$X = \left[\cdots \quad X_{-1} \quad \boxed{X_0} \quad X_1 \quad \cdots \right],$$

in which each of the entries X_i is an element of a (row) vector space \mathbf{C}^{N_i} , with varying dimensions $N_i \in \mathbf{N}$, and such that the total energy $\|X\|_2^2 = \sum_{-\infty}^{\infty} \|X_i\|_2^2$ is bounded. In

the above expression, the square identifies the position of the 0-th entry. We denote the set $(\mathbf{Z} \rightarrow \mathbf{N})$ of index sequences by \mathcal{I} , and with $N \in \mathcal{I}$ say that the above X is an element of $\ell_2(\mathbf{C}^N)$, or \mathbf{C}_2^N for brevity. We adopt the shorthand “ $\cdot n$ ” for the index sequence N with all N_i equal to n . Hence, *e.g.*, $\mathbf{C}_2^{\cdot 1}$ is the set of the usual ℓ_2 sequences.

Let $N, P \in \mathcal{I}$. Following [1], we denote by $\mathcal{X}(\mathbf{C}^N, \mathbf{C}^P)$ the class of bounded operators $(\mathbf{C}_2^N \rightarrow \mathbf{C}_2^P)$. *E.g.*, a system transfer operator with n_1 input ports and n_0 output ports is an operator in $\mathcal{X}(\mathcal{N}^{\cdot 1}, \mathcal{P}^{\cdot 1})$, with $\mathcal{N} = \mathbf{C}^{n_1}$ and $\mathcal{P} = \mathbf{C}^{n_0}$. An operator $A \in \mathcal{X}(\mathbf{C}^N, \mathbf{C}^P)$ may be represented by a doubly infinite matrix with entries $A_{ij} : \mathbf{C}^{N_i} \rightarrow \mathbf{C}^{P_j}$, and may as well be represented by the shorthand $A_{N \times P}$. For example, with $N = [\cdots 1 \boxed{3} 2 \cdots]$ and $P = [\cdots 2 \boxed{1} 3 \cdots]$, $X \in \mathbf{C}_2^N$ will have the form

$$X = \left[\cdots \square \quad \square\square\square \quad \square\square \quad \cdots \right].$$

and $A \in \mathcal{X}(\mathbf{C}^N, \mathbf{C}^P)$

$$A = \begin{bmatrix} \ddots & \vdots & & \vdots & & \\ \cdots & \square\square & \square & \square\square\square & \cdots & \\ & \square\square & \square & \square\square\square & & \\ & \square\square & \square & \square\square\square & & \\ \cdots & \square\square & \square & \square\square\square & \cdots & \\ & \vdots & & \vdots & \ddots & \end{bmatrix}.$$

where in this case each box represents a complex number. We think of A as acting on row inputs and producing row outputs. The i -th row of A will be in $\ell_2(\mathbf{C}^{N_i}, \mathbf{C}^P) = (\mathbf{C}_2^P)^{N_i}$ and will be bounded by $\|A\|$. The converse is certainly not true, as can be seen when A is Toeplitz and upper triangular, for in that case A will correspond to a classical ℓ_2 system, and boundedness in ℓ_2 of the impulse response is known not to be a sufficient condition for boundedness of the system transfer operator.

Shifts and Constructors

For every index sequence $N \in \mathcal{I}$, $N = [\cdots N_{-1} \boxed{N_0} N_1 \cdots]$ we indicate the k -th shift by

$$N^{(k)} = \left[\cdots N_{-k-1} \boxed{N_{-k}} N_{-k+1} \cdots \right].$$

We will use the shorthand X^+ for $X^{(1)}$, and likewise $X^- = X^{(-1)}$.

We define the shift operator $Z : \mathbf{C}^N \rightarrow \mathbf{C}^{N^+}$ as

$$(XZ)_i = X_{i-1}.$$

The shift operator is of course bounded. It is even unitary, meaning that $Z_{N \times N^+} (Z^*)_{N^+ \times N} = I_{N \times N}$, and $(Z^*)_{N^+ \times N} Z_{N \times N^+} = I_{N^+ \times N^+}$. We denote by $Z^{[k]}$ the k -times repeated application of Z :

$$Z^{[k]} = Z_{N \times N^+} Z_{N^+ \times N^{(2)}} \cdots Z_{N^{(k-1)} \times N^{(k)}}.$$

(Note that formally Z^k is not well defined because dimensions do not match. Nonetheless, we will in the next sections usually suppress dimension information in formulas and just write Z^k when we mean $Z^{[k]}$.) The entries Z_{ij} of Z satisfy

$$\begin{aligned} Z_{ij} &= I_{N_i \times N_i} & j = i + 1, \\ &= 0 & \text{otherwise,} \end{aligned}$$

and hence Z can be pictured as the infinite size matrix

$$Z = \begin{bmatrix} \ddots & \vdots & & & \vdots & & \\ \cdots & 0 & I_{N_{-1} \times N_{-1}} & & 0 & \cdots & \\ & & \boxed{0} & I_{N_0 \times N_0} & & & \\ & & & 0 & I_{N_1 \times N_1} & & \\ \cdots & 0 & & & 0 & \cdots & \\ & \vdots & & & \vdots & \ddots & \end{bmatrix}.$$

Following [1], we define the operators

$$\pi : u \in \mathbf{C}^{N_0} \rightarrow f \in \mathbf{C}^N : \begin{cases} f_0 = u \\ f_i = 0, & i \neq 0 \end{cases}$$

with adjoint

$$\pi^* : f \in \mathbf{C}^N \rightarrow u \in \mathbf{C}^{N_0} : u = f_0.$$

For the entry (i, j) of the matrix representation of an operator A we may write

$$A_{ij} = \pi^* Z^{[i]} \cdot A \cdot Z^{[j]} \pi,$$

and define an operator $\pi_i^* = \pi^* Z^{[i]}$ to select the i -th row of its operand accordingly. Next, we define the k -th diagonal shift on $A \in \mathcal{X}(\mathbf{C}^N, \mathbf{C}^P)$ by

$$A^{(k)} = Z^{*[k]} A Z^{[k]},$$

which will be in $\mathcal{X}(\mathbf{C}^{N^{(k)}}, \mathbf{C}^{P^{(k)}})$. We adopt, as with index sets, the shorthand A^+ for $A^{(+1)}$, A^- for $A^{(-1)}$. Hence $(A^+)_ij = A_{i-1, j-1}$.

Spaces for Upper, Lower and Diagonal Operators

As in [1] we define subset of upper, lower and diagonal operators in \mathcal{X} as

$$\begin{aligned} \mathcal{U} &= \{A \in \mathcal{X} : A_{ij} = 0, i > j\} \\ \mathcal{L} &= \{A \in \mathcal{X} : A_{ij} = 0, i < j\} \\ \mathcal{D} &= \mathcal{U} \cap \mathcal{L}. \end{aligned}$$

For $A \in \mathcal{D}$, “ A_i ” will serve as shorthand for the entry A_{ii} , and we shall write

$$A = \text{diag} \left[\cdots \quad A_{-1} \quad \boxed{A_0} \quad A_1 \quad \cdots \right] = \text{diag}(A_i).$$

Let $A \in \mathcal{X}$. We define the j -th diagonal $A_{[j]} \in \mathcal{D}$ of A by

$$(A_{[j]})_i = A_{i-j, i}.$$

Hence $A_{[0]}$ is the main diagonal of the operator A , and for positive j , $A_{[j]}$ is the j -th subdiagonal above $A_{[0]}$. With this notation, A can formally be written in terms of its diagonals as

$$A = \sum_{j=-\infty}^{\infty} Z^{[j]} A_{[j]},$$

although this expression need not converge at all. A class of operators that do allow this representation are the Hilbert-Schmidt operators [1]:

$$\mathcal{X}_2 = \{A \in \mathcal{X} : \|A\|_{HS}^2 = \sum_{i,j} \|A_{ij}\|_2^2 < \infty\}$$

along with inner product $\langle A, B \rangle = \text{trace}(AB^*)$, and norm $\|A\|_{HS}^2 = \langle A, A \rangle = \text{trace}(AA^*)$.

A subset \mathcal{H} in \mathcal{X}_2 is a left D -invariant subspace in \mathcal{X}_2 if

$$A \in \mathcal{H}, B \in \mathcal{H} \Rightarrow D_1 A + D_2 B \in \mathcal{H} \quad \text{all } D_{1,2} \in \mathcal{D}.$$

We can define orthogonal projectors $\mathbf{P}_{\mathcal{H}}$ onto these subspaces, according to the natural Hilbert-Schmidt metric. Standard subspaces are

$$\begin{aligned} \mathcal{U}_2 &= \mathcal{U} \cap \mathcal{X}_2 \\ \mathcal{L}_2 &= \mathcal{L} \cap \mathcal{X}_2 \\ \mathcal{D}_2 &= \mathcal{L}_2 \cap \mathcal{U}_2 \end{aligned}$$

and standard projectors that go with these spaces are $\mathbf{P}_0 = \mathbf{P}_{\mathcal{D}_2}$ and $\mathbf{P} = \mathbf{P}_{\mathcal{U}_2}$:

$$\begin{aligned} \mathbf{P}_0 : \mathcal{X}_2 &\rightarrow \mathcal{D}_2 : & \mathbf{P}_0(A) &= A_{[0]}, \\ \mathbf{P} : \mathcal{X}_2 &\rightarrow \mathcal{U}_2 : & \mathbf{P}(A) &= \sum_{j=0}^{\infty} Z^{[j]} A_{[j]}. \end{aligned}$$

It is a fundamental fact (and proven in [1]) that

$$\mathcal{X}_2 = \mathcal{L}_2 Z^{-1} \oplus \mathcal{D}_2 \oplus \mathcal{U}_2 Z,$$

where “ \oplus ” indicates orthogonal composition of spaces.

Diagonal Inner Product

For $A, B \in \mathcal{X}_2$, define the “diagonal” or “brace” inner product $\{A, B\}$ as

$$\{A, B\} = \mathbf{P}_0(AB^*)$$

It follows that, with $A, B \in \mathcal{X}_2$, $\{A, B\} \in \mathcal{D}_2$, and that $\langle A, B \rangle = \text{trace}\{A, B\}$. In particular, we have that

$$\begin{aligned} A = 0 &\Leftrightarrow \langle A, A \rangle = 0 \Leftrightarrow \{A, A\} = 0 \\ D_1 \langle A, B \rangle D_2 = 0 \quad (\text{all } D_{1,2} \in \mathcal{D}) &\Leftrightarrow \{A, B\} = 0 \end{aligned}$$

so that orthogonality of left D -invariant subspaces is the same in each of these inner products. The observation that the diagonal inner product does not render a single number but rather a full diagonal of rowwise inner products will be useful in the determination of projections onto subspaces. The above expressions show that two left D -invariant subspaces are orthogonal iff they are orthogonal rowwise.

