ELSEVIER Linear Algebra and its Applications 332–334 (2001) 265–353

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

On the parametrization of conditioned invariant subspaces and observer theory☆

P.A. Fuhrmann [a,*], U. Helmke [b]

[a] *Department of Mathematics, Ben-Gurion University of the Negev, 84105 Beer Sheva, Israel*
[b] *Institut für Mathematik, Universität Würzburg, Würzburg, Germany*

Received 17 December 1999; accepted 17 November 2000

Submitted by F. Puerta

Abstract

The present paper is an in depth analysis of the set of conditioned invariant subspaces of a given observable pair $(\mathscr{C}, \mathscr{A})$. We do this analysis in two different ways, one based on polynomial models starting with a characterization obtained in [P.A. Fuhrmann, Linear Operators and Systems in Hilbert Space, 1981; IEEE Trans. Automat. Control AC-26 (1981) 284], the other being a state space approach. Toeplitz operators, projections in polynomial and rational models, Wiener–Hopf factorizations and factorization indices all appear and are tools in the characterizations. We single out an important subclass of conditioned invariant subspaces, namely the tight ones which already made an appearance in [P.A. Fuhrmann, U. Helmke, Systems Control Lett. 30 (1997) 217], a precursor of the present paper. Of particular importance for the study of the parametrization of the set of conditioned invariant subspaces of an observable pair (C, A) is the structural map that associates with any reachable pair, with only the input dimension constrained, a uniquely determined conditioned invariant subspace. The construction of this map uses polynomial models and the shift realization. New objects, the partial observability and reachability matrices are introduced which are needed for the state space characterizations. Kernel and image representations for conditioned invariant subspaces are derived. Uniqueness of a kernel representation of a conditioned invariant subspace is shown to be equivalent to tightness. We pass on to an analysis and derivation of the Kronecker–Hermite canonical form for full column rank, rectangular polynomial matrices. This extends the work of A.E. Eckberg (A characterization of linear systems via polynomial matrices and module theory, Ph.D. Thesis, MIT, Cambridge, MA, 1974), G.D. Forney [SIAM J. Control Optim. 13 (1973) 493] and D. Hinrichsen H.F. Münzner and D. Prätzel-Wolters [Systems

☆ Partially supported by GIF under grant no. I-526-034.

* Corresponding author. Tel.: +972-7-6461616; fax: +972-7-6469343.
E-mail address: paf@math.bgu.ac.il (P.A. Fuhrmann).

PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 2 4 8 - 8

Control Lett. 1 (1981) 192]. We proceed to give a parametrization of such matrices. Based on this and utilizing insights from D. Hinrichsen et al. [Systems Control Lett. 1 (1981) 192], we parametrize the set of conditioned invariant subspaces. We relate this to image representations, making contact with the work of J. Ferrer et al. [Linear Algebra Appl. 275/276 (1998) 161; Stratification of the set of general (A, B)-invariant subspaces (1999)]. We add a new angle by being able to parametrize the set of all reachable pairs in a kernel representation, this via an embedding result for rectangular polynomial matrices in square ones. As a by product we redo observer theory in a unified way, giving a new insight into the connection to geometric control and to the stable partial realization problem.

Keywords: Geometric control; Conditioned invariant subspaces; Observers; Polynomial models; Parametrization problems; Linear systems

1. Introduction

Geometric control theory in the sense of Basile and Marro [2] and Wonham [38], which was developed as a tool for solving basic control problems alongside frequency domain and module theoretic methods, has been rather neglected lately. This mainly as a result of the increasing popularity of the H^∞ theory. Our belief is that this area is far from exhausting its usefulness and it provides a very attractive research area with many open problems. The basic objects of study in geometric control are controlled and conditioned invariant subspaces and various variants, or subclasses, of these spaces. The structure theory of these subspaces has been developed close to two decades ago by Antoulas [1], Emre and Hautus [4], Fuhrmann and Willems [20,21], Fuhrmann [11,12], Hinrichsen et al. [28], and others. However, many questions remain still open. To some extent this paper is a continuation of Fuhrmann and Helmke [16], where a class of conditioned invariant subspaces was studied in a generic situation, but the scope of this paper is much larger as will hopefully become clear reading this introduction.

Our principal objective in this paper is a comprehensive study of the set of all conditioned invariant subspaces of a given observable pair $(\mathscr{C}, \mathscr{A})$. We shall tackle this problem by two different methods, namely state space theory on the one hand and a module theoretic, that is basically a frequency domain method on the other. The link between the two methods is the theory of polynomial and rational models as developed in [8,11]. The study of the set of all conditioned invariant subspaces of a given observable pair is not new. It has been initiated in [28], using module theory, where a complete parametrization has been obtained. This is a path breaking paper that has been overlooked for a long time. Lately the work of Ferrer et al. [5,6] is an alternative approach using state space techniques. Our point of departure is a characterization of conditioned invariant subspaces, obtained in [11,12], in the context of polynomial models and the use of the shift realization, see [8,9]. An observable pair (C, A) determines a nonsingular polynomial matrix D, unique up to a left unimodular

factor, such that the corresponding shift realization, in the state space X_D, would be isomorphic to the original pair. In this term any conditioned invariant subspace $\mathscr{V}$ has a, not necessarily unique, representation in the form

$$\mathscr{V} = X_D \cap TF^p[z]. \tag{1}$$

This turns out to be an unexpectedly rich representation and it is quite a bit of a surprise how much information is encoded in it. To a certain extent, much of this paper is an effort at uncoding this information. In the process we will encounter Toeplitz operators, Wiener–Hopf factorizations, reduced observability indices, module representations, canonical forms for observable pairs.

Let us proceed with an outline of the paper. In Section 2 we collect the needed background from polynomial model theory. We introduce polynomial and rational models, define the shift operator, as well as Toeplitz and Hankel operators. Duality theory is outlined, based on the duality in $F^p((z^{-1}))$, the $F[z]$-module of truncated vector Laurent series. The shift realization is introduced. We characterize rational models in terms of nonsingular polynomial matrices, or equivalently by similarity classes of observable pairs. Next we give the model characterizations of controlled and conditioned invariant subspaces and define the structural map that is all important in this paper. Elements of Toeplitz operator theory, in the polynomial context, are discussed via Wiener–Hopf factorization techniques. We recall the connection of factorization indices and observability indices.

Section 3 begins with the definition of the structural map. We assume that we are given an observable pair $(\mathscr{C}, \mathscr{A})$ which, without loss of generality, is given in dual Brunovsky form. The structural map associates with appropriately sized, otherwise unrestricted, pairs (A, B) a conditioned invariant subspace for the pair $(\mathscr{C}, \mathscr{A})$. This allows the study of the set of conditioned invariant subspaces via the set of reachable pairs of appropriate dimension. In this connection see Helmke [26]. We relate the set of conditioned invariant subspaces to kernels of certain Toeplitz operators as well as to those of restrictions of projection operators. Conditions for a codimension formula to hold as well as to the uniqueness of representations of the form (1) are obtained. Given a conditioned invariant subspace, we introduce the corresponding restricted observability indices and prove a majorization result in Proposition 3.5. This is dual to a result of Loiseau [30] obtained by completely different methods. The central results of this section are Theorem 3.3 characterizing the conditions under which the codimension formula $\operatorname{codim} X_D \cap TF^p[z] = \deg\det T$ holds and Theorem 3.5 characterizing the conditions under which the representation (1) is unique.

Section 4 is devoted to state space results. The central tools are new objects, the partial reachability and observability matrices. These are instrumental in the derivation of image and kernel representation of conditioned invariant subspaces. The special subclass of tight subspaces can also be characterized in these terms. One can use these objects to an ab initio analysis. However we look also at the partial observability and reachability matrices as matrix representations, with respect to canonical bases, of certain projection operators in polynomial and rational models. The central

result is Theorem 4.1. As an application we introduce the concept of a preobserver for a linear function of the state of a linear system. This is formally an observer without stability considerations. In Theorem 5.1 we characterize the existence of a preobserver in geometric terms. Next, in Theorem 5.2 existence of preobservers is related to partial realizations. This is a preview of a full study of observers taken up in Section 2.

In Section 5, we apply the previous results to the study of observers. The principle result here is Theorem 5.3, that gives the equivalence of several different characterizations for the existence of an asymptotic observer for a linear function of the state. These characterizations are scattered in the literature, not always given with correct proofs. To conclude we clarify the links between the existence of observers and the solvability of a related stable partial realization problem. The result itself is not new, however our approach seems to be new and quite clear.

We pass on, in Section 6, to the study of the Kronecker–Hermite form for full column rank polynomial matrices under the action of right multiplication by unimodular matrices. This form, given in Theorem 6.1 is an extension, to the rectangular polynomial matrix setting, of the column echelon form for constant matrices over a field. This problem has been studied before by Eckberg [3], Forney [7] and Hinrichsen and Prätzel-Wolters [27]. We hope that our presentation is relatively easy to digest. We add an enumeration formula for the number of continuous parameters in the parametrization in terms of the integral invariants.

Section 7 is devoted to the central result, namely the parametrization of the set of all conditioned invariant subspaces of a given observable pair. The set of these subspaces is given as the union of cells, determined by integral invariants. In the derivation of this parametrization, image representations are derived and the relations between them are established. Our treatment unifies the tight and nontight cases and gives also the parametrization of all minimal McMillan degree kernel representations in the nontight situation. This is done in terms of special embeddings of rectangular proper rational matrices in square ones. We show how one can avoid using the rational setting by introducing D-Kronecker–Hermite forms. Examples are provided.

2. Preliminaries

We begin by introducing polynomial and rational models. We will denote by F^m the space of all column m-vectors with coordinates in a field F. Let π_+ and π_- denote the projections of $F^m((z^{-1}))$, the space of truncated Laurent series, onto $F^m[z]$ and $z^{-1}F^m[[z^{-1}]]$, the space of polynomials and of formal power series vanishing at infinity, respectively. Since

$$F^m((z^{-1})) = F^m[z] \oplus z^{-1}F^m[[z^{-1}]], \tag{2}$$

π_+ and π_- are complementary projections. Given a nonsingular polynomial matrix D in $F^{m\times m}[z]$ we define two projections π_D in $F^m[z]$ and π^D in $z^{-1}F^m[[z^{-1}]]$ by

$$\pi_D f = D\pi_-(D^{-1}f) \quad \text{for } f \in F^m[z], \tag{3}$$

$$\pi^D h = \pi_-(D^{-1}\pi_+(Dh)) \quad \text{for } h \in z^{-1}F^m[[z^{-1}]], \tag{4}$$

and define two linear subspaces of $F^m[z]$ and $z^{-1}F^m[[z^{-1}]]$ by

$$X_D = \operatorname{Im}\pi_D \tag{5}$$

and

$$X^D = \operatorname{Im}\pi^D. \tag{6}$$

An element f of $F^m[z]$ belongs to X_D if and only if $\pi_+ D^{-1}f = 0$, i.e. if and only if $D^{-1}f$ is a strictly proper rational vector function. Thus we have also the following description of the polynomial model X_D:

$$X_D = \{f \in F^m[z] \mid f = Dh,\ h \in z^{-1}F^m[[z^{-1}]]\}. \tag{7}$$

We refer to X_D as *polynomial models* whereas to X^D as *rational models*.