Basis Representations of Subspaces

Let \mathcal{H} be a left D -invariant subspace in \mathcal{X}_2 . Because of the *left* D -invariance, \mathcal{H} falls naturally apart into “slices” $\mathcal{H}_i = \pi_i^* \mathcal{H}$. Each such \mathcal{H}_i is an ordinary subspace in ℓ_2 . If each of these subspaces is finite dimensional, say $\dim(\mathcal{H}_i) = N_i$, then we shall say that \mathcal{H} is of local finite dimension. Each \mathcal{H}_i has a finite orthonormal basis $\{(q_i)_1, \dots, (q_i)_{N_i}\}$, with $(q_i)_k \in \ell_2$, and hence is generated by a sequence $\mathbf{Q}_i \in \ell_2(\mathbf{C}^{N_i})$ whose rows are the $(q_i)_k$ ($k = 1 \dots N_i$), such that

$$\mathcal{H}_i = \{D_i \mathbf{Q}_i : D_i \in \mathbf{C}^{1 \times N_i}\},$$

and $\mathbf{Q}_i \mathbf{Q}_i^* = I_{N_i \times N_i}$. Stack the \mathbf{Q}_i to arrive at an operator \mathbf{Q} whose i -th row $\pi_i^* \mathbf{Q}$ is \mathbf{Q}_i . This \mathbf{Q} is not necessarily a bounded ($\mathcal{X}_2 \rightarrow \mathcal{X}_2$) operator, but with domain restricted to \mathcal{D}_2 it is a bounded operator in ($\mathcal{D}_2 \rightarrow \mathcal{X}_2$) — in fact, an isometry — with range \mathcal{H} : $\mathcal{H} = \mathcal{D}_2(\mathbf{C}^1, \mathbf{C}^N) \cdot \mathbf{Q}$. In addition, \mathbf{Q} is orthonormal in the sense that $\Lambda_{\mathbf{Q}} := \mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*) = I_{N \times N}$.

(*Remark.* Since \mathbf{Q} need not be a bounded \mathcal{X}_2 operator, but is known to be a bounded ($\mathcal{D}_2 \rightarrow \mathcal{X}_2$)-operator, the value of an expression like $\mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*)$ should be interpreted as $\mathbf{P}_0(D\mathbf{Q}\mathbf{Q}^*) = D\mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*)$, for all $D \in \mathcal{D}_2$. Technically speaking, the “ \mathbf{P}_0 ” in $\mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*)$ could be dropped as $\mathbf{Q}^* : \mathcal{X}_2 \rightarrow \mathcal{D}_2$ already, but then the notation would lead to confusion and not be compatible with the previous cases, especially since the domain of \mathbf{Q} can be extended. In this respect, if $X \in \mathcal{X}_2$ then the product $X\mathbf{Q}^*$ is interpreted as $X\mathbf{Q}^* = \sum Z^{[k]} \mathbf{P}_0(Z^{-k}X\mathbf{Q}^*)$, which is compatible with the usual definition when $\mathbf{Q} \in \mathcal{X}$.)

The above construction is summarized in the following proposition:

Proposition 1. *If a left D -invariant subspace \mathcal{H} in \mathcal{X}_2 has finite local dimension $N \in \mathcal{I}$, then there exists an operator \mathbf{Q} , bounded in ($\mathcal{D}_2(\mathbf{C}^1, \mathbf{C}^N) \rightarrow \mathcal{X}_2$), with $\Lambda_{\mathbf{Q}} = \mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*) = I$, such that*

$$\mathcal{H} = \mathcal{D}_2(\mathbf{C}^1, \mathbf{C}^N) \cdot \mathbf{Q}.$$

\mathbf{Q} is said to be an orthonormal basis representation of \mathcal{H} .

More generally, let \mathbf{F} be a bounded ($\mathcal{D}_2 \rightarrow \mathcal{X}_2$)-operator such that

$$\mathcal{H} = \mathcal{D}_2(\mathbf{C}^1, \mathbf{C}^N) \cdot \mathbf{F}_{N \times 1},$$

and $\Lambda_{\mathbf{F}} = \mathbf{P}_0(\mathbf{F}\mathbf{F}^*) \in \mathcal{D}(\mathbf{C}^N, \mathbf{C}^N)$ is uniformly positive (meaning that $\Lambda_{\mathbf{F}}$ is boundedly invertible as well; we write $\Lambda_{\mathbf{F}} \gg 0$). Then also \mathbf{F} generates \mathcal{H} and is said to be a strong basis representation of \mathcal{H} . A Gram-Schmidt orthogonalization on each of the rows \mathbf{F}_i will yield $\mathbf{F} = R\mathbf{Q}$, where \mathbf{Q} is an orthonormal basis representation of \mathcal{H} , and $R \in \mathcal{D}(\mathbf{C}^N, \mathbf{C}^N)$ is a boundedly invertible positive factor of $\Lambda_{\mathbf{F}}$, since $\Lambda_{\mathbf{F}} = \mathbf{P}_0(\mathbf{F}\mathbf{F}^*) = \mathbf{P}_0(R\mathbf{Q}\mathbf{Q}^*R^*) = RR^* \gg 0$.

Projection onto Subspaces

Lemma 2. *If \mathcal{H} is a subspace in \mathcal{X}_2 , generated by an orthonormal basis representation \mathbf{Q} , then (for $X \in \mathcal{X}_2$),*

$$X \perp \mathcal{H} \quad \Leftrightarrow \quad \mathbf{P}_0(X\mathbf{Q}^*) = 0.$$

PROOF Any Y in \mathcal{H} can be written as $Y = D\mathbf{Q}$, for some $D \in \mathcal{D}_2$. Then $X \perp Y \Leftrightarrow \{X, Y\} = \mathbf{P}_0(XY^*) = 0$, and $\mathbf{P}_0(XY^*) = \mathbf{P}_0(X\mathbf{Q}^*D^*) = \mathbf{P}_0(X\mathbf{Q}^*)D^*$. Since this is 0 for all D in \mathcal{D}_2 , it follows that $\mathbf{P}_0(X\mathbf{Q}^*) = 0$.

Lemma 3. *Let \mathcal{H} be a left D -invariant subspace in \mathcal{X}_2 , generated by an orthonormal basis representation \mathbf{Q} . The projection of any $X \in \mathcal{X}_2$ onto \mathcal{H} is given by*

$$\mathbf{P}_{\mathcal{H}}(X) = D\mathbf{Q},$$

with $D = \mathbf{P}_0(X\mathbf{Q}^*)$.

PROOF An operator \mathbf{P} is a projector onto a subspace \mathcal{H} if it is idempotent: $\mathbf{P}\mathbf{P} = \mathbf{P}$, and if its range is \mathcal{H} . This last requirement is true because $\mathbf{P}_{\mathcal{H}}(X) = D\mathbf{Q}$, with $D = \mathbf{P}_0(X\mathbf{Q}^*) \in \mathcal{D}_2$, and all elements in \mathcal{D}_2 can be reached this way.

$\mathbf{P}_{\mathcal{H}}$ is idempotent: if $X \in \mathcal{H}$, then $\mathbf{P}_{\mathcal{H}}(X) = X$. Indeed, by proposition 1, $X = D_1\mathbf{Q}$ for some $D_1 \in \mathcal{X}_2$, since \mathbf{Q} is a basis. In fact D_1 is equal to $\mathbf{P}_0(X\mathbf{Q}^*)$: $\mathbf{P}_0(X\mathbf{Q}^*) = \mathbf{P}_0(D_1\mathbf{Q}\mathbf{Q}^*) = D_1\mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*) = D_1$, hence $\mathbf{P}_{\mathcal{H}}(X) = D\mathbf{Q} = D_1\mathbf{Q} = X$.

Finally, the projector is orthogonal: if $X \in \mathcal{X}_2$, then $X - \mathbf{P}_{\mathcal{H}}(X) \in \mathcal{H}^\perp$ because

$$\begin{aligned} \mathbf{P}_0((X - \mathbf{P}_0(X\mathbf{Q}^*)\mathbf{Q})\mathbf{Q}^*) &= \mathbf{P}_0(X\mathbf{Q}^*) - \mathbf{P}_0(\mathbf{P}_0(X\mathbf{Q}^*) \cdot \mathbf{Q}\mathbf{Q}^*) \\ &= \mathbf{P}_0(X\mathbf{Q}^*) - \mathbf{P}_0(X\mathbf{Q}^*)\mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*) \\ &= \mathbf{P}_0(X\mathbf{Q}^*) - \mathbf{P}_0(X\mathbf{Q}^*) = 0. \end{aligned}$$

□

If \mathbf{F} is a strong basis representation generating \mathcal{H} , then

$$\mathbf{P}_{\mathcal{H}}(X) = \mathbf{P}_0(X\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1}\mathbf{F}$$

is also a projection onto \mathcal{H} . This can be derived from the orthogonal projection by putting $\mathbf{F} = R\mathbf{Q}$, where $\Lambda_{\mathbf{F}} = RR^*$ must be boundedly invertible.

3. NERODE STATE SPACE DEFINITIONS

Let be given a bounded linear causal time-varying system with n_1 input ports and n_0 output ports, and with transfer operator T in $\mathcal{U}(\mathcal{N}^{\bullet 1}, \mathcal{P}^{\bullet 1})$, where $\mathcal{N} = \mathbf{C}^{n_1}$ and $\mathcal{P} = \mathbf{C}^{n_0}$. We will derive a state space description for T , *i.e.*, some representation of T such that when $u \in \ell_2(\mathbf{C}^1, \mathcal{N}^{\bullet 1})$ is an input sequence and $y = uT$ is its corresponding output, we can recover any entry y_k of y from knowledge of u_k and a compact (state) representation of $\{u_i : i \leq k-1\}$, the “past” of u with respect to instant k . It is of course not enough to consider only one pair u, y and hope to recover a state space description from it, or to consider only one time instant k . One approach is to let u range over all ℓ_2 , and to consider, for each time instant, the relation between inputs applied until instant $k-1$ (*i.e.*, the projection of ℓ_2 onto this subspace) and corresponding outputs from instant k on (the projection of y onto “the future”). This is akin to a Hilbert resolution space approach and is described in detail in a monograph on time-varying system theory by Feintuch and Saeks [6]. The approach we take here is (necessarily) strongly related to this resolution method, yet has a few additional merits.

We consider inputs and corresponding outputs as elements of \mathcal{X}_2 , *i.e.*, an infinite collection of ℓ_2 input sequences such that the energy (Hilbert-Schmidt norm) of the total collection is bounded. Since the operators in \mathcal{X}_2 admit a decomposition into diagonals, and projections onto \mathcal{L}_2Z^{-1} and \mathcal{U}_2 or even \mathcal{U}_2Z are well defined, we avoid much of the problems of causality and strict causality to which a major part of [6] is devoted. A second advantage is that, in order to arrive at a state space description, it is enough to consider the effect of inputs in \mathcal{L}_2Z^{-1} (the “past”) onto the projection onto \mathcal{U}_2 (the future part) of their corresponding outputs, *i.e.*, to study operators $\mathbf{P}(UT) : U \in \mathcal{L}_2Z^{-1}$. In this way, the notion of time is avoided almost completely, and as a consequence the use of indices representing time is often not needed. The resulting theory is elegant and in a natural way almost looks like a time invariant theory with non-commutative operators.

Let be given a bounded linear causal time-varying system with transfer operator T in $\mathcal{U}(\mathcal{N}^{\bullet 1}, \mathcal{P}^{\bullet 1})$, for \mathcal{N}, \mathcal{P} some finite-dimensional Hilbert spaces. Define the Hankel operator H_T associated to T to be

$$H_T : \mathcal{L}_2Z^{-1} \rightarrow \mathcal{U}_2 : \quad UH_T = \mathbf{P}(UT)$$

We consider the effect of inputs in \mathcal{L}_2Z^{-1} onto outputs in \mathcal{U}_2 , *i.e.*, we study the range and kernel of the operators H_T and H_T^* .

We say that an input U_1 is Nerode equivalent to U_2 , $U_1 \stackrel{N}{\simeq} U_2$, for $U_{1,2} \in \mathcal{L}_2Z^{-1}$, if $\mathbf{P}((U_1 - U_2)T) = 0$. Accordingly, $U \stackrel{N}{\simeq} 0$ if $U \in \mathcal{L}_2Z^{-1}$ and $\mathbf{P}(UT) = 0$, *i.e.*, if U is in the kernel of H_T . Define $\mathcal{M} = \{U : U \stackrel{N}{\simeq} 0\} = \{U \in \mathcal{L}_2Z^{-1} : \mathbf{P}(UT) = 0\}$. \mathcal{M} is called the input null space. It is a left D-invariant subspace in \mathcal{L}_2Z^{-1} . Denote the complement of \mathcal{M} in \mathcal{L}_2Z^{-1} by \mathcal{H} (called the input state space):

$$\mathcal{L}_2Z^{-1} = \mathcal{H} \oplus \mathcal{M}.$$

Define the natural output space \mathcal{H}_0 in \mathcal{U}_2 to be the range of the operator H_T :

$$\mathcal{H}_0 = \{\mathbf{P}(UT) : U \in \mathcal{L}_2Z^{-1}\}.$$

\mathcal{H}_0 is the left D-invariant subspace containing the projection in \mathcal{U}_2 of all outputs of the system that can be generated from inputs in \mathcal{L}_2Z^{-1} . Denote the complement of \mathcal{H}_0 in \mathcal{U}_2 by \mathcal{M}_0 :

$$\mathcal{U}_2 = \mathcal{H}_0 \oplus \mathcal{M}_0$$

From these definitions, the relations

$$\begin{aligned} \mathbf{P}(\mathcal{L}_2Z^{-1}T) &= \mathbf{P}(\mathcal{H}T) + \mathbf{P}(\mathcal{M}T) \\ &= \mathbf{P}(\mathcal{H}T) \\ &= \mathcal{H}_0, \end{aligned}$$

follow immediately, and with slightly more work,

$$\begin{aligned} \mathbf{P}_{\mathcal{L}_2Z^{-1}}(\mathcal{U}_2T^*) &= \mathbf{P}_{\mathcal{L}_2Z^{-1}}(\mathcal{H}_0T^*) + \mathbf{P}_{\mathcal{L}_2Z^{-1}}(\mathcal{M}_0T^*) \\ &= \mathbf{P}_{\mathcal{L}_2Z^{-1}}(\mathcal{H}_0T^*) \\ &= \mathcal{H}. \end{aligned}$$

The following shift invariance properties are fundamental from the definitions of the above

subspaces.

Theorem 4. *The spaces $\mathcal{H}, \mathcal{M}, \mathcal{H}_0, \mathcal{M}_0$ are left D -invariant subspaces satisfying the shift-invariance properties:*

$$\begin{cases} Z^{-1}\mathcal{M} \subset \mathcal{M} \\ \mathbf{R}'\mathcal{H} \subset \mathcal{H} \end{cases} \quad \begin{cases} Z\mathcal{M}_0 \subset \mathcal{M}_0 \\ \mathbf{R}\mathcal{H}_0 \subset \mathcal{H}_0 \end{cases}$$

in which the restricted shift operator \mathbf{R}' is defined by $\mathbf{R}'U = \mathbf{P}_{\mathcal{L}_2Z^{-1}}(ZU) = ZU - \mathbf{P}_0(ZU)$ (for $U \in \mathcal{L}_2Z^{-1}$), and \mathbf{R} on \mathcal{U}_2 defined by $\mathbf{R}Y = \mathbf{P}(Z^{-1}Y)$,

PROOF

1. $Z^{-1}\mathcal{M} \subset \mathcal{M}$. If $U \in \mathcal{M}$, so that $\mathbf{P}(UT) = 0$, then $UT \in \mathcal{L}_2Z^{-1}$, from which it follows that $Z^{-1}UT \in \mathcal{L}_2Z^{-1}$ also, and $\mathbf{P}(Z^{-1}UT) = 0$.
2. $\mathbf{R}'\mathcal{H} \subset \mathcal{H}$. This is a consequence of the shift invariance of \mathcal{M} in the following way. If $U \in \mathcal{H}$, then $\mathbf{P}_{\mathcal{L}_2Z^{-1}}(ZU) \in \mathcal{L}_2Z^{-1}$ by definition, and $\mathbf{P}_{\mathcal{L}_2Z^{-1}}(ZU) \perp \mathcal{M}$ because for all $X \in \mathcal{M}$, $\{\mathbf{P}_{\mathcal{L}_2Z^{-1}}(ZU), X\} = \{ZU, X\} - \{\mathbf{P}_0(ZU), X\} = \{U, Z^{-1}X\}^{(-1)}$. Since \mathcal{M} is shift-invariant, $Z^{-1}X \in \mathcal{M}$, and $\{\mathbf{R}'U, X\} = 0$.
3. $\mathbf{R}\mathcal{H}_0 \subset \mathcal{H}_0$:

$$\begin{aligned} \mathbf{P}(Z^{-1}\mathcal{H}_0) &= \{\mathbf{P}(Z^{-1}\mathbf{P}(UT)) : U \in \mathcal{L}_2Z^{-1}\} \\ &= \{\mathbf{P}(Z^{-1}UT) : U \in \mathcal{L}_2Z^{-1}\} \\ &\subset \mathcal{H}_0, \end{aligned}$$

because $Z^{-1}U \in \mathcal{L}_2Z^{-1}$.

4. $Z\mathcal{M}_0 \subset \mathcal{M}_0$. Since $Z\mathcal{M}_0 \subset Z\mathcal{U}_2 \subset \mathcal{U}_2$, we only have to prove that $Z\mathcal{M}_0 \perp \mathcal{H}_0$. For any $Y \in \mathcal{M}_0, X \in \mathcal{H}_0$,

$$\begin{aligned} \{ZY, X\} &= \mathbf{P}_0(ZY \cdot X^*) \\ &= \mathbf{P}_0\left((Y \cdot X^*)^{(-1)}Z\right) = \mathbf{P}_0(Y \cdot X^*Z)^{(-1)} \\ &= \{Y, Z^{-1}X\}^{(-1)} = \{Y, Z^{-1}(X - X_{[0]})\}^{(-1)}. \end{aligned}$$

Use is made of the fact that $ZDZ^{-1} = D^{(-1)}$. \mathcal{H}_0 is left Z^{-1} -invariant: $Z^{-1}(X - X_{[0]}) \in \mathcal{H}_0$. Hence, since $Y \perp \mathcal{H}_0$, $\{ZY, X\} = 0$. \square

4. CANONICAL STATE SPACE REALIZATIONS

Let T be a given bounded linear causal time-varying system transfer operator in $\mathcal{U}(\mathcal{N}^{\bullet 1}, \mathcal{P}^{\bullet 1})$, and assume that its shift-invariant input/output state and null spaces, $\mathcal{H}, \mathcal{H}_0, \mathcal{M}$ and \mathcal{M}_0 , are known. \mathcal{H} is such that $\mathbf{P}(\mathcal{L}_2Z^{-1}T) = \mathbf{P}(\mathcal{H}T)$, hence the effect of any input in the past (\mathcal{L}_2Z^{-1}) onto the future output in \mathcal{U}_2 is equivalently described by a (unique) representative element \mathbf{X} of \mathcal{H} , called the state. The point is that \mathcal{H} is assumed to be a much smaller dimensional space than \mathcal{L}_2Z^{-1} , so that the state indeed ‘‘summarizes’’ the past input. A refinement of these

observations leads to the construction of a operator state space model, in a way that is already familiar from a number of other contexts as well. By choosing a basis in either the input state space or the output state space, the desired result, a minimal state space realization involving only diagonal operators, is obtained.

4.1. “Canonical Controller” State Space Realization

For a given input U in \mathcal{X}_2 and instant k , define the past input $U_{-(k)}$ (with respect to instant k) to be $U_{-(k)} = \mathbf{P}_{\mathcal{L}_2 Z^{-1}}(Z^{-k}U)$. Define the state $\mathbf{X}_k \in \mathcal{H}$ at instant k to be the projection of the past input onto \mathcal{H} : $\mathbf{X}_k = \mathbf{P}_{\mathcal{H}}(U_{-(k)}) = \mathbf{P}_{\mathcal{H}}(Z^{-k}U) \in \mathcal{L}_2 Z^{-1}$.