We turn X_D into an $F[z]$-module by defining

$$p \cdot f = \pi_D(pf) \quad \text{for } p \in F[z], \quad f \in X_D. \tag{8}$$

Since $\operatorname{Ker}\pi_D = DF^m[z]$ it follows that X_D is isomorphic to the quotient module $F^m[z]/DF^m[z]$. Similarly, we introduce in X^D an $F[z]$-module structure by

$$p \cdot h = \pi_-(ph) \quad \text{for } p \in F[z], \quad h \in X^D. \tag{9}$$

In X_D we will focus on a special map S_D, a generalization of the classical companion matrix, which corresponds to the shift action of the polynomial z, i.e.,

$$S_D f = \pi_D z f \quad \text{for } f \in X_D.$$

Thus the module structure in X_D is identical to the module structure induced by S_D through $p \cdot f = p(S_D)f$. With this definition the study of S_D is identical to the study of the module structure of X_D. In particular the invariant subspaces of S_D are just the submodules of X_D. They are related to factorization of polynomial matrices.

Similarly, we introduce in X^D a module structure, given by

$$S^D h = \pi_- z h \quad \text{for } h \in X^D. \tag{10}$$

Polynomial and rational models are closely related. In fact, the map $\rho_D : X^D \longrightarrow X_D$ given by $h \mapsto Dh$ is an intertwining isomorphism, i.e. it satisfies $S_D\rho_D = \rho_D S^D$.

We shall need a few results about duality in the context of polynomial models. A full duality theory in this context was developed in [11,12]. We define the following pairing on $F^p((z^{-1})) \times F^p((z^{-1}))$, with $f \in F^p((z^{-1}))$, $g \in F^p((z^{-1}))$, by

$$[f, g] = \sum_{i=1}^{\infty} \tilde{g}_{-i-1} f_i. \tag{11}$$

Here $\tilde{g}$ the transpose of a vector g. We note that the defining sum has only a finite number of nonzero terms. Furthermore, the annihilator of $F^p[z]$ in $F^p((z^{-1}))$

is given by $(F^p[z])^\perp = F^p[z]$. Also, we can identify $(F^p[z])^*$, the dual space to $F^p[z]$, with $z^{-1}F^p[[z^{-1}]]$. Next we note that, given a polynomial matrix T, we have $(TF^p[z])^\perp = X^{\tilde{T}}$. This leads to the identification of the dual space to a polynomial model with a rational model, i.e.

$$(X_T)^* = X^{\tilde{T}}. \tag{12}$$

Hankel and Toeplitz operators are basic tools for the study of matrix rational functions and hence of linear systems.

For the analysis of the uniqueness of a representation the use of Wiener–Hopf factorization indices is a central tool. There exists, see [20,21], a very close connection between the analysis of state feedback, Toeplitz operators and Wiener–Hopf factorizations. This we proceed to discuss.

Let G be a $p \times m$ rational matrix, We define the *Toeplitz operator* $\mathscr{T}_G : F^m[z] \longrightarrow F^p[z]$, with *symbol* G, by

$$\mathscr{T}_G f = \pi_+ G f. \tag{13}$$

The analysis of Toeplitz operators is closely related to the study of Wiener–Hopf factorizations. These are introduced next.

Definition 2.1. Let $G \in F^{p\times m}((z^{-1}))$ be rational. A *left Wiener–Hopf factorization at infinity* is a factorization of G of the form

$$G = G_- D G_+ \tag{14}$$

with $G_+ \in F^{m\times m}[z]$ unimodular, $G_- \in F^{p\times p}[[z^{-1}]]$ biproper and

$$D(z) = \begin{pmatrix} \Delta(z) & 0 \\ 0 & 0 \end{pmatrix},$$

where $\Delta(z) = \operatorname{diag}(z^{\kappa_1}, \ldots, z^{\kappa_r})$. The integers κ_i, assumed decreasingly ordered, are called the *left factorization indices* at infinity. A *right factorization* and the *right factorization indices* are analogously defined with the plus and minus signs in (14) reversed.

We note that if $T(z)$ is a polynomial matrix, then all its right Wiener–Hopf factorization indices are nonnegative. Similarly, if $G(z)$ is a strictly proper rational function, then all its right Wiener–Hopf factorization indices are negative. The first statement follows from the fact that the factorization indices are just the row indices of a row reduced polynomial matrix obtained from T by elementary row operations. To see the other statement, we set $G(z) = K(z)/t(z)$, with K a polynomial matrix and t the l.c.m. of the denominators of all entries of G. Clearly, $t(z) = z^\tau \gamma(z)$, for some nonnegative τ and a scalar, rational biproper γ. Let $K(z) = U(z)\Delta(z)\Gamma(z)$ be a right Wiener–Hopf factorization and $\Delta(z) = \operatorname{diag}(z^{\nu_1}, \ldots, z^{\nu_p})$. So

$$G(z) = U(z)\operatorname{diag}(z^{\nu_1-\tau}, \ldots, z^{\nu_p-\tau})\frac{\Gamma(z)}{\gamma(z)}.$$

This shows that

$$U(z)\mathrm{diag}(z^{\nu_1-\tau},\dots,z^{\nu_p-\tau}) = \gamma(z)G(z)\Gamma(z)^{-1}$$

is strictly proper. Since $U(z)$ is a polynomial matrix, we must have $\nu_i - \tau < 0$, and these are the factorization indices of G.

Wiener–Hopf factorizations are a useful tool in the characterization of the elements of the polynomial model X_T.

Proposition 2.1. *Let D be a $p \times p$ nonsingular polynomial matrix. Let $D = U\Delta\Gamma$ be a right Wiener–Hopf facorization with $U(z)$ unimodular, $\Gamma(z)$ biproper and $\Delta(z) = \mathrm{diag}(z^{\nu_1},\dots,z^{\nu_p})$. Then*

1. *We have $f \in X_D$ if and only if*

$$f = Ug,\ g = \begin{pmatrix} g_1 \\ \cdot \\ \cdot \\ \cdot \\ g_p \end{pmatrix},$$

where g_i are polynomials satisfying $\deg g_i < \nu_i$. *Equivalently,*

$$X_D = UX_\Delta. \tag{15}$$

2. *If D is row proper with row indices $\nu_1,\dots,\nu_p$, then*

$$X_D = \left\{ f \,|\, f = \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_p \end{pmatrix},\ \deg f_i < \nu_i \right\} \tag{16}$$

3. *We have*

$$\dim X_D = \deg\det D = \sum_{i=1}^{p} \nu_i. \tag{17}$$

Proof.

1. We have, by the characterization (7) of the elements of X_D and the fact that $\Gamma z^{-1}F^p[[z^{-1}]] = z^{-1}F^p[[z^{-1}]]$, that

$$\begin{aligned} X_D &= \{f \mid f = U\Delta\Gamma h,\ h \in z^{-1}F^p[[z^{-1}]]\} \\ &= \{f \mid f = U\Delta h,\ h \in z^{-1}F^p[[z^{-1}]]\} \\ &= U\{g \mid g = \Delta h,\ h \in z^{-1}F^p[[z^{-1}]]\} = UX_\Delta. \end{aligned}$$

   We also used the fact that for a unimodular U we have $Ug \in F^p[z]$ if and only if $g \in F^p[z]$.
2. If D is row proper with row indices $\nu_1,\dots,\nu_p$, then we can write $D = \Delta\Gamma$, with $\Delta(z) = \mathrm{diag}(z^{\nu_1},\dots,z^{\nu_p})$ and Γ biproper. Thus $X_D = X_\Delta$. However, clearly we have $X_{z^\nu} = \{f \in F[z] \mid \deg f < \nu\}$ and the result follows.
3. Follows from the fact that $\dim X_{z^\nu} = \dim\{f \in F[z] \mid \deg f < \nu\} = \nu$. $\square$

Clearly, if G is singular, then the Toeplitz operator $\mathscr{T}_G$ has an infinite dimensional kernel or cokernel. The case of interest for us is that of a nonsingular G. In this case we have the coprime factorization $G = D^{-1}T$ with D, T square, nonsingular polynomial matrices. The following is well known.

Proposition 2.2. *Let G be a $p \times p$ nonsingular rational function and assume $G = G_- D G_+$ is a left Wiener–Hopf factorization with $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ and $\mu_1 \geqslant \cdots \geqslant \mu_p$. Then:*

1. *We have*
$$\dim \operatorname{Ker} \mathscr{T}_G = -\sum_{\mu_i < 0} \mu_i. \tag{18}$$
2. *We have*
$$\dim \operatorname{coker} \mathscr{T}_G = \sum_{\mu_i > 0} \mu_i. \tag{19}$$
3. *The Toeplitz operator $\mathscr{T}_G$ is invertible if and only if all left Wiener–Hopf factorization indices are* 0, *i.e.* $\mu_1 = \cdots = \mu_p = 0$.
In this case we have $G = G_- G_+$ and
$$\mathscr{T}_G^{-1} f = G_+^{-1} \pi_+ G_-^{-1} f, \quad f \in F^p[z]. \tag{20}$$

The next result, quoted from [20], characterizes the observability indices of an observable pair in terms of Wiener–Hopf factorization indices.

Proposition 2.3. *Let (C, A) be an observable pair, and let $T(z)^{-1} H(z) = C(zI - A)^{-1}$ be a left coprime factorization. Then the observability indices of (C, A) are equal to the right factorization indices of T.*

Given the $p \times m$ proper rational matrix function G, we define the *Hankel operator* $H_G : F^m[z] \longrightarrow z^{-1} F^p[[z^{-1}]]$ by
$$H_G f = \pi_- G f, \quad f \in F^m[z]. \tag{21}$$
It is easily seen that $\operatorname{Ker} H_G$ and $\operatorname{Im} H_G$ are $F[z]$-submodules of $F^m[z]$ and $z^{-1} F^p [[z^{-1}]]$, respectively. Rationality of G implies $\operatorname{Ker} H_G$ is a full submodule which is equivalent to the quotient module $F^m[z]/\operatorname{Ker} H_G$ being a torsion module. Similarly, rationality means that $\operatorname{Im} H_G$ is a torsion submodule of $z^{-1} F^p[[z^{-1}]]$. There exist therefore two nonsingular polynomial matrices $D_r \in F^{m \times m}[z]$ and $D_l \in F^{p \times p}[z]$ such that
$$\begin{aligned} \operatorname{Ker} H_G &= D_r F^m[z], \\ \operatorname{Im} H_G &= X^{D_l}. \end{aligned} \tag{22}$$
This leads to the coprime factorizations
$$G = N_r D_r^{-1} = D_l^{-1} N_l. \tag{23}$$

These coprime factorizations are the key, see [8], to the construction of, minimal in this case, realizations of G using shift operators. For extensions of this realization procedure, see [9].

In the next theorem we describe the shift realization associated with matrix fraction representations of rational transfer functions. We use the notation

$$G = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$

to indicate that (A, B, C, D) is a realization of a transfer function G, i.e.

$$G(z) = D + C(zI - A)^{-1}B.$$

Theorem 2.1. *Let*

$$G = D^{-1}E = \overline{ED}^{-1}$$

be, not necessarily coprime, matrix fraction representation of a proper, $p \times m$ rational function.