Theorem 5. *Given a bounded system transfer operator $T \in \mathcal{U}(\mathcal{N}^{\bullet 1}, \mathcal{P}^{\bullet 1})$ with input state space \mathcal{H} , then with the above definition of $\mathbf{X}_k \in \mathcal{H}$, we have the “operator state space” realization*

$$Y = UT \quad \iff \quad \begin{cases} \mathbf{X}_{k+1} = \mathbf{X}_k \mathbf{A} + U_{[k]} \mathbf{B} \\ Y_{[k]} = \mathbf{X}_k \mathbf{C} + U_{[k]} \mathbf{D} \end{cases}$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are bounded operators satisfying

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{\mathcal{H}}(Z^{-1}\cdot) & \mathbf{P}_0(\cdot T) \\ \mathbf{P}_{\mathcal{H}}(Z^{-1}\cdot) & \mathbf{P}_0(\cdot T) \end{bmatrix}.$$

PROOF Recall that since $U_{-(k)} \in \mathcal{L}_2 Z^{-1} = \mathcal{H} \oplus \mathcal{M}$, and $\mathbf{P}_0(\mathcal{M}T) = 0$ by definition of \mathcal{M} , we have $\mathbf{P}_0(U_{-(k)}T) = \mathbf{P}_0(\mathbf{P}_{\mathcal{H}}(U_{-(k)})T + \mathbf{P}_{\mathcal{M}}(U_{-(k)})T) = \mathbf{P}_0(\mathbf{X}_k T)$.

1. $Y = UT \iff Y_{[k]} = \mathbf{P}_0(Z^{-k}Y)$
 $= \mathbf{P}_0(Z^{-k}UT)$
 $= \mathbf{P}_0(U_{-(k)}T) + \mathbf{P}_0(U_{[k]}T)$
 $= \mathbf{P}_0(\mathbf{X}_k T) + U_{[k]} \mathbf{P}_0 T.$
2. $\mathbf{X}_{k+1} = \mathbf{P}_{\mathcal{H}}(U_{-(k+1)})$
 $= \mathbf{P}_{\mathcal{H}}(Z^{-k-1}U)$
 $= \mathbf{P}_{\mathcal{H}}(Z^{-1}U_{-(k)} + Z^{-1}U_{[k]})$
 $= \mathbf{P}_{\mathcal{H}}\left(Z^{-1}\mathbf{P}_{\mathcal{H}}(U_{-(k)}) + Z^{-1}\mathbf{P}_{\mathcal{M}}(U_{-(k)})\right) + \mathbf{P}_{\mathcal{H}}(Z^{-1}U_{[k]})$
 $= \mathbf{P}_{\mathcal{H}}(Z^{-1}\mathbf{X}_k) + \mathbf{P}_{\mathcal{H}}(Z^{-1}U_{[k]}),$

where in making the last step the fact is used that \mathcal{M} is shift-invariant ($Z^{-1}\mathcal{M} \subset \mathcal{M}$) and that $\mathcal{H} \perp \mathcal{M}$. \square

It is clear that $\|\mathbf{A}\| \leq 1$, and that if there exists an $\hat{\mathbf{X}} \in \mathcal{H}$ such that $Z^{-1}\hat{\mathbf{X}} \in \mathcal{H}$, then $\|\mathbf{A}\| = 1$. Let $r(\mathbf{A})$ denote the spectral radius of \mathbf{A} :

$$r(\mathbf{A}) = \lim_{n \rightarrow \infty} \|\mathbf{A}^n\|^{1/n}.$$

Since $\|\mathbf{A}\| \leq 1$ we have that $r(\mathbf{A}) \leq 1$ also.

The above state space description in terms of operators is not yet very useful. By choosing an orthogonal basis \mathbf{Q} in \mathcal{H} , it is possible to “precompute” the effect of the operators \mathbf{A}, \mathbf{B} and T on \mathbf{Q} , and arrive at a state space description with diagonal operators A, B, C, D only. This

is demonstrated in the following theorem. Some care must be taken if \mathbf{Q} is an unbounded operator on \mathcal{X}_2 . It can be shown that this happens only if $r(\mathbf{A}) = 1$, and that $r(\mathbf{A}) = 1$ coincides with $\ell_A = 1$, where $\ell_A = r(ZA)$ is the spectral radius of the operator ZA . Nonetheless, \mathbf{Q} is bounded as a $(\mathcal{D}_2 \rightarrow \mathcal{X}_2)$ operator, and this property is sufficient to prove the theorem.

Theorem 6. *Given a bounded system transfer operator $T \in \mathcal{U}(\mathcal{N}^{\bullet 1}, \mathcal{P}^{\bullet 1})$, and assume that the input state space \mathcal{H} of T is locally finite dimensional. Let $N = \dim(\mathcal{H})$, and let \mathbf{Q} represent an orthonormal N -dimensional basis of \mathcal{H} , such that $\Lambda_{\mathbf{Q}} = \mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*) = I$.*

1. T admits a state space realization

$$Y = UT \iff \begin{cases} XZ^{-1} = XA + UB \\ Y = XC + UD \end{cases}, \quad (1)$$

where

$$\begin{aligned} A &= \mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*Z^{-1}) \in \mathcal{D}(\mathbf{C}^N, \mathbf{C}^{N^-}) & C &= \mathbf{P}_0(\mathbf{Q}T) \in \mathcal{D}(\mathbf{C}^N, \mathcal{P}^{\bullet 1}) \\ B &= \mathbf{P}_0(\mathbf{Q}^*Z^{-1}) \in \mathcal{D}(\mathcal{N}^{\bullet 1}, \mathbf{C}^{N^-}) & D &= \mathbf{P}_0(T) \in \mathcal{D}(\mathcal{N}^{\bullet 1}, \mathcal{P}^{\bullet 1}). \end{aligned}$$

2. The realization satisfies the following relations:

$$\|A\| \leq 1,$$

$$\begin{aligned} \mathbf{Q}^* &= \mathbf{Q}^*AZ + BZ \\ T &= \mathbf{Q}^*C + D \end{aligned} \quad (2)$$

$$A^*A + B^*B = I \quad (3)$$

3. If $\ell_A = r(ZA) < 1$, then

$$\begin{aligned} \mathbf{Q} &= [B(I - ZA)^{-1}Z]^* \\ T &= D + B(I - ZA)^{-1}ZC \end{aligned} \quad (4)$$

so that \mathbf{Q} is a bounded operator in $\mathcal{L}Z^{-1}$, and $X \in \mathcal{X}_2(\mathbf{C}^{\bullet 1}, \mathbf{C}^N)$.

PROOF

1. Expanding X into its diagonals, $X = \sum_{-\infty}^{\infty} Z^k X_{[k]}$, we will derive the equivalent relation

$$Y = UT \iff \begin{cases} X_{[k+1]}^{(-)} = X_{[k]}A + U_{[k]}B \\ Y_{[k]} = X_{[k]}C + U_{[k]}D \end{cases}.$$

For a given \mathbf{X}_k in \mathcal{H} , it is possible to write \mathbf{X}_k in terms of the basis \mathbf{Q} of \mathcal{H} : $\mathbf{X}_k = X_{[k]}\mathbf{Q}$, for some $X_{[k]} \in \mathcal{D}_2(\mathbf{C}^{\bullet 1}, \mathbf{C}^N)$. Starting, for a certain k and \mathbf{X}_k , with the realization in theorem 5, write the new state \mathbf{X}_{k+1} as $\mathbf{X}_{k+1} = X_{[k+1]}\mathbf{Q}$. Then

$$\begin{aligned} \mathbf{X}_{k+1} = X_{[k+1]}\mathbf{Q} &= \mathbf{P}_{\mathcal{H}}(Z^{-1}\mathbf{X}_k) && + \mathbf{P}_{\mathcal{H}}(Z^{-1}U_{[k]}) \\ &= \mathbf{P}_0(Z^{-1}\mathbf{X}_k\mathbf{Q}^*)\mathbf{Q} && + \mathbf{P}_0(Z^{-1}U_{[k]}\mathbf{Q}^*)\mathbf{Q} \\ &= \mathbf{P}_0(Z^{-1}X_{[k]}\mathbf{Q}\mathbf{Q}^*)\mathbf{Q} && + \mathbf{P}_0(Z^{-1}U_{[k]}\mathbf{Q}^*)\mathbf{Q} \\ &= X_{[k]}^+\mathbf{P}_0(Z^{-1}\mathbf{Q}\mathbf{Q}^*)\mathbf{Q} && + U_{[k]}^+\mathbf{P}_0(Z^{-1}\mathbf{Q}^*)\mathbf{Q} \\ X_{[k+1]} &= X_{[k]}^+\mathbf{P}_0(Z^{-1}\mathbf{Q}\mathbf{Q}^*) + U_{[k]}^+\mathbf{P}_0(Z^{-1}\mathbf{Q}^*). \end{aligned}$$

Putting $A^+ = \mathbf{P}_0(Z^{-1}\mathbf{Q}\mathbf{Q}^*)$ and $B^+ = \mathbf{P}_0(Z^{-1}\mathbf{Q}^*)$, i.e., $A = \mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*Z^{-1})$ and $B = \mathbf{P}_0(\mathbf{Q}^*Z^{-1})$, gives the first part of the result. In the same way, $C = \mathbf{P}_0(\mathbf{Q}T)$ is derived via

$$\begin{aligned}\mathbf{P}_0(\mathbf{X}_k T) &= \mathbf{P}_0(X_{[k]} \mathbf{Q}T) \\ &= X_{[k]} \mathbf{P}_0(\mathbf{Q}T).\end{aligned}$$

2. From the above formula we have that $\|A\| = \sup_i \|A_i\| = \sup_i \|\mathbf{Q}_i \mathbf{Q}_{i+1}^*\| \leq 1$ since $\mathbf{Q}_i \mathbf{Q}_i^* = I$ for all i . Continuing, since

$$\begin{aligned}X_{[k]} &= \mathbf{P}_0(\mathbf{X}_k \mathbf{Q}^*) \\ &= \mathbf{P}_0(U_{-(k)} \mathbf{Q}^*) \\ &= \mathbf{P}_0(Z^{-k} U \mathbf{Q}^*)\end{aligned}$$

and $X = \sum Z^k X_{[k]}$, it follows that $X = U \mathbf{Q}^*$. Combining this with the state equations (1) yields

$$\begin{bmatrix} U \mathbf{Q}^* Z^{-1} &= U \mathbf{Q}^* A + UB \\ UT &= U \mathbf{Q}^* C + UD \end{bmatrix},$$

(for all $U \in \mathcal{X}_2$), or

$$\begin{bmatrix} \mathbf{Q}^* Z^{-1} &= \mathbf{Q}^* A + B \\ T &= \mathbf{Q}^* C + D \end{bmatrix}.$$

This proves (2). Equation (3) follows by using the expression on \mathbf{Q}^* in the computation of $\Lambda_{\mathbf{Q}} = I$:

$$\begin{aligned}\Lambda_{\mathbf{Q}} &= \mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*) \\ &= \mathbf{P}_0([Z^* A^* \mathbf{Q} + Z^* B^*][\mathbf{Q}^* A Z + B Z]) \\ &= Z^* A^* \mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*) A Z + Z^* B^* B Z \\ &= Z^* A^* A Z + Z^* B^* B Z = I \\ \Rightarrow &A^* A + B^* B = I\end{aligned}$$

3. Assuming $\ell_A < 1$ so that $(I - ZA)^{-1}$ is bounded, equation (2) can be rewritten via $\mathbf{Q}^* = BZ(I - AZ)^{-1}$ into equation (4). This shows that \mathbf{Q} is a bounded operator, hence $X = U \mathbf{Q}^*$ is bounded in the Hilbert-Schmidt norm. \square

Definition 7. (Bounded State Equivalence) A realization $\{A_1, B_1, C_1, D_1\}$ is said to be boundedly state-equivalent to a given realization $\{A, B, C, D\}$, if there exists a boundedly invertible state transformation operator $R \in \mathcal{D}(\mathbf{C}^N, \mathbf{C}^N)$, such that

$$\begin{bmatrix} A_1 & C_1 \\ B_1 & D_1 \end{bmatrix} = \begin{bmatrix} R & \\ & I \end{bmatrix} \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} R^{(-1)} & \\ & I \end{bmatrix}$$

To see the rationale behind this definition, start with the given realization

$$\begin{cases} XZ^{-1} &= XA + UB \\ Y &= XC + UD \end{cases}$$

and map X to an equivalent state vector X_1 via $X = X_1R$, with R a boundedly invertible diagonal operator. Then

$$\begin{aligned} &\begin{cases} X_1RZ^{-1} &= X_1RA + UB \\ Y &= X_1RC + UD \end{cases} \\ \Leftrightarrow &\begin{cases} X_1Z^{-1} &= X_1RAR^{(-1)} + UBR^{(-1)} \\ Y &= X_1RC + UD \end{cases} \\ \Leftrightarrow &\begin{cases} X_1Z^{-1} &= X_1A_1 + UB_1 \\ Y &= X_1C_1 + UD_1 \end{cases} \end{aligned}$$

Theorem 8. *Given a bounded system transfer operator $T \in \mathcal{U}$, and assume that the input state space \mathcal{H} of T is finite dimensional. Let $N = \dim(\mathcal{H})$, and let \mathbf{F} be the representation of a strong N -dimensional bounded basis of \mathcal{H} , such that $\Lambda_{\mathbf{F}} = \mathbf{P}_0(\mathbf{F}\mathbf{F}^*) \gg 0$ and $\Lambda_{\mathbf{F}} < \infty$. Then T admits a state space realization*

$$\begin{aligned} A_1 &= \mathbf{P}_0(\mathbf{F}\mathbf{F}^*Z^{-1}) \cdot \Lambda_{\mathbf{F}}^{(-1)} & C_1 &= \mathbf{P}_0(\mathbf{F}T) \\ B_1 &= \mathbf{P}_0(\mathbf{F}^*Z^{-1}) \cdot \Lambda_{\mathbf{F}}^{(-1)} & D_1 &= \mathbf{P}_0(T) \end{aligned}$$

and $\ell_{A_1} \leq 1$ and independent of the choice of the strong basis in \mathcal{H} . If $\ell_{A_1} < 1$, then

$$\begin{aligned} \mathbf{F}^* \Lambda_{\mathbf{F}}^{-1} &= [B_1(I - ZA_1)^{-1}Z] \\ T &= D_1 + B_1(I - ZA_1)^{-1}ZC_1, \end{aligned} \tag{5}$$

so that \mathbf{F} is a bounded operator in $\mathcal{L}Z^{-1}$, and $X \in \mathcal{X}_2(\mathbf{C}^1, \mathbf{C}^N)$.

PROOF The realization follows from theorem 5 in the same way as the realization in theorem 6 has been derived, but now with the projector onto \mathcal{H} written in terms of \mathbf{F} : $\mathbf{P}_{\mathcal{H}}(\cdot) = \mathbf{P}_0(\cdot\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1}\mathbf{F}$. (Rest of proof omitted.) When \mathbf{F} is written in terms of an orthonormal basis \mathbf{Q} of \mathcal{H} ,

$$\begin{aligned} \mathbf{F} &= R\mathbf{Q} \\ \Lambda_{\mathbf{F}} &= \mathbf{P}_0(\mathbf{F}\mathbf{F}^*) = RR^* \end{aligned}$$

(where $R \in \mathcal{D}(\mathbf{C}^N, \mathbf{C}^N)$ is a boundedly invertible positive factor of $\Lambda_{\mathbf{F}}$), then the above realization on \mathbf{F} can also be derived via a state transformation $X \rightarrow X_1R$ of the realization $\{A, B, C, D\}$ on \mathbf{Q} in theorem 6, e.g.,

$$\begin{aligned} A_1 = RAR^{(-1)} &= R\mathbf{P}_0(\mathbf{Q}\mathbf{Q}^*Z^{-1})R^{(-1)} \\ &= \mathbf{P}_0(R\mathbf{Q}\mathbf{Q}^*R^*Z^{-1})R^{*(-1)}R^{(-1)} \\ &= \mathbf{P}_0(\mathbf{F}\mathbf{F}^*Z^{-1})\Lambda_{\mathbf{F}}^{(-1)}. \end{aligned}$$

The other relations mentioned in the theorem follow from the application of this state transformation to the corresponding relations in theorem 6. Finally, the fact that ℓ_{A_1} is independent

of the choice of \mathbf{F} (or of R), as long as it is a strong basis, is derived from

$$\begin{aligned} \|[ZA_1]^n\|^{1/n} &= \|[ZR^{-*}AR^{*(-1)}]^n\|^{1/n} \\ &= \|[R^{*(-1)}(ZA)R^{*(-1)}]^n\|^{1/n} \\ &= \|R^{*(-1)}[ZA]^nR^{*(-1)}\|^{1/n}. \end{aligned}$$

For $n \rightarrow \infty$ and R, R^{-1} both uniformly bounded, it follows that $\ell_{A_1} = \ell_A$. \square

4.2. “Canonical Observer” State Space Realization

To obtain a realization in the observer form, define the state \mathbf{X}_k to be in the output state space \mathcal{H}_0 : again with $U_{-(k)} = \mathbf{P}_{\mathcal{L}_2Z^{-1}}(Z^{-k}U)$,

$$\mathbf{X}_k = \mathbf{P}(U_{-(k)}T) \in \mathcal{H}_0.$$

Theorem 9. *Given a bounded system transfer operator $T \in \mathcal{U}$ with output state space \mathcal{H}_0 , then with the above definition of \mathbf{X}_k , we have the “operator state space” realization*

$$Y = UT \iff \begin{cases} \mathbf{X}_{k+1} = \mathbf{X}_k\mathbf{A} + U_{[k]}\mathbf{B} \\ Y_{[k]} = \mathbf{X}_k\mathbf{C} + U_{[k]}\mathbf{D} \end{cases}$$

with

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{P}(Z^{-1}\cdot) & \mathbf{P}_0(\cdot) \\ \mathbf{P}(Z^{-1}\cdot T) & \mathbf{P}_0(\cdot T) \end{bmatrix}$$

PROOF

$$\begin{aligned} 1. \mathbf{X}_{k+1} &= \mathbf{P}(U_{-(k+1)} \cdot T) \\ &= \mathbf{P}(\mathbf{P}_{\mathcal{L}_2Z^{-1}}(Z^{-k-1}U) \cdot T) \\ &= \mathbf{P}([Z^{-1}\mathbf{P}_{\mathcal{L}_2Z^{-1}}(Z^{-k}U) + Z^{-1}U_{[k]}] \cdot T) \\ &= \mathbf{P}(Z^{-1}U_{-(k)}T + Z^{-1}U_{[k]}T) \\ &= \mathbf{P}(Z^{-1}U_{-(k)}T) + \mathbf{P}(Z^{-1}U_{[k]}T) \\ &= \mathbf{P}(Z^{-1}\mathbf{P}(U_{-(k)}T)) + \mathbf{P}(Z^{-1}U_{[k]}T) \\ &= \mathbf{P}(Z^{-1}\mathbf{X}_k) + \mathbf{P}(Z^{-1}U_{[k]}T). \end{aligned}$$

$$\begin{aligned} 2. Y_{[k]} &= \mathbf{P}_0(Z^{-k}UT) \\ &= \mathbf{P}_0(U_{-(k)}T) + \mathbf{P}_0(U_{[k]}T) \\ &= \mathbf{P}_0(\mathbf{X}_k) + U_{[k]}\mathbf{P}_0(T). \end{aligned}$$

Theorem 10. *Let be given a bounded system transfer operator $T \in \mathcal{U}$, and assume that the output state space \mathcal{H}_0 of T is known and of finite local dimension. Let $N = \dim(\mathcal{H}_0)$, and let \mathbf{G} represent an orthogonal N -dimensional basis of \mathcal{H}_0 , such that $\mathbf{P}_0(\mathbf{G}\mathbf{G}^*) = I$.*

1. A state space realization of T is

$$Y = UT \iff \begin{cases} XZ^{-1} = XA + UB \\ Y = XC + UD \end{cases}, \quad (6)$$

where

$$\begin{aligned} A &= \mathbf{P}_0(\mathbf{G}\mathbf{G}^*Z^{-1}) \in \mathcal{D}(\mathbf{C}^N, \mathbf{C}^{N^-}) & C &= \mathbf{P}_0(\mathbf{G}) \in \mathcal{D}(\mathbf{C}^N, \mathcal{P}^{\bullet 1}) \\ B &= \mathbf{P}_0(T\mathbf{G}^*Z^{-1}) \in \mathcal{D}(\mathcal{N}^{\bullet 1}, \mathbf{C}^{N^-}) & D &= \mathbf{P}_0(T) \in \mathcal{D}(\mathcal{N}^{\bullet 1}, \mathcal{P}^{\bullet 1}). \end{aligned}$$

2. The realization satisfies the following relations:

$$\|A\| \leq 1,$$

$$\begin{aligned} \mathbf{G} &= C + AZ\mathbf{G} \\ T &= BZ\mathbf{G} + D \end{aligned} \quad (7)$$

$$AA^* + CC^* = I. \quad (8)$$

3. If $\ell_A = r(ZA) < 1$, then

$$\begin{aligned} \mathbf{G} &= (I - AZ)^{-1}C \\ T &= D + B(I - ZA)^{-1}ZC, \end{aligned} \quad (9)$$

so that \mathbf{G} is a bounded operator in \mathcal{X} , and $X \in \mathcal{X}_2(\mathbf{C}^1, \mathbf{C}^N)$.