1. *In the state space X_D a system is defined by*

$$\begin{aligned} A &= S_D, \\ B\xi &= \pi_D \overline{E}\xi, \\ Cp &= (D^{-1}p)_{-1}, \\ D &= G(\infty). \end{aligned} \tag{24}$$

Then

$$G = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right);$$

this realization is observable and it is reachable if and only if E and D are left coprime.

2. *In the state space $X_{\overline{D}}$ a system is defined by*

$$\begin{aligned} A &= S_{\overline{D}}, \\ B\xi &= \pi_{\overline{D}}\xi, \\ Cf &= (\overline{ED}^{-1}f)_{-1}, \\ D &= G(\infty). \end{aligned} \tag{25}$$

Then

$$G = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right);$$

this realization is reachable and it is observable if and only if $\overline{E}$ and $\overline{D}$ are right coprime.

3. *In the state space $X^{\overline{D}}$ a system is defined by*

$$\begin{aligned} A &= S^{\overline{D}}, \\ B\xi &= \pi_- \overline{D}^{-1}\xi, \\ Ch &= (\overline{E}h)_{-1}, \\ D &= G(\infty). \end{aligned} \tag{26}$$

Then

$$G = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right);$$

this realization is reachable and it is observable if and only if $\overline{E}$ *and* $\overline{D}$ *are right coprime.*

Note that in the realization (24) the pair (C, A) depends only on D, and we will denote it by (C_D, A_D). Similarly, for the realizations (25) and (26), the pairs (A, B) depend only on $\overline{D}$ and we will denote it by $(A_{\overline{D}}, B_{\overline{D}})$ and $(A^{\overline{D}}, B^{\overline{D}})$, respectively.

Polynomial and rational models have convenient representations in terms of realizations. For example, if $(C, A) \in F^{p\times k} \times F^{k\times k}$ is an observable pair and $T(z)^{-1} H(z)$ is a left coprime factorization of $C(zI - A)^{-1}$, then it is well known, e.g. [23,36], that

$$X^T = \{C(zI - A)^{-1}\xi \mid \xi \in F^k\}. \tag{27}$$

Moreover, as a result of the state space isomorphism theorem, the columns of H form a basis for X_T or, equivalently, the columns of $C(zI - A)^{-1} = T(z)^{-1}H(z)$ form a basis for X^T. We note that the pair (C, A) determines the nonsingular polynomial matrix T up to a left unimodular factor. Conversely, every nonsingular polynomial matrix T determines, via the shift realization, an observable pair (C, A) which is unique up to a state space isomorphism.

3. Conditioned invariant subspaces and their representations

We quote now the characterizations of controlled and conditioned invariant subspaces, relative to the shift realizations. The first one is taken from [21], whereas the second from [11,12].

Theorem 3.1. *Let G be a $p \times m$ proper rational function having the polynomial coprime factorization $G = D^{-1}E = \overline{ED}^{-1}$. Then*:

1. *With respect to realization* (26) *in the state space $X^{\overline{D}}$, a subspace $V \subset X^{\overline{D}}$ is controlled invariant if and only if*

$$\mathscr{V} = \pi^{\overline{D}} L \tag{28}$$

for some submodule $L \subset z^{-1}F^m[[z^{-1}]]$.

2. *With respect to realization* (24) *in the state space X_D, a subspace $V \subset X_D$ is conditioned invariant if and only if*

$$\mathscr{V} = X_D \cap M \tag{29}$$

for some submodule $M \subset F^p[z]$.

These two characterizations of controlled and conditioned invariant subspaces given by (28) and (29), respectively, are as basis free as one can get, with all the corresponding advantages and disadvantages. The biggest disadvantage of these representations is that they are not specific enough. Many different submodules can correspond to the same controlled or conditioned invariant subspace, as the case may be. We note however that the submodule $L \subset z^{-1}F^m[[z^{-1}]]$ appearing in (28) can, without loss of generality, be taken as a torsion submodule. The submodule $M \subset F^p[z]$ appearing in (29) can be uniquely defined by taking $M = M_{\mathscr{V}}$, where $M_{\mathscr{V}}$ is the submodule of $F^m[z]$ generated by $\mathscr{V}$. We say a subspace $\mathscr{V} \subset F^p[z]$ has multiplicity l if the submodule $M_{\mathscr{V}}$ has rank l.

Proposition 3.1. *Given a nonsingular $p \times p$ polynomial matrix D, let (C, A) be the unique pair, defined by* (24), *that minimally realizes D in the state space X_D. Then*:

1. *A subspace $\mathscr{V} \subset X_D$ is conditioned invariant if and only if it has a, not necessarily unique, representation of the form*

$$\mathscr{V} = X_D \cap T(z)F^p[z], \tag{30}$$

where T is a nonsingular polynomial matrix. In this case we have

$$\mathscr{V} = \operatorname{Ker} \pi_T \mid X_D = X_D \cap T(z)F^p[z]. \tag{31}$$

2. *A subspace $\mathscr{V} \subset X^{\overline{D}}$ is controlled invariant if and only if it has a, not necessarily unique, representation of the form*

$$\mathscr{V} = \pi^{\overline{D}} X^{\overline{T}} \tag{32}$$

for some nonsingular polynomial matrix $\overline{T} \in F^{p\times p}[z]$.

Proof.
1. For a proof see [20].
2. For a proof see [11,12]. □

The representation formula (30) is at the heart of the paper and opens up many interesting questions, some of which we describe later. We note first that given a unimodular polynomial matrix U, we have $TF^p[z] = TUF^p[z]$ and $X^{U\overline{T}} = X^{\overline{T}}$. Thus it is not important to distinguish between representing polynomial matrices up to a one sided unimodular factor. The representations of controlled and conditioned invariant subspaces given in Proposition 3.1 have the advantage of using nonsingular polynomial matrices in the representations. The disadvantage is the nonuniqueness of the representing polynomial matrices. We can however recover uniqueness, modulo unimodular factors, by going over to the use of rectangular polynomial matrices. This we do next.

As we noted already, the submodule M in the representation of the form $\mathscr{V} = X_D \cap M$ is in general not unique. To get a unique representation we need to associate

with a conditioned invariant subspace of X_D a unique submodule, and what is more natural than the submodule of $F^p[z]$ generated by $\mathscr{V}$. The following result is adapted from [28].

Proposition 3.2. *Let $\mathscr{V} \subset X_D$ be a conditioned invariant subspaces. Let $\langle \mathscr{V} \rangle$ be the submodule of $F^p[z]$ generated by $\mathscr{V}$, that is the smallest submodule of $F^p[z]$ that contains $\mathscr{V}$. Then*

$$\mathscr{V} = X_D \cap \langle \mathscr{V} \rangle. \tag{33}$$

Proof. Assume $\mathscr{V} = X_D \cap M$ for some submodule of $F^p[z]$. Clearly, $\mathscr{V} \subset M$ and hence $\langle \mathscr{V} \rangle \subset M$ and so $\mathscr{V} \subset \langle \mathscr{V} \rangle \subset M$, which in turn implies

$$\mathscr{V} \subset X_D \cap \langle \mathscr{V} \rangle \subset X_D \cap M = \mathscr{V}. \qquad \square$$

Corollary 3.1. *If $E \subset X_D$ is a subspace, then $X_D \cap \langle E \rangle$ is the smallest conditioned invariant subspace of X_D that contains E.*

Proof. $X_D \cap \langle E \rangle$ is conditioned invariant subspace and contains E. Let $\mathscr{W}$ be any other conditioned invariant subspace containing E. Then $\langle E \rangle \subset \langle \mathscr{W} \rangle$ and hence

$$X_D \cap \langle E \rangle \subset X_D \cap \langle \mathscr{W} \rangle = \mathscr{W}. \qquad \square$$

Proposition 3.3. *A subspace $\mathscr{V} \subset X_D$ is a conditioned invariant subspace if and only if it has a representation of the form*

$$\mathscr{V} = X_D \cap H(z)F^l[z], \tag{34}$$

where $H(z)$ is a full column rank $p \times l$ polynomial matrix whose columns are in $\mathscr{V}$. $H(z)$ is uniquely determined up to a right $l \times l$ unimodular factor.

Proof. Follows from Theorem 3.1 and the representation of submodules of $F^p[z]$. $\square$

So, to proceed, let $\mathscr{V} \subset X_D$ be a conditioned invariant subspace and let $\langle \mathscr{V} \rangle$ be the submodule of $F^p[z]$ generated by it. Let $H(z)$ be a basis matrix for $\langle \mathscr{V} \rangle$ with its columns contained in $\mathscr{V}$. In particular, this implies that $G = D^{-1}H$ is strictly proper.

Given an observable pair $(\mathscr{C}, \mathscr{A})$, let $\mathscr{V}$ be a conditioned invariant subspace. Let J be an output injection map such that $(\mathscr{A} + J\mathscr{C})\mathscr{V} \subset \mathscr{V}$. Then the *restricted pair* $(\overline{A}_{\mathscr{V}}, \overline{C}_{\mathscr{V}})$ acting in the state space $\mathscr{V}$ is defined by

$$\overline{A}_{\mathscr{V}} = (\mathscr{A} + J\mathscr{C}) \mid \mathscr{V},$$
$$\overline{C}_{\mathscr{V}} = \mathscr{C} \mid \mathscr{V}.$$

The pair $(\overline{C}_{\mathscr{V}}, \overline{A}_{\mathscr{V}})$ is also observable and has a set of observability indices, $\lambda_1, \ldots, \lambda_p \geqslant 0$, associated with it. We will refer to the λ_i as the *reduced observability indices*.

If in the representation $\mathscr{V} = X_D \cap T(z)F^p[z]$ of a conditioned invariant subspace the polynomial matrix T is nonsingular, then we can associate with T an isomorphism class of reachable pairs (A, B) acting in F^k, where $k = \deg\det T$. The column indices of T, namely $\kappa_1 \geqslant \cdots \geqslant \kappa_p \geqslant 0$, are invariant under right multiplication by unimodular matrices. They are equal to the controllability indices of (A_T, B_T). We refer to (A_T, B_T) as the *coreduced system* and to the κ_i as the *coreduced reachability indices*.

Theorem 3.2. *Let $\mathscr{V} \subset X_D$ be a conditioned invariant subspace and let $\langle\mathscr{V}\rangle$ be the submodule of $F^p[z]$ generated by it. Let $H(z)$ be a $p \times l$ basis matrix for $\langle\mathscr{V}\rangle$, i.e. the submodule $\langle\mathscr{V}\rangle$ has rank l. Let*

$$D^{-1}H = \Gamma(z)\begin{pmatrix} z^{-\delta_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z^{-\delta_l} \\ & & 0 & & \end{pmatrix} U(z), \tag{35}$$

where Γ is biproper and U polynomially unimodular, be a left Wiener–Hopf factorization. Then:

1. $\delta_i > 0$ *for all* $i = 1, \ldots, l$.
2. *We have* $\dim \mathscr{V} = d = \sum_{i=1}^{l} \delta_i$.
3. *There exists a basis matrix H for $\langle\mathscr{V}\rangle$ for which all the columns of H are contained in $\mathscr{V}$.*
4. *Let H be a basis matrix for $\langle\mathscr{V}\rangle$ for which all its columns H are contained in $\mathscr{V}$. Let $g_1, \ldots, g_l$ be the columns of $G = D^{-1}H$. Then the set of vectors*

$$\left\{\begin{matrix} g_1, zg_1, \ldots, z^{\delta_1-1}g_1 \\ \cdot \\ \cdot \\ \cdot \\ g_l, zg_l, \ldots, z^{\delta_l-1}g_l \end{matrix}\right\} \tag{36}$$

is a vector space basis for $\mathscr{W} = D^{-1}\mathscr{V}$.

5. *With respect to the shift realization corresponding to D in the state space X^D, we have the following matrix representation, with respect to the previous basis, of the realization of the reduced system.*

$$A^D = \left(\begin{array}{ccccc|ccc|ccccc}
0 & & & & * & & & & & & & & * \\
1 & . & & & * & & & & & & & & * \\
 & . & . & & * & & & & & & & & * \\
 & & . & . & * & & & & & & & & * \\
 & & & 1 & * & & & & & & & & * \\
\hline
 & & & & * & . & & & & & & & * \\
 & & & & * & & . & & & & & & * \\
 & & & & * & & & . & & & & & * \\
\hline
 & & & & * & & & & 0 & & & & * \\
 & & & & * & & & & 1 & . & . & & * \\
 & & & & * & & & & & & . & . & * \\
 & & & & * & & & & & & & . & * \\
 & & & & * & & & & & & & 1 & *
\end{array}\right), \tag{37}$$

$$C^D = \left(\begin{array}{ccccc|ccc|ccccc}
0 & . & . & 0 & 1 & & & & 0 & . & . & . & 0 \\
\hline
 & & & & & . & & & & & & & \\
 & & & & & & . & & & & & & \\
 & & & & & & & . & & & & & \\
\hline
0 & . & . & . & 0 & & & & 0 & . & . & . & 1
\end{array}\right).$$

Here, the principal blocks of A^D are of size $\delta_i \times \delta_i$ and those of C^D of size $1 \times \delta_i$.

6. *The indices δ_i, $i = 1, \ldots, l$, are the reduced observability indices corresponding to the conditioned invariant subspace $\mathscr{V}$.*

Proof.

1. Clearly, the basis matrix H is determined only up to a right unimodular factor. Setting $\mathscr{W} = D^{-1}\mathscr{V}$ and $G = D^{-1}H$, we have $\mathscr{W} = X^D \cap GF^l[z] = X^D \cap \langle G \rangle$. If one of the indices is nonpositive, say $\delta_i \leqslant 0$, then this would imply that $\mathscr{W}$ is generated by less than l elements, which is impossible as the rank is assumed to be l.
2. Clearly, $\dim \operatorname{Ker} \mathscr{T}_{D^{-1}H} = \sum_{\delta_i > 0} \delta_i = \sum_{i=1}^{l} \delta_i = d$. We use now Lemma 3.1 to infer that $\dim \mathscr{V} = \dim X_D \cap HF^l[z] = d$.
3. Rewrite the factorization (35) as

$$D^{-1}HU(z)^{-1} = \overline{\Gamma}(z) \begin{pmatrix} z^{-\delta_1} & & & \\ & . & & \\ & & . & \\ & & & . \\ & & & & z^{-\delta_l} \end{pmatrix},$$

where $\overline{\Gamma}$ is the $p \times l$ left biproper matrix consisting of the first l columns of Γ. Then, as the right-hand side is strictly proper, $HU(z)^{-1}$ is such a basis. We redefine H to incorporate the unimodular factor.

4. The vectors in (36) are linearly independent by the fact that $g_1, \ldots, g_l$ are a basis for $\langle \mathscr{W} \rangle$. They are all in $X^D \cap \langle \mathscr{W} \rangle$ and there are exactly d of them.
5. Obvious from the structure of the basis (36).
6. From the matrix representation (37), it is clear that there exists an output injection map that reduces (C^D, A^D) to dual Brunovsky form with observability indices $\delta_1, \ldots, \delta_l$. $\square$

The use of rectangular polynomial matrices is crucial for the parametrization of the set of conditioned invariant subspaces, a topic we will discuss in detail in Section 7. By Proposition 3.1, any conditioned invariant subspace represented by a rectangular polynomial matrix H, as in (34), has also a representation in terms of a nonsingular polynomial matrix T. This leads to different isomorphism classes of reachable pair. To keep degrees fixed we will have to analyze embeddings of H that do not increase McMillan degree.

However, the analysis for the case of a nonsingular T as in (30) is of great importance too. One of our first goals therefore is to determine the conditions equivalent to the uniqueness of the determination of a nonsingular T by the condition invariant subspace $\mathscr{V}$. This leads to the introduction and analysis of tight subspaces. Once we are in the situation of uniqueness, up to a right unimodular factor, of the nonsingular polynomial matrix T, we can associate with it a unique isomorphism class of reachable pairs (A, B) that are associated with the nonsingular polynomial matrix T. The connection is given via the relation

$$(zI - A)^{-1}B = \Phi(z)T(z)^{-1}, \tag{38}$$

i.e. via the shift realization procedure. Here the matrix fraction representation $\Phi(z)T(z)^{-1}$ is assumed to be right coprime. Thus the concept of tightness introduced formally below will allow us to study the set of conditioned invariant subspaces via reachable pairs.

On the other hand, one might start with a reachable pair, using Eq. (38) to obtain a nonsingular polynomial matrix T, unique up to a right unimodular factor and define a conditioned invariant subspace via (30). Exploring these possibilities will be one of the main goals of this paper. It is of interest to characterize the conditions that guarrantee that the such obtained subspace is tight.

Our main objective is to parametrize, given an observable pair $(\mathscr{C}, \mathscr{A})$, the set of all conditioned invariant subspaces of a given (co)dimension. Without loss of generality we can assume that $(\mathscr{C}, \mathscr{A})$ is in Brunovsky canonical form. This means that, with the coprime factorizations $\mathscr{C}(zI - \mathscr{A})^{-1} = D(z)^{-1}\Theta(z)$, we have

$$D(z) = \begin{pmatrix} z^{\mu_1} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & z^{\mu_p} \end{pmatrix}, \tag{39}$$

with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$, the observability indices of $(\mathscr{C}, \mathscr{A})$. By Theorem 3.1, any conditioned invariant subspace $\mathscr{V} \subset X_D$ has a representation of the form $\mathscr{V} = X_D \cap T(z)F^p[z]$, for some polynomial matrix $T(z)$.

Let $(A, B) \in F^{k\times k} \times F^{k\times p}$ be an arbitrary reachable pair. We consider the coprime factorizations $(zI - A)^{-1}B = H(z)T(z)^{-1}$. Using the characterizations of Theorem 3.1, we have a mapping from reachable pairs to conditioned invariant subspaces, given by

$$(A, B) \mapsto X_D \cap T(z)F^p[z]. \tag{40}$$

We will refer to this map as the *structural map*. Note that $T(z)$ is defined only up to a right unimodular factor, however the conditioned invariant subspace $X_D \cap T(z)F^p[z]$ is uniquely determined by the state space similarity equivalence class of (A, B). The structural map is all important in our study. Among other things it allows us to study the set of conditioned invariant subspaces via the set of reachable pairs of appropriate dimension. The uniqueness question for representations of the form $\mathscr{V} = X_D \cap TF^p[z]$ with T nonsingular is fully analyzed in Theorem 3.5. However, this theorem treats essentially only the case where the submodule of $F^p[z]$ generated by the conditioned invariant subspace $\mathscr{V}$ is a full submodule. We return to the general case in Theorem 3.2.

A moments reflection will show that the structural map may have large fibers. For it to become an interesting object, we need to restrict its domain of definition, i.e. the class of observable pairs under consideration. In fact, we will show that if A is a $k \times k$ matrix, then, for $k < \mu_p$, the structural map is bijective. Moreover, we have the codimension formula $\operatorname{codim} \mathscr{V} = \deg\det T(z)$. To get a better understanding, we connect the study of the structural map with that of Toeplitz operators. We introduce a special case of Toeplitz operators that fits the algebraic context in which we are interested. A convenient general reference for Wiener–Hopf theory is Ref. [22].

Lemma 3.1. *Let D and T be $p \times p$ and $m \times p$ polynomial matrices, respectively, and assume D is nonsingular. Let $\mathscr{T}_{D^{-1}T}$ be the Toeplitz operator defined in* (13). *The map $\psi : \operatorname{Ker}\mathscr{T}_{D^{-1}T} \longrightarrow X_D \cap TF^p[z]$ defined by*

$$\psi(p) = Tp$$

is a surjective linear map. If T has full column rank, then ψ is also injective and we have

$$\dim \operatorname{Ker}\mathscr{T}_{D^{-1}T} = \dim X_D \cap TF^p[z]. \tag{41}$$

Proof. Let $p \in \operatorname{Ker}\mathscr{T}_{D^{-1}T}$, i.e. $\mathscr{T}_{D^{-1}T}p = \pi_+ D^{-1}Tp = 0$. Setting $f = \psi(p) = Tp$, we get $f \in X_D$ as well as $f \in TF^p[z]$, that is $f \in X_D \cap TF^p[z]$.

Conversely, assume $f \in X_D \cap TF^p[z]$. This implies that for some strictly proper h and polynomial p we have $f = Dh = Tp = \psi(p)$. So $h = D^{-1}Tp$ and hence $p \in \operatorname{Ker}\mathscr{T}_{D^{-1}T}$. This shows that the map ψ is surjective.

If T has full column rank, it is left invertible. Hence $Tp = 0$ implies $p = 0$, i.e. ψ is injective. This implies the two spaces are isomorphic, and hence equality (41) follows. $\square$

The following theorem has been proved in [11,12]. We quote it, omitting its simple proof.

Lemma 3.2. *Let T, D be nonsingular polynomial matrices. Then there exists a polynomial matrix S for which $D^{-1}TS$ is biproper if and only if all left Wiener–Hopf factorization indices of $D(z)^{-1}T(z)$ are nonpositive.*

The Wiener–Hopf factorization indices give insight into several basic questions. The connection, via coprime factorizations, to the analysis of controllability and observability indices is well known, see [21]. In the present analysis we use them to derive two important results. First, they give a necessary and sufficient condition for the computation of the codimension of a conditioned invariant subspace $\mathscr{V} = X_D \cap TF^p[z]$ in terms of T. This is done in Proposition 3.4. Secondly, we shall look into the essential uniqueness of a representation of a conditioned invariant subspace in the form $\mathscr{V} = X_D \cap TF^p[z]$. This question will be taken up in Theorem 3.5.

Proposition 3.4. *Let D and T be nonsingular, $p \times p$ polynomial matrices. Then, all left Wiener–Hopf factorization indices of $D(z)^{-1}T(z)$ are nonpositive if and only if we have the codimension formula*

$$\operatorname{codim} X_D \cap TF^p[z] = \deg\det T. \tag{42}$$

Proof. Assume all left Wiener–Hopf factorization indices of $D(z)^{-1}T(z)$ are nonpositive. By Lemma 3.2, there exists a polynomial matrix S for which $D^{-1}TS$ is biproper. We have in this case

$$\mathscr{V} = X_D \cap T(z)F^p[z] = TX_S$$

and $\dim \mathscr{V} = \dim TX_S = \deg\det S$. Now, since $D^{-1}TS$ is biproper, we have

$$\deg\det D = \deg\det(TS) = \deg\det T + \deg\det S.$$

So, $\operatorname{codim} \mathscr{V} = \deg\det D - \deg\det S = \deg\det T$.