PROOF

1. Expanding X into its diagonals, $X = \sum_{-\infty}^{\infty} Z^k X_{[k]}$, we will derive the equivalent relation

$$Y = UT \iff \begin{cases} X_{[k+1]}^{(-1)} = X_{[k]}A + U_{[k]}B \\ Y_{[k]} = X_{[k]}C + U_{[k]}D \end{cases}. \quad (10)$$

The proof follows closely that of theorem 6. For a given \mathbf{X}_k in \mathcal{H}_0 , put $\mathbf{X}_k = X_{[k]}\mathbf{G}$, for some $X_{[k]} \in \mathcal{D}_2(\mathbf{C}^1, \mathbf{C}^N)$. Then

$$\begin{aligned} \mathbf{X}_{k+1} = X_{[k+1]}\mathbf{G} &= \mathbf{P}(Z^{-1}\mathbf{X}_k) && + \mathbf{P}(Z^{-1}U_{[k]}T) \\ &= \mathbf{P}_{\mathcal{H}_0}(Z^{-1}\mathbf{X}_k) && + \mathbf{P}_{\mathcal{H}_0}(Z^{-1}U_{[k]}T) \\ &= \mathbf{P}_{\mathcal{H}_0}(Z^{-1}X_{[k]}\mathbf{G}) && + \mathbf{P}_{\mathcal{H}_0}(Z^{-1}U_{[k]}T) \\ &= \mathbf{P}_0(Z^{-1}X_{[k]}\mathbf{G}\mathbf{G}^*)\mathbf{G} && + \mathbf{P}_0(Z^{-1}U_{[k]}T\mathbf{G}^*)\mathbf{G} \\ &= X_{[k]}^+ \mathbf{P}_0(Z^{-1}\mathbf{G}\mathbf{G}^*)\mathbf{G} && + U_{[k]}^+ \mathbf{P}_0(Z^{-1}T\mathbf{G}^*)\mathbf{G}. \end{aligned}$$

Hence $A = \mathbf{P}_0(\mathbf{G}\mathbf{G}^*Z^{-1})$ and $B = \mathbf{P}_0(T\mathbf{G}^*Z^{-1})$. In the same way,

$$\begin{aligned} \mathbf{P}_0(\mathbf{X}_k) &= \mathbf{P}_0(X_{[k]}\mathbf{G}) \\ &= X_{[k]}\mathbf{P}_0(\mathbf{G}), \end{aligned}$$

hence $C = \mathbf{P}_0(\mathbf{G})$.

2. $\|A\| \leq 1$ follows as in theorem 6. To show that $\mathbf{G} = C + AZ\mathbf{G}$, put $Y_{+(k)} = \mathbf{P}(U_{-(k)}T)$. Then on the one hand, $Y_{+(k)} = \mathbf{X}_k = X_{[k]}\mathbf{G}$, on the other hand, it can be shown (using (10)) that $Y_{+(k)} = X_{[k]}C + X_{[k]}AZ\mathbf{G}$. Hence $\mathbf{G} = C + AZ\mathbf{G}$. $T = BZ\mathbf{G} + D$ then follows from substituting this relation (in the form $C = (I - AZ)\mathbf{G}$) into equation (6):

$$\begin{cases} XZ^{-1} = XA + UB \\ Y = XC + UD \end{cases}$$

⇔

$$\begin{cases} X(I - AZ) = UBZ \\ Y = X(I - AZ)\mathbf{G} + UD \end{cases}$$

$$\Rightarrow Y = U(BZ\mathbf{G} + D).$$

Finally, $AA^* + CC^* = I$ follows by substituting the relation $\mathbf{G} = C + AZ\mathbf{G}$ in the expression $\Lambda_{\mathbf{G}}\mathbf{P}_0(\mathbf{G}\mathbf{G}^*) = I$.

3. $X \in \mathcal{X}_2$ if $\ell_A < 1$ follows directly once it has been established that $X = Y\mathbf{G}^*$. The proof of this property is dual to that in theorem 6 and is omitted here. \square

Theorem 11. *Given a bounded system transfer operator $T \in \mathcal{U}$, and assume that the output state space \mathcal{H}_0 of T is finite dimensional. Let $N = \dim(\mathcal{H})$, and let \mathbf{F}_0 represent a strong N -dimensional basis of \mathcal{H} , such that $\Lambda_{\mathbf{F}_0} = \mathbf{P}_0(\mathbf{F}_0\mathbf{F}_0^*) \gg 0$ and $\Lambda_{\mathbf{F}_0} < \infty$. Then T admits a state space realization*

$$\begin{aligned} A_1 &= \mathbf{P}_0(\mathbf{F}_0\mathbf{F}_0^*Z^{-1}) \cdot \Lambda_{\mathbf{F}_0}^{(-1)} & C_1 &= \mathbf{P}_0(\mathbf{F}_0) \\ B_1 &= \mathbf{P}_0(T\mathbf{F}_0^*Z^{-1}) \cdot \Lambda_{\mathbf{F}_0}^{(-1)} & D_1 &= \mathbf{P}_0(T) = T_{[0]} \end{aligned}$$

and $\ell_{A_1} \leq 1$ and independent of the choice of the basis, as long as $\Lambda_{\mathbf{F}_0} \gg 0$. If $\ell_{A_1} < 1$, then

$$\begin{aligned} \mathbf{F}_0 &= (I - A_1Z)^{-1}C_1 \\ T &= D_1 + B_1(I - ZA_1)^{-1}ZC_1. \end{aligned}$$

PROOF The proof follows from theorem 10 and goes along the lines of the proof of theorem 8, with state transformation $X = X_1R$, and orthogonal basis \mathbf{G} such that $\mathbf{F}_0 = R\mathbf{G}$. \square

Theorem 12. *Given a bounded system transfer operator $T \in \mathcal{U}$ with finite dimensional state spaces \mathcal{H} and \mathcal{H}_0 . Let \mathbf{F} be the representation of a strong basis in \mathcal{H} . Let*

$$\mathbf{F}_0 = \mathbf{P}(\mathbf{F}T)$$

and suppose that \mathbf{F}_0 represents a strong basis ($\Lambda_{\mathbf{F}_0} \gg 0$). Then the canonical realization based on \mathbf{F} (theorem 8) is identical to the canonical realization based on \mathbf{F}_0 (theorem 11).

The Hankel operator $H_T = \mathbf{P}(\cdot T)$ on \mathcal{L}_2Z^{-1} has a decomposition in terms of \mathbf{F} , \mathbf{F}_0 as

$$\begin{aligned} H_T : \mathcal{L}_2Z^{-1} \rightarrow \mathcal{U}_2 : Y = UH_T &= \mathbf{P}_0(U\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1} \cdot \mathbf{F}_0 \\ &\Leftrightarrow \begin{cases} X = \mathbf{P}_0(U\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1} \\ Y = X\mathbf{F}_0 \end{cases} \end{aligned}$$

PROOF Let X be the state of the realization on \mathbf{F} , and \bar{X} be that of \mathbf{F}_0 . We will prove that, when $\mathbf{F}_0 = \mathbf{P}(\mathbf{F}T)$, these states are the same. The proof hinges on the fact that $\mathbf{P}(U_{-(k)}T) = \mathbf{P}(\mathbf{P}_{\mathcal{H}}(U_{-(k)}T))$ by definition of \mathcal{H} . Let

$$\begin{bmatrix} \mathbf{X}_k \\ \bar{\mathbf{X}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{\mathcal{H}}(U_{-(k)}) \\ \mathbf{P}(U_{-(k)}T) \end{bmatrix}, \quad \begin{bmatrix} \mathbf{X}_k \\ \bar{\mathbf{X}}_k \end{bmatrix} = \begin{bmatrix} X_{[k]}\mathbf{F} \\ \bar{X}_{[k]}\mathbf{F}_0 \end{bmatrix}$$

(according to the definitions leading to theorems 8 and 11). Then

$$\begin{aligned}
\bar{\mathbf{X}}_k &= \mathbf{P}(U_{-(k)}T) \\
&= \mathbf{P}(\mathbf{P}_{\mathcal{H}}(U_{-(k)})T) \\
&= \mathbf{P}(\mathbf{X}_kT) \\
&= \mathbf{P}(X_{[k]}\mathbf{F}T) \\
&= X_{[k]}\mathbf{P}(\mathbf{F}T) \\
&= X_{[k]}\mathbf{F}_0
\end{aligned}$$

If \mathbf{F}_0 is strong, then $\bar{X}_{[k]} = X_{[k]}$.

To prove that $Y = UH_T = \mathbf{P}_0(U\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1}\mathbf{F}_0$, (where this $U \in \mathcal{L}_2Z^{-1}$ plays the role of any $U_{-(k)}$ of the expressions above), notice that we defined $\mathbf{X} = \mathbf{P}_{\mathcal{H}}(U) = \mathbf{P}_0(U\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1}\mathbf{F} = \mathbf{X}\mathbf{F}$, and hence $X = \mathbf{P}_0(U\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1}$ for the controller realizations, and $Y = \bar{X} = \bar{X}\mathbf{F}_0$ for the observer realizations. Since these states X, \bar{X} are the same when $\mathbf{F}_0 = \mathbf{P}(\mathbf{F}T)$, the result follows. \square

The above decomposition of the Hankel operator proves to be essential in the actual computation of a realization of a given transfer operator T , as is shown in the next section.

5. FROM TRANSFER OPERATOR TO REALIZATION

In this section we shall consider how a realization can actually be computed from the data in a transfer operator T . The Hankel operator H_T will play an important role in the computations, just as it did in the related generalized Wiener-Hopf theory developed in [13] and [14].

Diagonal Expansion of the Hankel Operator

If the operator $X \in \mathcal{U}_2$, then the diagonal expansion of X is \tilde{X} , defined by

$$\begin{aligned}
X &= X_{[0]} + ZX_{[1]} + Z^2X_{[2]} + \dots = X_{[0]} + X_{[1]}^{(-1)}Z + X_{[2]}^{(-2)}Z^2 + \dots \\
\tilde{X} &= \begin{bmatrix} X_{[0]} & X_{[1]}^{(-1)} & X_{[2]}^{(-2)} & \dots \end{bmatrix}.
\end{aligned}$$

\tilde{X} is an alternative representation of X which we still will denote as belonging to \mathcal{U}_2 . If the operator $X \in \mathcal{L}_2Z^{-1}$, then the diagonal expansion of X is also designated by \tilde{X} , now defined by

$$\begin{aligned}
X &= Z^{-1}X_{[-1]} + Z^{-2}X_{[-2]} + \dots = X_{[-1]}^{(+1)}Z^{-1} + X_{[-2]}^{(+2)}Z^{-2} + \dots \\
\tilde{X} &= \begin{bmatrix} X_{[-1]}^{(+1)} & X_{[-2]}^{(+2)} & \dots \end{bmatrix}
\end{aligned}$$

Here also is \tilde{X} an alternative representation of X .