Conversely, assume the codimension formula (42) holds. Thus necessarily T is nonsingular. Assume the left factorization indices of $D^{-1}T$ are $\lambda_1 \geqslant \cdots \geqslant \lambda_k > 0 \geqslant \lambda_{k+1} \geqslant \cdots \geqslant \lambda_p$. Set

$$\Delta_+(z) = \operatorname{diag}(z^{\lambda_1}, \ldots, z^{\lambda_k}, 1, \ldots, 1)$$

and

$$\Delta_-(z) = \operatorname{diag}(1, \ldots, 1, z^{\lambda_{k+1}}, \ldots, z^{\lambda_p}).$$

The left Wiener–Hopf factorization of $D^{-1}T$ has therefore the form

$$D^{-1}T = \Gamma\Delta_-\Delta_+U, \tag{43}$$

with U unimodular and Γ biproper. Using our assumption and applying Proposition 3.4 twice, we compute

$$\begin{aligned}\deg\det D - \deg\det T &= \dim X_D \cap T(z)F^p[z] = \dim\operatorname{Ker}\mathscr{T}_{D^{-1}T}\\ &= \dim\operatorname{Ker}\mathscr{T}_{\Gamma\Delta_-\Delta_+U} = \dim\operatorname{Ker}\mathscr{T}_{\Delta_-\Delta_+}\\ &= \dim X_{\Delta_-^{-1}} \cap \Delta_+(z)F^p[z] = \dim X_{\Delta_-^{-1}}\\ &= -\sum_{i=k+1}^{p}\lambda_i.\end{aligned}$$

Thus we conclude that $\deg\det D + \sum_{i=k+1}^{p}\lambda_i = \deg\det T$, i.e., $\dim\mathscr{V} = -\sum_{i=k+1}^{p}\lambda_i$. On the other hand, from the factorization (43), we obtain

$$\deg\det T = \deg\det D + \sum_{i=1}^{k}\lambda_i - \sum_{i=k+1}^{p}\lambda_i.$$

Comparing the two expressions, we conclude that $\sum_{i=1}^{k}\lambda_i = 0$, i.e. all the factorization indices are nonpositive. □

It will be convenient for the rest of the paper to introduce new concepts.

Definition 3.1. Given nonsingular polynomial matrices $D, T \in F^{p\times p}[z]$, we will say that

1. T is (D-proper) (D-strictly proper) if $D^{-1}T$ is (strictly) proper.
2. T is D-regular if all the left Wiener–Hopf factorization indices of $D^{-1}T$ are nonpositive.
3. T is D-tight if all the left Wiener–Hopf factorization indices of $D^{-1}T$ are negative.

Clearly, D-properness implies D-regularity, and D-strict properness implies D-tightness. Moreover, D-regularity of T is equivalent to the codimension formula (42) holding.

Our next aim is to discuss the conditions under which the representation of a conditioned invariant subspace $\mathscr{V}$ in the form $\mathscr{V} = X_D \cap TF^p[z]$ is essentially unique, i.e. T is determined up to a left unimodular factor. That such a representation is not unique in general is easily seen. In fact, for every polynomial matrix E, we have $X_D \cap DEF^p[z] = \{0\}$. The answer to this question has to do with reduced observability indices and, in turn, with Wiener–Hopf factorizations.

We state and prove now a result on the majorization of the observability indices over the reduced observability indices. This was stated without proof in [16]. A dual result has been proved earlier in [30] by different methods. This result is of interest not only by itself, but also because of the proof technique which we now give in full.

Proposition 3.5. *Let (C, A) be an observable pair with the observability indices $\mu_1 \geqslant \cdots \geqslant \mu_p$, and let $\mathscr{V}$ be a conditioned invariant subspace. Let J be a friend of $\mathscr{V}$, i.e. an output injection map such that $(A + JC)\mathscr{V} \subset \mathscr{V}$. Then:*

1. *The restricted pair (A_1, C_1) acting in the state space $\mathscr{V}$ and defined by*

$$A_1 = (A + JC) \mid \mathscr{V},$$
$$C_1 = C \mid \mathscr{V}$$

is observable.

2. *The reduced observability indices $\lambda_1 \geqslant \cdots \geqslant \lambda_p$ satisfy*

$$\lambda_i \leqslant \mu_i, \quad i = 1, \ldots, p. \tag{44}$$

Proof.

1. Choosing an arbitrary complementary subspace to $\mathscr{V}$, we obtain the block matrix representations

$$A + JC = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}, \quad C = \begin{pmatrix} C_1 & C_2 \end{pmatrix}.$$

Assume $\xi \in \bigcap_i \operatorname{Ker} C_1 A_1^i$, then $\binom{\xi}{0} \in \bigcap_i \operatorname{Ker} CA^i$ and hence necessarily 0. So $\xi = 0$ and (A_1, C_1) is observable.

2. Let $T(z)^{-1}H(z)$ be a left coprime factorization of $C(zI - A)^{-1}$. Without loss of generality we can assume that T is in row proper form with row indices $\mu_1 \geqslant \cdots \geqslant \mu_p$, X_T is the state space and the pair (C, A) is given by $A = S_T$ and $Cf = (T^{-1}f)_{-1}$ for all $f \in X_T$. The assumption of row properness implies that the μ_i are the observability indices of the pair (C, A).

Since $\mathscr{V}$ is conditioned invariant, there exists a nonsingular polynomial matrix $T_1 = E_1 F_1$ such that $T_1^{-1}T$ is biproper and $\mathscr{V} = E_1 X_{F_1}$. With $E_1 X_{F_1}$ as the state space, the reduced system is given by $A_1 = S_{T_1} \mid E_1 X_{F_1}$ and $C_1 E_1 f = (T_1^{-1} E_1 f)_{-1} = (F_1^{-1} f)_{-1}$. Thus (A_1, C_1) is isomorphic to the pair $(S_{F_1}, (F_1^{-1}\cdot)_{-1})$. Again, without loss of generality, we may assume that F_1 is row proper and $\lambda_1 \geqslant \cdots \geqslant \lambda_p$ are its row indices and hence equal to the observability indices of the reduced system. By our assumption of row properness we get

$$X_{T_1} = X_{z^{\mu_1}} \oplus \cdots \oplus X_{z^{\mu_p}} = \left\{ \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_p \end{pmatrix} \middle| \deg f_i < \mu_i \right\}.$$

Similarly,

$$X_{F_1} = X_{z^{\lambda_1}} \oplus \cdots \oplus X_{z^{\lambda_p}} = \left\{ \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_p \end{pmatrix} \middle| \deg f_i < \lambda_i \right\},$$

and of course we have $E_1 X_{F_1} \subset X_{T_1}$. We will show, by induction, that $\lambda_i \leqslant \mu_i$. If we assume $\lambda_1 > \mu_1$, then, for the first column of the polynomial matrix E_1, we must have

$$E_1 z^i = \begin{pmatrix} e_1^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ e_p^{(1)} \end{pmatrix} z^i \in X_{z^{\mu_1}} \oplus \cdots \oplus X_{z^{\mu_p}}$$

for $i = 0, \ldots, \lambda_1 - 1$, and in particular for $i = \mu_1$. Thus the first column is necessarily 0, contradicting the nonsingularity of E_1.

Assume now that $\lambda_i \leqslant \mu_i$ for $i = 1, \ldots, k$ and $\lambda_{k+1} > \mu_{k+1}$. Note that our assumptions imply $\mu_k > \mu_{k+1}$. Considerations as before imply $e_j^{(i)} = 0$ for $i = 1, \ldots, k$; $j = k+1, \ldots, p$. Furthermore, for the $k+1$th column we get $e_j^{(k+1)} = 0$ for $j = k, \ldots, p$. Once again, this implies the singularity of E_1 and leads to contradiction. Thus $\lambda_i \leqslant \mu_i$ for all i. $\square$

The factorization indices corresponding to a conditioned invariant subspace have a nice interpretation in terms of the reduced observability indices.

Proposition 3.6. *Given a nonsingular, $p \times p$ polynomial matrix D, let $\mathscr{V} = X_D \cap TF^p[z]$ be a conditioned invariant subspace. We assume without loss of generality that all left Wiener–Hopf factorization indices of $D^{-1}T$ are nonpositive. Then the reduced observability indices are the negatives of the left Wiener–Hopf factorization indices of $D^{-1}T$.*

Proof. By Lemma 3.2, there exists a polynomial matrix R such that, with $D_1 := TR$, $D_1^{-1}D$ is biproper. This implies $\mathscr{V} = TX_R$. The reduced system is given by

$$\begin{aligned} &\overline{A} = S_{D_1} \mid TX_R \simeq S_R = A_R, \\ &\overline{C}(Tg) = (D_1^{-1}Tg)_{-1} = (R^{-1}g)_{-1} \simeq C_R. \end{aligned}$$

So $(\overline{C}, \overline{A})$ is isomorphic to (C_R, A_R). Since the observability indices of (C_R, A_R) are equal to the row indices of R, it follows that the row indices of R are equal to the reduced observability indices.

Now, let $R = U_R \Delta \Gamma_R$ be a right Wiener–Hopf factorization. Since $D_1 = TR$, it follows that $D^{-1}D_1 = D^{-1}TR$ or $D^{-1}T = (D^{-1}D_1)R^{-1} = (\Gamma\Gamma_R^{-1})\Delta^{-1}U_R^{-1}$. This is a left Wiener–Hopf factorization of $D^{-1}T$ and the factorization indices are the negatives of the right factorization indices of R. $\square$

Lemma 3.3. *Let $D(z) = \mathrm{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$ and $\mu_1 + \cdots + \mu_p = n$. Assume $(C, A) \in F^{p\times k} \times F^{k\times k}$ is an observable pair. Let $T(z)^{-1}H(z)$ be a left coprime factorization of $C(zI - A)^{-1}$. Then, with $\pi^D : z^{-1}F^p[[z^{-1}]] \longrightarrow z^{-1}F^p[[z^{-1}]]$, the projection defined by* (4), *i.e. by*

$$\pi^D h = \pi_- D^{-1}\pi_+ Dh \quad \textit{for } h \in z^{-1}F^p[[z^{-1}]],$$

the following diagram is commutative.

$$\begin{tikzcd}
X_T \arrow[r, "\pi_D \mathcal{T}_{DT^{-1}}|X_T"] \arrow[d, "T^{-1}"'] & X_D \arrow[d, "D^{-1}"] \\
X^T \arrow[r, "\pi^D|X^T"] & X^D
\end{tikzcd}$$

Proof. Note that $f \in X_T$ implies that $T^{-1}f$ is strictly proper. Hence, we compute

$$D^{-1}\pi_D \mathscr{T}_{DT^{-1}} f = D^{-1}D\pi_- D^{-1}\pi_+ DT^{-1}f = \pi^D(T^{-1}f). \qquad \square$$

Theorem 3.3. *Let $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$ and $\mu_1 + \cdots + \mu_p = n$. Assume $(A, B) \in F^{k\times k} \times F^{k\times p}$ is a reachable pair. Let $H(z)T(z)^{-1}$ be a right coprime factorization of $(zI - A)^{-1}B$. Then the following statements are equivalent*:

1. *All the left factorization indices of $D^{-1}T$, $-\lambda_i$ are nonpositive. We assume $\lambda_1 \geqslant \cdots \geqslant \lambda_p \geqslant 0$.*
2. *The Toeplitz map $\mathscr{T}_{D^{-1}T} : F^m[z] \longrightarrow F^m[z]$ is surjective.*
3. *The Toeplitz map $\hat{\mathscr{T}}_{\tilde{T}\tilde{D}^{-1}} : z^{-1}F^m[[z^{-1}]] \longrightarrow z^{-1}F^m[[z^{-1}]]$ defined by $h \mapsto \pi_-\tilde{T}\tilde{D}^{-1}h$ is injective.*
4. *The Toeplitz map $\mathscr{T}_{\tilde{D}\tilde{T}^{-1}} : F^m[z] \longrightarrow F^m[z]$ is injective.*
5. *The induced Toeplitz map $\pi_{\tilde{D}}\mathscr{T}_{\tilde{D}\tilde{T}^{-1}} : X_{\tilde{T}} \longrightarrow X_{\tilde{D}}$ is injective.*
6. *The map $\pi^{\tilde{D}} : X^{\tilde{T}} \longrightarrow X^{\tilde{D}}$ is injective.*
7. *The map $\pi_T : X_D \longrightarrow X_T$ is surjective.*
8. *We have* $\operatorname{codim} X_D \cap T(z)F^p[z] = \deg\det T$.