These definitions keep entries of X that are on the same i -th row π_i^*X in X also on the same row $\pi_i^*\tilde{X}$ in \tilde{X} . This is seen directly from the second expansion of X in Z , since a multiplication of a diagonal on the right by Z will only shift its columns. In addition, we have that $\mathbf{P}_0(XX^*) = \tilde{X}\tilde{X}^*$.

Using diagonal expansions, we can associate an operator \tilde{H}_T to the Hankel operator H_T of a system T , in the sense that \tilde{H}_T maps the diagonal expansion of U to the diagonal expansion of Y . This diagonal expansion of the operator H_T has a matrix representation with entries in

\mathcal{D} and can be specified in terms of the entries of T :

Theorem 13. Let $T \in \mathcal{U}$, and $Y = UH_T$ with $U \in \mathcal{L}_2\mathcal{Z}^{-1}$. The matrix representation of the operator \tilde{H}_T such that $\tilde{Y} = \tilde{U}\tilde{H}_T$ is given by

$$\tilde{H}_T = \begin{bmatrix} T_{[1]} & T_{[2]}^{(-1)} & T_{[3]}^{(-2)} & \cdots \\ T_{[2]} & T_{[3]}^{(-1)} & & \\ T_{[3]} & & \ddots & \\ \vdots & & & \end{bmatrix}$$

PROOF The multiplication UT can be broken down into operations on diagonals of U : $Y = UT = \sum Z^k (U_{[k]}T)$. It follows that

$$Y_{[0]} = \begin{bmatrix} U_{[-1]}^{(+1)} & U_{[-2]}^{(+2)} & \cdots \end{bmatrix} \begin{bmatrix} T_{[1]} \\ T_{[2]} \\ \vdots \end{bmatrix}, \quad Y_{[1]} = \begin{bmatrix} U_{[-1]}^{(+2)} & U_{[-2]}^{(+3)} & \cdots \end{bmatrix} \begin{bmatrix} T_{[2]} \\ T_{[3]} \\ \vdots \end{bmatrix},$$

etc. Hence

$$\begin{bmatrix} Y_{[0]} & Y_{[1]}^{(-1)} & Y_{[2]}^{(-2)} & \cdots \end{bmatrix} = \begin{bmatrix} U_{[-1]}^{(+1)} & U_{[-2]}^{(+2)} & U_{[-3]}^{(+3)} & \cdots \end{bmatrix} \tilde{H}_T$$

with \tilde{H}_T as claimed. \square

A nice connection of T with \tilde{H}_T is obtained by construct (infinite size) submatrices H_i ($-\infty < i < \infty$) of \tilde{H}_T by selecting the i -th entry of each diagonal in \tilde{H}_T . The H_i can be viewed as *time-varying* Hankel matrices. The entries of H_i are entries of T , e.g.,

$$H_0 = \begin{bmatrix} T_{-1,0} & T_{-1,1} & T_{-1,2} & \cdots \\ T_{-2,0} & T_{-2,1} & & \\ T_{-3,0} & & \ddots & \\ \vdots & & & \end{bmatrix}$$

Hence the rows of H_i are parts of the rows of T , and in fact the H_i are mirrored submatrices of T , as seen in figure 4. The mirroring effect is introduced by definition of the diagonal expansion of operators in $\mathcal{L}_2\mathcal{Z}^{-1}$.

Hankel Matrix Decompositions

We will need the following results from the previous chapter. Given a bounded system transfer operator $T \in \mathcal{U}$, and assume that the input/output state spaces \mathcal{H} and \mathcal{H}_0 of T are finite dimensional. Let \mathbf{F} be the representation of a strong N -dimensional basis of \mathcal{H} , and \mathbf{F}_0 the representation of a strong N -dimensional basis of \mathcal{H}_0 . Then a state space realization of T based on \mathbf{F} is

$$\begin{aligned} A &= \mathbf{P}_0(\mathbf{F}\mathbf{F}^*Z^{-1}) \cdot \Lambda_{\mathbf{F}}^{(-1)} & C &= \mathbf{P}_0(\mathbf{F}T) \\ B &= \mathbf{P}_0(\mathbf{F}^*Z^{-1}) \cdot \Lambda_{\mathbf{F}}^{(-1)} & D &= \mathbf{P}_0(T) \end{aligned} \quad (11)$$

and (assuming $\ell_A < 1$) satisfies $\Lambda_{\mathbf{F}}^{-1}\mathbf{F}^* = B(I - ZA)^{-1}Z$.

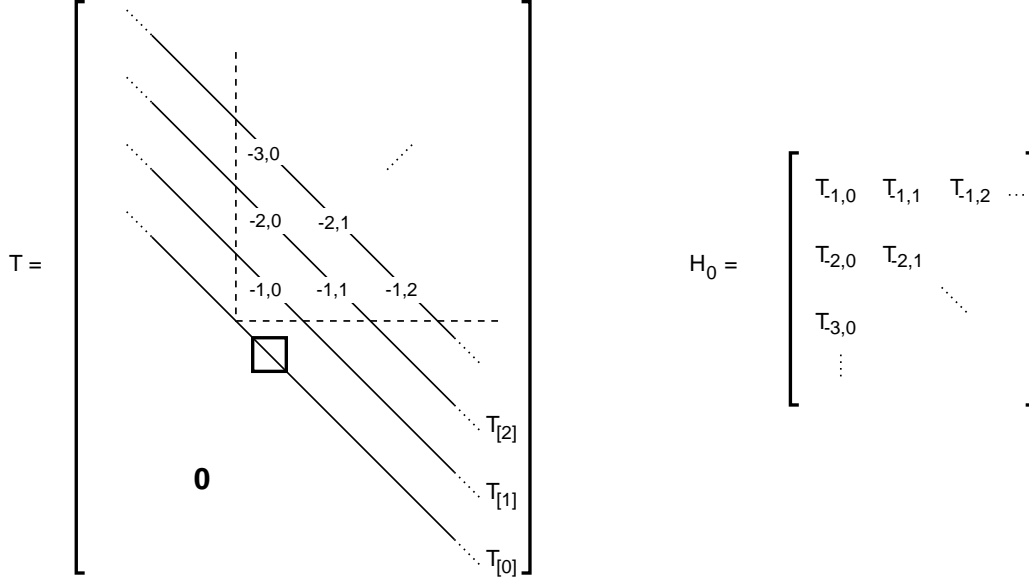


Figure 4. Hankel matrices are submatrices of T .

A second realization that is based on \mathbf{F} follows from the above realization after applying a state transformation by $\Lambda_{\mathbf{F}}^{-1}$:

$$\begin{aligned} A &= \Lambda_{\mathbf{F}}^{-1} \mathbf{P}_0(\mathbf{F}\mathbf{F}^*Z^{-1}) & C &= \Lambda_{\mathbf{F}}^{-1} \mathbf{P}_0(\mathbf{F}T) \\ B &= \mathbf{P}_0(\mathbf{F}^*Z^{-1}) & D &= \mathbf{P}_0(T) \end{aligned} \quad (12)$$

and (assuming $\ell_A < 1$) satisfies $\mathbf{F}^* = B(I - ZA)^{-1}Z$.

A third state space realization of T is based on \mathbf{F}_0 :

$$\begin{aligned} A &= \mathbf{P}_0(\mathbf{F}_0\mathbf{F}_0^*Z^{-1}) \cdot \Lambda_{\mathbf{F}_0}^{(-1)} & C &= \mathbf{P}_0(\mathbf{F}_0) \\ B &= \mathbf{P}_0(T\mathbf{F}_0^*Z^{-1}) \cdot \Lambda_{\mathbf{F}_0}^{(-1)} & D &= \mathbf{P}_0(T) = T_{[0]} \end{aligned} \quad (13)$$

and (assuming $\ell_A < 1$) satisfies $\mathbf{F}_0 = (I - AZ)^{-1}C$.

Realization (11) is equal to realization (13) if $\mathbf{F}_0 = \mathbf{P}(\mathbf{F}T)$ is taken, and if this \mathbf{F}_0 is a strong basis representation. With $\mathbf{F}_0 = \mathbf{P}(\mathbf{F}T)$ we have a decomposition of H_T as (theorem 12)

$$H_T = \mathbf{P}_0(\cdot \mathbf{F}^*) \Lambda_{\mathbf{F}}^{-1} \cdot \mathbf{F}_0$$

Switching to diagonal expansions, this decomposition turns into a decomposition of the diagonal expansion of H_T and leads to an expression that is familiar in the time-invariant case:

Theorem 14. *Let $T \in \mathcal{U}$ be the transfer operator of a bounded system. If $\{A, B, C, D\}$ is a state space realization of T , then \tilde{H}_T has a decomposition*

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

where $\mathcal{C} : \mathcal{L}_2 Z^{-1} \rightarrow \mathcal{D}_2$, $\mathcal{O} : \mathcal{D}_2 \rightarrow \mathcal{U}_2$ are defined as

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

If the realization is given by equation (11), based on a strong basis representation \mathbf{F} generating the input state space, then \mathcal{C}^* is equal to the diagonal expansion of $\Lambda_{\mathbf{F}}^{-1}\mathbf{F}$.

If the realization is given by equation (12), again based on a strong basis representation \mathbf{F} generating the input state space, then $\tilde{\mathbf{F}} = \mathcal{C}^*$.

If the realization is given by (13), based on a strong basis representation \mathbf{F}_0 generating the output state space, then $\tilde{\mathbf{F}}_0 = \mathcal{O}$.

PROOF From $T = D + B(I - ZA)^{-1}ZC$ follows

$$\begin{aligned} T_{[1]} &= B^{(+1)}C & T_{[2]}^{(-1)} &= B^{(+1)}AC^{(-1)} & \dots \\ T_{[2]} &= B^{(+2)}A^{(+1)}C & T_{[3]}^{(-1)} &= B^{(+2)}A^{(+1)}AC^{(-1)} \\ T_{[3]} &= B^{(+3)}A^{(+2)}A^{(+1)}C \\ &\dots \end{aligned}$$

Application of theorem 13 shows that \tilde{H}_T has the claimed decomposition. (With slightly more effort, the same can be shown in case $\ell_A = 1$.)

For $\ell_A < 1$, the second part of the theorem can be inferred from the relations $\Lambda_{\mathbf{F}}^{-1}\mathbf{F} = [B(I - ZA)^{-1}Z]^*$, $\mathbf{F} = [B(I - ZA)^{-1}Z]^*$, and $\mathbf{F}_0 = (I - AZ)^{-1}C$ respectively. The theorem is formally verified by using the decomposition of the Hankel operator (theorem 12) and looking at the relation between the ordinary and the diagonally expanded Hankel operator. For $U \in \mathcal{L}_2 Z^{-1}$, realization (11) and (13) follow from

$$\begin{bmatrix} Y = UH_T \\ \tilde{Y} = \tilde{U}\tilde{H}_T \end{bmatrix} \Leftrightarrow \begin{bmatrix} X = \mathbf{P}_0(U\mathbf{F}^*)\Lambda_{\mathbf{F}}^{-1} = \tilde{U}(\tilde{\mathbf{F}})^*\Lambda_{\mathbf{F}}^{-1} = \tilde{U}\mathcal{C} \\ Y = X\mathbf{F}_0 \\ \tilde{Y} = X\tilde{\mathbf{F}}_0 = X\mathcal{O} \end{bmatrix}$$

showing that $\Lambda_{\mathbf{F}}^{-1}\tilde{\mathbf{F}} = \mathcal{C}^*$, $\tilde{\mathbf{F}}_0 = \mathcal{O}$. Realization (12) is slightly different due to a state transformation by $\Lambda_{\mathbf{F}}^{-1}$:

$$\begin{bmatrix} Y = UH_T \\ \tilde{Y} = \tilde{U}\tilde{H}_T \end{bmatrix} \Leftrightarrow \begin{bmatrix} X = \mathbf{P}_0(U\mathbf{F}^*) = \tilde{U}\tilde{\mathbf{F}}^* = \tilde{U}\mathcal{C} \\ Y = X\Lambda_{\mathbf{F}}^{-1}\mathbf{F}_0 \\ \tilde{Y} = X\Lambda_{\mathbf{F}}^{-1}\tilde{\mathbf{F}}_0 = X\mathcal{O} \end{bmatrix}$$

showing that, for this realization, $\tilde{\mathbf{F}} = \mathcal{C}^*$. □

\mathcal{C} is the controllability matrix, \mathcal{O} is the observability matrix in the present context. A realization $\{A, B, C, D\}$ is called a controllable realization if $\mathcal{C}^*\mathcal{C} > 0$, and uniformly controllable if $\mathcal{C}^*\mathcal{C} \gg 0$. In view of theorem 14, it follows straightforwardly that the second realization on a strong basis \mathbf{F} in \mathcal{H} as given by equation (12) is uniformly controllable by construction: $\Lambda_{\mathbf{F}} = \mathbf{P}_0(\mathbf{F}\mathbf{F}^*) = \mathcal{C}^*\mathcal{C}$ and $\Lambda_{\mathbf{F}} \gg 0$ yields $\mathcal{C}^*\mathcal{C} \gg 0$. If $\Lambda_{\mathbf{F}}^{-1} \gg 0$ (i.e., the state transformation connecting realization (12) to realization (11) is boundedly invertible), then the realization

(11) is also uniformly controllable, since for this realization it holds that $\Lambda_{\mathbf{F}}^{-1} = \mathcal{C}^* \mathcal{C}$.

Along the same lines, a realization $\{A, B, C, D\}$ is called observable if $\mathcal{O} \mathcal{O}^* > 0$, and uniformly observable if $\mathcal{O} \mathcal{O}^* \gg 0$. A realization based on a strong basis \mathbf{F}_0 generating \mathcal{H}_0 via equation (13) is uniformly observable by construction, since $\Lambda_{\mathbf{F}_0} = \mathbf{P}_0(\mathbf{F}_0 \mathbf{F}_0^*) = \mathcal{O} \mathcal{O}^*$ and $\Lambda_{\mathbf{F}_0} \gg 0$. A realization is called minimal if it is both controllable and observable.

The rank of the Hankel operator \tilde{H}_T is defined to be $N \in \mathcal{I}$ such that H_i has rank N_i , for $i = \dots, -1, 0, 1, \dots$. Since $\mathcal{H}_0 = (\mathcal{L}_2 Z^{-1}) H_T$ it follows immediately from the relations between an operator and its diagonal expansion that $\text{rank}(\tilde{H}_T) = \dim(\mathcal{H}_0)$. Since $\text{rank}(\tilde{H}_T) = \text{rank}(\tilde{H}_T^*)$, it follows that $\text{rank}(\tilde{H}_T) = \dim(\mathcal{H})$, and hence $\dim(\mathcal{H}_0) = \dim(\mathcal{H})$.

Theorem 15. *Let T be a bounded linear causal time-varying system transfer operator in \mathcal{U} .*

1. *If \tilde{H}_T has local finite rank $N \in \mathcal{I}$, then there exist minimal state space realizations of order N . This is a Kronecker-type result.*
2. *These realizations can be obtained from any decomposition of \tilde{H}_T into $\tilde{H}_T = \mathcal{C} \cdot \mathcal{O}$, where \mathcal{C} has column rank N , \mathcal{O} has row rank N (i.e., with $0 < \mathcal{C}^* \mathcal{C} < \infty$, $0 < \mathcal{O} \mathcal{O}^* < \infty$), whenever at least one of these products is taken uniformly positive, as follows*
 - *If $\mathcal{C}^* \mathcal{C} \gg 0$, then take $\mathbf{F} \in \mathcal{L}_2 Z^{-1}$ such that its diagonal expansion $\tilde{\mathbf{F}} = \mathcal{C}^*$. This \mathbf{F} is a strong basis representation generating the input state space \mathcal{H} of T . A realization of T is given by equation (12) and is uniformly controllable by construction.*
 - *If $\mathcal{O} \mathcal{O}^* \gg 0$, then take $\mathbf{F}_0 \in \mathcal{U}_2$ such that its diagonal expansion $\tilde{\mathbf{F}}_0 = \mathcal{O}$. This \mathbf{F}_0 is a strong basis representation generating the output state space \mathcal{H}_0 of T . A realization of T is given by (13) and is uniformly observable by construction.*
3. *Existence of a realization that is uniformly controllable and uniformly observable is a system property: it depends only on T . If it exists then a realization based on \mathbf{F} is also uniformly observable, and a realization based on \mathbf{F}_0 is also uniformly controllable.*

PROOF

1. This will follow from the construction in step 2.
2. The decomposition can be constructed via decompositions of the H_i , which is a standard linear algebra problem (typically using SVDs). The choice for \mathbf{F} and \mathbf{F}_0 is motivated by theorem 14 and the discussion following it.
3. The condition for existence of a realization that is both uniformly controllable and uniformly observable is that, given a strong basis \mathbf{F} , then $\mathbf{F}_0 = \mathbf{P}(\mathbf{F}T)$ should be a strong basis in the output state space: $\Lambda_{\mathbf{F}_0} \gg 0$. Because of the definition of input and output state space, we have at least that $\Lambda_{\mathbf{F}_0} > 0$, but it need not necessarily be uniformly positive. If it isn't, then no boundedly invertible state transformation R applied to the realization on \mathbf{F} (making it a realization based on $R\mathbf{F}$) will make it uniformly positive: $\Lambda_{\mathbf{F}'_0} := R\Lambda_{\mathbf{F}_0}R^*$. Since all realizations based on strong bases \mathbf{F} are

connected via boundedly invertible state transformations, and these realizations are the only ones that are uniformly controllable, the conclusion is that there either exists a realization that is uniformly observable in addition, or it does not exist, depending on T . \square

ACKNOWLEDGEMENT

This research was supported by the commission of the EC under the ESPRIT BRA program 3280 (NANA). It was performed partly at the Department of Theoretical Mathematics, Weizmann Institute of Science, Rehovot, Israel, whence the first author would like to thank support by The Karyn Kupcinec Summer School program, and the second author gratefully acknowledges a grant from the Meyerhoff foundation.

REFERENCES

- [1] D. Alpay, P. Dewilde, and H. Dym, "Lossless Inverse Scattering and Reproducing Kernels for Upper Triangular Operators," *Operator Theory Advances and Applications*, **47**:61–135, 1990.
- [2] D. Alpay and P. Dewilde, "Time-varying Signal Approximation and Estimation," In *Proc. MTNS-89: Signal Processing, Scattering and Operator Theory, and Numerical Methods*. Birkhäuser, 1990.
- [3] D. Alpay, P. Dewilde, and H. Dym, "Lossless Inverse Scattering, Reproducing Kernels and Nevanlinna-Pick Interpolation for Upper Triangular Operators," (*to appear*).
- [4] P. Dewilde, "A Course on the Algebraic Schur and Nevanlinna-Pick Interpolation Problems," In Ed. F. Deprettere and A.J. van der Veen, editors, *Algorithms and Parallel VLSI Architectures*. Elsevier, 1991.
- [5] L.A. Zadeh, "Time-Varying Networks, I," *Proc. IRE*, **49**:1488–1503, October 1961.
- [6] A. Feintuch and R. Saeks, "*System Theory: A Hilbert Space Approach*," Academic Press, 1982.
- [7] E.W. Kamen and K.M. Hafez, "A Transfer-Function Approach to Linear Time-Varying Discrete-Time Systems," *SIAM J. Control and Optimization*, **23**(4):550–565, 1985.
- [8] E.W. Kamen, "The Poles and Zeros of a Linear Time-Varying System," *Lin. Alg. Applications*, **98**:263–289, 1988.
- [9] B.D.O. Anderson and J.B. Moore, "Detectability and Stabilizability of Time-Varying Discrete-Time Linear Systems," *SIAM J. Control and Optimization*, **19**(1):20–32, 1981.
- [10] A.J. van der Veen and P.M. Dewilde, "Orthogonal Embedding Theory for Contractive Time-Varying Systems," In H. Kimura, editor, *Proc. Int. Symp. on MTNS*. MITA Press, Japan, 1991.
- [11] L. Kronecker, "Algebraische Reduction der Schaaren Bilinearer Formen," *S.B. Akad. Berlin*, pp. 663–776, 1890.
- [12] B.L. Ho and R.E. Kalman, "Effective Construction of Linear, State-Variable Models from Input/Output Functions," *Regelungstechnik*, **14**:545–548, 1966.
- [13] H. Nelis, "*Sparse Approximations of Inverse Matrices*," PhD thesis, Delft Univ. Techn., The Netherlands, 1989.
- [14] H. Woerdeman, "*Matrix and Operator Extensions*," PhD thesis, Dept. Math. Comp Sci., Free University, Amsterdam, The Netherlands, 1989.