Proof. (1) $\Leftrightarrow$ (2) Follows from Proposition 2.2.

(2) $\Leftrightarrow$ (3) Follows by duality.

(1) $\Leftrightarrow$ (4) The left Wiener–Hopf factorization indices of $\tilde{D}\tilde{T}^{-1}$ are the negatives of the left Wiener–Hopf factorization indices of $D^{-1}T$ and hence are nonnegative. This is equivalent to the injectivity of $\mathscr{T}_{\tilde{D}\tilde{T}^{-1}}$.

(4) $\Leftrightarrow$ (5) We clearly have $\mathscr{T}_{\tilde{D}\tilde{T}^{-1}}(\tilde{T}F^p[z]) = \tilde{D}F^p[z]$. This shows that the induced map $\overline{\mathscr{T}}_{\tilde{D}\tilde{T}^{-1}} : F^p[z]/\tilde{T}F^p[z] \longrightarrow F^p[z]/\tilde{D}F^p[z]$ is injective if and only if the Toeplitz map $\mathscr{T}_{\tilde{D}\tilde{T}^{-1}}$ is injective. The two-factor spaces are isomorphic to the polynomial models $X_{\tilde{T}}$ and $X_{\tilde{D}}$, respectively. The isomorphism from $X_{\tilde{T}}$ to

$F^p[z]/\tilde{T}F^p[z]$ is given by $f \mapsto [f]$, whereas the isomorphism from $F^p[z]/\tilde{D}F^p[z]$ to $X_{\tilde{D}}$ is given by $[g] \mapsto \pi_{\tilde{D}}g$. We compute, for $f \in X_{\tilde{T}}$, the composition of the three injective maps

$$f \mapsto [f] \mapsto \mathscr{T}_{\tilde{D}\tilde{T}^{-1}}[f] = [\mathscr{T}_{\tilde{D}\tilde{T}^{-1}}f] \mapsto \pi_{\tilde{D}}\mathscr{T}_{\tilde{D}\tilde{T}^{-1}}f,$$

which shows the injectivity of $\pi_{\tilde{D}}\mathscr{T}_{\tilde{D}\tilde{T}^{-1}} : X_{\tilde{T}} \longrightarrow X_{\tilde{D}}$. From the commutativity of the diagram in Lemma 3.3, it follows that the map $\pi^{\tilde{D}} : X^{\tilde{T}} \longrightarrow X^{\tilde{D}}$ is also injective.

(5) $\Leftrightarrow$ (6) Follows from the isomorphism of $\pi_{\tilde{D}}\mathscr{T}_{\tilde{D}\tilde{T}^{-1}}$ and $\pi^{\tilde{D}} \mid X^{\tilde{T}}$ given by the commutativity of Diagram 1.

(6) $\Leftrightarrow$ (7) Follows from the fact that $\pi_T \mid X_D$ is the dual to $\pi^{\tilde{D}} \mid X^{\tilde{T}}$.

(7) $\Leftrightarrow$ (8) Clearly

$$\operatorname{Ker}\pi_T \mid X_D = X_D \cap T(z)F^p[z].$$

If $\pi_T \mid X_D$ is surjective, then

$$\dim X_T = \deg\det T = \deg\det D - \dim X_D \cap T(z)F^p[z]$$

or $\operatorname{codim} X_D \cap T(z)F^p[z] = \deg\det T$.

This argument is reversible. $\square$

This theorem has a counterpart in the context of rational models and controlled invariant subspaces. We omit the proof.

Theorem 3.4. *Let $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$ and $\mu_1 + \cdots + \mu_p = n$. Assume $(A, B) \in F^{k\times k} \times F^{k\times p}$ is a reachable pair. Let $H(z)T(z)^{-1}$ be a right coprime factorization of $(zI - A)^{-1}B$. Then the following statements are equivalent*:

1. *All the right factorization indices of TD^{-1}, are nonpositive. We assume $\lambda_1 \geqslant \cdots \geqslant \lambda_p \geqslant 0$.*
2. *$\pi^D \mid X^T$ is injective.*
3. *We have $\dim \pi^D X^T = \deg\det T$.*

Corollary 3.2. *Let the polynomial matrices D and T be defined as in Theorem* 3.3. *A sufficient condition for the injectivity of $\pi^D \mid X^T$ is $\delta = \deg\det T \leqslant \mu_p$.*

Proof. A strictly proper function

$$h = \begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ \cdot \\ h_p \end{pmatrix}$$

belongs to $\operatorname{Ker}\pi^D$ if and only if $h_i \in z^{-(\mu_i+1)}F[[z^{-1}]]$. Now $h \in X^T$ if and only if $h = T^{-1}g$ for some polynomial vector g. By Cramer's rule, we have $T^{-1}g =$

$\operatorname{adj} T\ g/\det T$. It follows that $h = T^{-1}g$, if it is nonzero, has a term of the form η/z^{α} for some $\alpha \leqslant \deg\det T$, and hence h cannot be in $\operatorname{Ker}\pi^{D}$. $\square$

We consider now conditioned invariant subspaces $\mathscr{V} \subset \mathscr{X}$, with respect to an observable pair $(\mathscr{C}, \mathscr{A})$. We say, see [16], that a conditioned invariant subspace $\mathscr{V}$ as a *tight* conditioned invariant subspace if

$$\mathscr{V} + \operatorname{Ker}\mathscr{C} = \mathscr{X}$$

holds, i.e. if $\mathscr{V}$ is transversal to $\operatorname{Ker}\mathscr{C}$. A tight subspace is the dual object to that of a coasting subspace, introduced in [35],

The following theorem, that generalizes a result from [16], clarifies the various conditions that characterize the uniqueness of a representation of a conditioned invariant subspace with respect to the shift realization.

Theorem 3.5. *Let* $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ *with* $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$, *and let*

$$\mathscr{V} = X_D \cap TF^p[z] \tag{45}$$

be a conditioned invariant subspace. Let $\lambda_1 \geqslant \cdots \geqslant \lambda_p$ *be the reduced observability indices, i.e. the observability indices of the system* (C_D, A_D) *reduced to* $\mathscr{V}$. *Then the following conditions are equivalent*:

1. *All left Wiener–Hopf indices of* $D^{-1}T$ *are negative.*
2. *The reduced observability indices are all positive, i.e.* $\lambda_p > 0$.
3. *Representation* (45) *of* $\mathscr{V}$ *is unique up to a right unimodular factor for T.*
4. $M_{\mathscr{V}}$, *the smallest submodule of* $F^p[z]$ *that includes* $\mathscr{V}$, *is full, i.e. has p generators which are linearly independent over* $F[z]$.
5. *The factor module* $F^p[z]/M_{\mathscr{V}}$ *is a torsion module.*
6. $\mathscr{V}$ *is a tight conditioned invariant subspace of* X_D.

Proof. (1) $\Leftrightarrow$ (2) Assume all the left Wiener–Hopf factorization indices of $D^{-1}T$ are negative. Thus we can write $D^{-1}T = \Gamma^{-1}\Delta^{-1}U^{-1}$, where

$$\Delta(z) = \begin{pmatrix} z^{\lambda_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z^{\lambda_p} \end{pmatrix} \tag{46}$$

with $\lambda_i > 0$, U unimodular and Γ biproper. Clearly, $S = U\Delta$ is a polynomial matrix and hence $D^{-1}TS = \Gamma^{-1}$ is biproper. This implies that the polynomial models X_{TS} and X_D have the same elements (although they carry different module structures). In this case we have

$$\mathscr{V} = X_D \cap T(z)F^p[z] = X_{TS} \cap T(z)F^p[z] = T(z)X_S.$$

Since $S = U\Delta$, the row indices of S are $\lambda_1 \geqslant \cdots \geqslant \lambda_p$. In fact, the row indices of S are the observability indices of the reduced system, hence are all positive.

Conversely, assume all the reduced observability indices are positive, i.e. $\lambda_i > 0$. Then $\mathscr{V}$ has a representation $\mathscr{V} = EX_S$ with the reduced observability indices, equal to the row indices of S, positive. Now $T^{-1}ES = \Gamma$, with Γ biproper. Let $S = U\Delta\Gamma_1$ be a right Wiener–Hopf factorization, with $\Delta(z) = \operatorname{diag}(z^{\lambda_1}, \dots, z^{\lambda_p})$. So $D^{-1}E = U^{-1}\Delta^{-1}\Gamma^{-1}\Gamma$, which is also a right Wiener–Hopf factorization. Thus, all the left Wiener–Hopf indices of $D^{-1}E$ are negative.

(2) $\Leftrightarrow$ (3) Assume $\lambda_1 \geqslant \cdots \geqslant \lambda_p > 0$. $\mathscr{V}$ has another representation of the form $\mathscr{V} = T_1X_{S_1} = X_D \cap T_1F^p[z]$, with $D^{-1}T_1S_1$ biproper. Without loss of generality we can assume that S_1 is row proper with row indices $\lambda_1 \geqslant \cdots \geqslant \lambda_p$. Now, by assumption,

$$\mathscr{V} = T_1X_{S_1} = X_D \cap T_1F^p[z] = X_D \cap TF^p[z]. \tag{47}$$

Since all λ_i are positive, X_{S_1} contains all constant polynomials. The previous equality implies therefore that $T_1(z) = T(z)E(z)$ for some polynomial matrix E. Thus, necessarily, both T and E are nonsingular. Defining $D_1 = T_1S_1$, we have $D_1 = T_1S_1 = TES_1$. Therefore we have

$$\begin{aligned}\mathscr{V} = X_D \cap TF^p[z] &= T_1X_{S_1} = TEX_{S_1}\\ &\subset TX_{ES_1} = X_D \cap TF^p[z].\end{aligned}$$

Thus we must have equality throughout and hence E is necessarily unimodular and $T_1 = TE$.

Conversely, we will show that if not all the reduced observability indices are positive then T is not uniquely determined. $\mathscr{V}$ has a representation of the form $\mathscr{V} = TX_S$, with $D^{-1}TS$ biproper. We can assume that S is column proper with column indices $\lambda_1 \geqslant \cdots \lambda_i > 0 \geqslant \lambda_{i+1} = \cdots = \lambda_p$. All elements of X_S are vector polynomials, with the last $p - i$ coordinates equal to 0. Thus, we can multiply T on the right by a polynomial matrix of the block form

$$\begin{pmatrix} I & 0 \\ 0 & R(z) \end{pmatrix},$$

with $R(z)$ an arbitrary $(p - i) \times (p - i)$ polynomial matrix.

This shows that T is not uniquely determined.

(3) $\Leftrightarrow$ (4) Assume that representation (45) is essentially unique. By the equivalence of (2) and (3), the reduced observability indices are all positive. Using the alternative representation $\mathscr{V} = TX_S$, where we assume without loss of generality that S is column proper, it follows that X_S contains all constant polynomials. Therefore M_{X_S}, the smallest submodule of $F^p[z]$ that contains X_S is $F^p[z]$. Hence

$$M_{\mathscr{V}} = M_{TX_S} = TM_{X_S} = TF^p[z],$$

which is a full submodule.

Conversely, assume that $M_{\mathscr{V}}$ is a full submodule. Let $f_1, \dots, f_p \in \mathscr{V}$ be linearly independent over $F[z]$. This implies that all observability indices of the reduced system are positive. By the equivalence of (2) and (3), T is uniquely determined.

(4) $\Leftrightarrow$ (5) That the quotient module $F^p[z]/M_{\mathscr{V}}$ is torsion if and only if $M_{\mathscr{V}}$ is full is well known, see [11,12].

(4) $\Leftrightarrow$ (6) Assume $M_{\mathscr{V}}$ is full. There exist therefore $g_1, \ldots, g_p \in \mathscr{V}$, linearly independent over $F[z]$. Since for every $g \in \mathscr{V} \cap \operatorname{Ker} C$, $zg \in \mathscr{V}$, we can multiply each g_i by a maximal power of z that leaves it in $\mathscr{V}$. This does not affect the linear independence of the g_i. Without loss of generality we may assume that the polynomial matrix G, whose columns are the g_i is column proper with highest coefficient matrix equal to I. Thus the column degrees are necessarily equal to $\mu_1 - 1, \ldots, \mu_p - 1$. This shows, noting that $\operatorname{Ker} C = \{g = (g_1, \ldots, g_p) \in X_D \mid \deg g_i < \mu_i\}$, that $\mathscr{V} + \operatorname{Ker} C = X_D$, i.e. $\mathscr{V}$ is tight.

Conversely, assume $\mathscr{V} + \operatorname{Ker} C = \mathscr{X} = X_D$. Since $\dim \operatorname{Ker} C = \deg \det T - p = n - p$, there exist $g_1, \ldots, g_p \in \mathscr{V}$, linearly independent, for which $\operatorname{Ker} C + L(g_1, \ldots, g_p) = X_D$. Let $e_1, \ldots, e_p$ be the unit vectors in F^p, and set $f_i = z^{\mu_i - 1} e_i$. Then $\operatorname{Ker} C + L(f_1, \ldots, f_p) = X_D$. Thus there exist constants α_{ij} and $q_i \in \operatorname{Ker} C$ such that $f_i = \sum_{j=1}^p \alpha_{ij} g_j + q_i$, or $h_i = f_i - q_i = \sum_{j=1}^p \alpha_{ij} g_j$. The polynomial matrix $H(z) = (h_1, \ldots, h_p)$ is obviously column proper. As $H(z) = AG(z)$, both A and $G(z)$ are nonsingular. Hence, $M_{L(g_1,\ldots,g_p)}$ is a full submodule, and because of the inclusion $M_{L(g_1,\ldots,g_p)} \subset M_{\mathscr{V}}$, so is $M_{\mathscr{V}}$. $\square$

One of the equivalent characterizations of tightness given in Theorem 3.5 was in terms of the reduced observability indices. One might expect that a characterization of tightness could also be given in terms of the coreduced reachability indices. This is not generally true, however we can state the following.

Proposition 3.7. *Let the polynomial matrix D be fully generic, i.e. $D(z) = z^\mu I$. Let $\mathscr{V} = X_D \cap T(z)F^p[z]$ with T nonsingular. Let $\lambda_1 \geqslant \cdots \geqslant \lambda_p$ be the reduced observability indices and let κ_i be the coreduced reachability indices, i.e. the column indices of T, which for convenience we take increasingly ordered, i.e. $\kappa_1 \leqslant \cdots \leqslant \kappa_p$. Then*:

1. *We have*

$$\lambda_i + \kappa_i = \mu, \qquad i = 1, \ldots, p. \tag{48}$$

2. *The inequality $\kappa_p < \mu$ implies the tightness of $\mathscr{V}$.*

Proof.

1. We have the left Wiener–Hopf factorization

$$D^{-1}T = \Gamma \operatorname{diag}(z^{-\lambda_1}, \ldots, z^{-\lambda_p})U.$$

Using the special form of D, we get

$$T = \Gamma \operatorname{diag}(z^{\mu-\lambda_1}, \ldots, z^{\mu-\lambda_p})U,$$

which shows that $\kappa_i = \mu - \lambda_i$.

2. If $\kappa_p < \mu$, then this inequality holds for all indices, i.e. $\kappa_i < \mu$. From the set of equalities (48) we conclude that $\lambda_i > 0$. Applying Theorem 3.5, we conclude that $\mathscr{V}$ is tight. $\square$

The equality $\lambda_i + \kappa_i = \mu_i$ which holds in the fully generic case discussed above is not true in general. For a counterexample, see Example 7.2.

We point out that with a slight modification the results of Proposition 3.7 hold in the generic case, where

$$\mu_i = \begin{cases} \mu, & i = 1, \dots, k, \\ \mu - 1, & i = k, \dots, p. \end{cases}$$

We omit the proof. Thus even in this case we have the equalities $\lambda_i + \kappa_i = \mu_i$. Thus, by the reasoning of Proposition 3.7, we have $\kappa_i < \mu_i$ implies tightness.

As in the case of Theorem 3.3, Theorem 3.5 has a counterpart in the context of rational models and controlled invariant subspaces. The proof is by use of duality and is omitted.

Theorem 3.6. *Let $D(z) = \operatorname{diag}(z^{\mu_1}, \dots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$, and let*

$$\mathscr{W} = \pi^D X^T \tag{49}$$

be a controlled invariant subspace. Let $\kappa_1 \geqslant \cdots \geqslant \kappa_p$ be the induced reachability indices associated with $T(z)$, i.e. the reachability indices of the system (A^T, B^T) induced in $\mathscr{W}$. Then the following conditions are equivalent:

1. *All right Wiener–Hopf indices of TD^{-1} are negative.*
2. *The induced controllability indices are all positive, i.e. $\kappa_p > 0$.*
3. *Representation* (49) *of $\mathscr{W}$ is unique up to a left unimodular factor for T.*
4. *$\mathscr{W}$ is a coasting controlled invariant subspace of X^D.*

The following proposition describes sufficient conditions for the codimension formula to hold as well as for the injectivity of the structural map defined in (40).

Proposition 3.8. *Let (A, B) be an arbitrary reachable pair, defined in a k-dimensional state space. Let $(zI - A)^{-1}B = H(z)T(z)^{-1}$ be coprime factorizations and assume $D(z) = \operatorname{diag}(z^{\mu_1}, \dots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$. Then*:

1. *If $k \leqslant \mu_p$, then the codimension formula* $\operatorname{codim} X_D \cap TF^p[z] = \deg\det T$ *holds.*
2. *If $k < \mu_p$, then the representation of the conditioned invariant subspace $\mathscr{V} = X_D \cap TF^p[z]$ is essentially unique, i.e. T is uniquely determined up to a right unimodular factor.*

Proof.

1. Since a polynomial matrix can be reduced to column proper form by elementary row operations, see [37], we can assume without loss of generality that T is in this form with column indices $\lambda_i \leqslant k < \mu_p$. Thus $D^{-1}T$ is proper and hence its right factorization indices are nonpositive. We apply now Lemma 3.4.

2. In this case we can assume without loss of generality that $D^{-1}T$ is strictly proper and hence the left factorization indices are all negative, so we can apply Theorem 3.5. $\square$

4. The state space approach

The characterization of conditioned invariant subspaces given in Section 3 used polynomial and rational models. Our intention in this section is to give a state space interpretation of these results. The key to this interpretation is the shift realization. To get matrix results, all one needs is to go to matrix representations with respect to appropriate choice of bases. Since duality theory is central, one needs to compute also dual bases. This sounds simple enough, however, putting it into practice can be quite intricate. We begin by collecting some relevant results and making some useful definitions.

Given a nonsingular $\overline{D} \in F^{p\times p}[z]$, we define the reachable pair $(A^{\overline{D}}, B^{\overline{D}})$ in the state space $X^{\overline{D}}$ via the shift realization (26). In Proposition 3.1, we showed that every nonsingular $\overline{T} \in F^{p\times p}[z]$ defines a controlled invariant subspace $\mathscr{V} \subset X^{\overline{D}}$ given by $\mathscr{V} = \pi^{\overline{D}} X^{\overline{T}}$. Conversely, every controlled invariant subspace has such a representation. This is true independently of any further assumptions on $\overline{T}$. Now any nonsingular polynomial matrix $\overline{T}$ with $\deg\det \overline{T} = k$ defines an observable pair $(C^{\overline{T}}, A^{\overline{T}}) \in F^{p\times k} \times F^{k\times k}$, unique up to similarity, for which $X^{\overline{T}} = \{C^{\overline{T}}(zI - A^{\overline{T}})^{-1}\xi \mid \xi \in F^k\}$. Conversely, every observable pair $(C^{\overline{T}}, A^{\overline{T}}) \in F^{p\times k} \times F^{k\times k}$, defines a nonsingular polynomial matrix $\overline{T}$, unique up to a left unimodular factor, via the coprime factorization $C^{\overline{T}}(zI - A^{\overline{T}})^{-1} = \overline{T}(z)^{-1}\Psi(z)$. Applying Proposition 3.1 once again, it follows that any condition invariant subspace of X_D has a representation of the form $\mathscr{V} = X_D \cap TF^p[z]$ with T nonsingular. Using the duality theory as described in Section 2, we have $\mathscr{V}^\perp = \pi^{\tilde{D}} X^{\tilde{T}} \subset X^{\tilde{D}}$ is condition invariant. Thus every reachable pair $(A, B) \in F^{k\times k} \times F^{k\times p}$ determines a unique condition invariant subspace of X_D. Conversely, given a condition invariant subspace of X_D, it determines a, possibly nonunique, reachable pair. This connection between conditioned invariant subspaces and reachable pairs we intend to explore in more detail. In this analysis, central tools are the partial observability and partial controllability matrices that we proceed to introduce.

To this end we assume without loss of generality that $D(z)$ is in Brunovsky form with observability indices $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$. This is equivalent to assuming that $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$. We assume that the ambient space X_D has dimension n, that is

$$\dim X_D = \deg\det D = \sum_{i=1}^{p} \mu_i = n. \tag{50}$$

We will frequently use the shorthand notation $\mu = (\mu_1, \ldots, \mu_p)$.

For the purpose of studying the set of all controlled (conditioned) invariant subspaces of fixed dimension (codimension), we would like to have a description of these subspaces also in state space terms. Indeed, this can be done through the use of two new objects—the *partial observability matrix* and the *partial reachability matrix*. A special instance of the partial observability matrix was introduced in [16]. Here we will remove all genericity restrictions on the reachability indices. It turns out that the partial observability matrix is very closely related to certain projections as well as to induced Toeplitz maps.

Duality plays an important role in this analysis. Since we have identified the dual space of a polynomial model X_T with a rational model $X^{\tilde{T}}$, it is only natural to make the identification $(F^n)^* = F^n$.

Definition 4.1. Let $\mu = (\mu_1, \ldots, \mu_p)$ with $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$.

1. Given an observable pair $(C, A) \in F^{p\times k} \times F^{k\times k}$, we define the *$\mu$-partial observability matrix* $\mathcal{O}_\mu(C, A) \in F^{n\times k}$ by

$$\mathcal{O}_\mu(C, A) = \mathcal{O}_{(\mu_1,\ldots,\mu_p)} = \mathcal{O}_\mu := \begin{pmatrix} C_1 \\ C_1 A \\ \cdot \\ \cdot \\ C_1 A^{\mu_1-1} \\ \cdot \\ \cdot \\ \cdot \\ C_p \\ \cdot \\ \cdot \\ C_p A^{\mu_p-1} \end{pmatrix} \tag{51}$$

   We can consider $\mathcal{O}_\mu$ as a map from F^k into F^n.

2. Given a reachable pair $(A, B) \in F^{k\times k} \times F^{k\times p}$, we define the *$\mu$-partial reachability matrix* $\mathcal{R}_\mu(A, B) \in F^{k\times n}$ by

$$\mathcal{R}_\mu(A, B) = (b_1, Ab_1, \ldots, A^{\mu_1-1}b_1, \ldots, b_p, \ldots, A^{\mu_p-1}b_p). \tag{52}$$

   We can consider $\mathcal{R}_\mu$ as a map from F^n into F^k.

Definition 4.2. Let $\mu = (\mu_1, \ldots, \mu_p)$ with $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$ and let $1 \leqslant k \leqslant n$.

1. A pair $(A, B) \in F^{k\times k} \times F^{k\times p}$ is called *μ-regular* if the $k \times n$ partial reachability matrix

$$\mathcal{R}_\mu(A, B) = (b_1, \ldots, A^{\mu_1-1}b_1, \ldots, b_p, \ldots, A^{\mu_p-1}b_p) \tag{53}$$

   has full rank k.

2. The pair (A, B) is called *μ-tight*, if it is $\mu - 1 = (\mu_1 - 1, \ldots, \mu_p - 1)$-regular, i.e. if

$$\operatorname{rank}\mathscr{R}_{\mu-1}(A,B)=\operatorname{rank}(b_1,\ldots,A^{\mu_1-2}b_1,\ldots,b_p,\ldots,A^{\mu_p-2}b_p)=k. \tag{54}$$

Obviously, any μ-tight pair is μ-regular and μ-regularity implies reachability. We now show that μ-tight pairs (A,B) give rise to tight subspaces of F^n.

Lemma 4.1. *The pair $(A,B)\in F^{k\times k}\times F^{k\times p}$ is μ-tight if and only if the subspace* $\operatorname{Ker}\mathscr{R}_\mu(A,B)$ *is tight, i.e. if and only if*

$$\operatorname{Ker}\mathscr{R}_\mu(A,B)+\operatorname{Ker}\mathscr{C}=F^n. \tag{55}$$

Proof. $\operatorname{Ker}\mathscr{R}_\mu(A,B)$ is tight if and only if

$$\begin{aligned}&\dim\operatorname{Ker}\mathscr{R}_\mu(A,B)+\dim\operatorname{Ker}\mathscr{C}-\dim(\operatorname{Ker}\mathscr{R}_\mu(A,B)\cap\operatorname{Ker}\mathscr{C})\\&\quad=\dim(\operatorname{Ker}\mathscr{R}_\mu(A,B)+\operatorname{Ker}\mathscr{C})=n,\end{aligned}$$

i.e. if and only if

$$\dim\operatorname{Ker}\begin{pmatrix}\mathscr{R}_\mu(A,B)\\ \mathscr{C}\end{pmatrix}=n-\operatorname{rank}\mathscr{R}_\mu(A,B)-\operatorname{rank}\mathscr{C}.$$

Equivalently, this holds if and only if

$$\operatorname{rank}\begin{pmatrix}\mathscr{R}_\mu(A,B)\\ \mathscr{C}\end{pmatrix}=\operatorname{rank}\mathscr{R}_\mu(A,B)+\operatorname{rank}\mathscr{C}.$$

Since $\mathscr{C}=\operatorname{diag}(\tilde{e}_{\mu_1},\ldots,\tilde{e}_{\mu_p})$, this is equivalent to

$$\begin{pmatrix}b_1 & \cdot & \cdot & \cdot & A^{\mu_1-1}b_1 & \cdot & \cdot & \cdot & b_p & \cdot & \cdot & \cdot & A^{\mu_1-1}b_p\\ 0 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 1\end{pmatrix}$$

be full row rank. But this just means that $\operatorname{rank}\mathscr{R}_{\mu-1}(A,B)=\operatorname{rank}\mathscr{R}_\mu(A,B)$, i.e. that (A,B) is tight. $\square$

To re-interpret partial observability/reachability matrices in the context of polynomial models, we need to introduce a few maps that connect standard coordinate spaces with polynomial or rational models. Assume we are given a nonsingular polynomial matrix in Brunovsky form, i.e. $D(z)=\operatorname{diag}(z^{\mu_1},\ldots,z^{\mu_p})$ with $\mu_1\geqslant\cdots\geqslant\mu_p\geqslant 0$ and $\sum_{i=1}^n\mu_i=n$. As in (26), we define a reachable pair $(\mathscr{A}^D,\mathscr{B}^D)$ using the shift realization (26) in the state space X^D. Using the standard basis in X^D, i.e.

$$\left\{\begin{pmatrix}1/z\\0\\\cdot\\\cdot\\0\end{pmatrix},\ldots,\begin{pmatrix}1/z^{\mu_1}\\0\\\cdot\\\cdot\\0\end{pmatrix},\ldots,\begin{pmatrix}0\\\cdot\\\cdot\\0\\1/z\end{pmatrix},\ldots,\begin{pmatrix}0\\\cdot\\\cdot\\0\\1/z^{\mu_p}\end{pmatrix}\right\},$$

we can identify this pair with its matrix representation in the above basis, i.e. with the Brunovsky pair

$$\mathscr{A}^{\mu} = \begin{pmatrix} J_{\mu_1} & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & J_{\mu_p} \end{pmatrix}_{n\times n}, \qquad \mathscr{B}^{\mu} = \begin{pmatrix} L_{\mu_1} & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & L_{\mu_p} \end{pmatrix}_{n\times p}, \tag{56}$$

where

$$J_{\mu_i} = \begin{pmatrix} 0 & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}_{\mu_i\times\mu_i}, \qquad L_{\mu_i} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix}_{\mu_i\times 1}. \tag{57}$$

Similarly, given the same polynomial matrix $D(z)$, we define an observable pair $(\mathscr{C}_D, \mathscr{A}_D)$ in the state space X_D using the shift realization (24). Using the standard basis in X_D, namely

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} z^{\mu_1-1} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ z^{\mu_p-1} \end{pmatrix} \right\}, \tag{58}$$

we can identify this pair with its matrix representation with respect to the basis defined above, i.e. with the dual Brunovsky pair defined by

$$\begin{aligned} A_{\mu} &= \tilde{A}^{\mu}, \\ C_{\mu} &= \tilde{B}^{\mu}. \end{aligned} \tag{59}$$

We expect to construct a map $J^D : F^n \longrightarrow X^D$ that will intertwine the pairs $(\mathscr{A}^{\mu}, \mathscr{B}^{\mu})$ and (A^D, B^D). To this end, we make the identification $F^n = F^{\mu_1} \times \cdots \times F^{\mu_p}$ and write, for $\eta \in F^n$,

$$\eta^T = (\eta_{11}, \ldots, \eta_{1\mu_1}, \eta_{21}, \ldots, \eta_{2\mu_2}, \ldots, \eta_{p1}, \ldots, \eta_{p\mu_p}).$$

With this notation, we define

$$\begin{aligned} &J^D : F^n \longrightarrow X^D, \\ &J^D\eta = \begin{pmatrix} \frac{\eta_{11}}{z} + \cdots + \frac{\eta_{1\mu_1}}{z^{\mu_1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\eta_{p1}}{z} + \cdots + \frac{\eta_{p\mu_p}}{z^{\mu_p}} \end{pmatrix} \end{aligned} \tag{60}$$

Similarly, given an observable pair (C, A) let $T(z)^{-1}H(z)$ be a left coprime facorization of $C(zI - A)^{-1}$. Altogether, we have four spaces F^k, F^n, X^T and X^D and a variety of maps connecting these spaces and relating to the various pairs. These we proceed to study and to facilitate this we shall need to choose a convenient basis in each one of these spaces.

In F^k we choose the *standard basis*, i.e.

$$\mathscr{B}_{\mathrm{st}} = \{e_1, \ldots, e_k\}. \tag{61}$$

In $F^n = F^{\mu_1} \times \cdots \times F^{\mu_p}$ we choose the *standard basis* but indexed differently, i.e.

$$\mathscr{B}_{\mathrm{ST}} = \{f_{11}, \ldots, f_{1\mu_1}, \ldots, f_{p1}, \ldots, f_{1\mu_p}\}. \tag{62}$$

In X^T we choose the *observer basis*, i.e.

$$\mathscr{B}_{\mathrm{ob}} = \{C(zI - A)^{-1}e_1, \ldots, C(zI - A)^{-1}e_k\}. \tag{63}$$

Finally, in X^D we choose the *Brunovsky basis*, i.e.

$$\mathscr{B}_{\mathrm{BR}} = \left\{ \begin{pmatrix} 1/z \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 1/z^{\mu_1} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1/z \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1/z^{\mu_p} \end{pmatrix} \right\}. \tag{64}$$

Clearly, $J^D\eta = f$ implies $[f]^{\mathrm{BR}} = \eta$, i.e. η is the vector of coordinates of $J^D\eta$ with respect to the Brunovsky basis (64). In particular $[J^D\eta]^{\mathrm{BR}} = \eta$, i.e. $(J^D)^{-1} = [\cdot]^{\mathrm{BR}}$. In an analogous way, given $\xi \in F^k$, we have $[J^T\xi]^{\mathrm{ob}} = \xi$, i.e. $(J^T)^{-1} = [\cdot]^{\mathrm{ob}}$.

We will study the Toeplitz operator with symbol TD^{-1} and relate the nonpositivity of the factorization indices of TD^{-1} with properties of the partial observability matrix. With these definitions, we can state:

Proposition 4.1. *Let $D(z) = \mathrm{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$ and $\mu_1 + \cdots + \mu_p = n$. Let $T(z)$ be a $p \times p$ nonsingular polynomial matrix. Assume T is row proper with row indices $\nu_1, \ldots, \nu_p$ satisfying*

$$\nu_1 + \cdots + \nu_p = \deg\det T = k. \tag{65}$$

Without loss of generality assume that the row highest coefficient matrix of T is the identity. Making the identification $F^k = F^{\nu_1} \times \cdots \times F^{\nu_p}$ we enumerate the standard basis of F^k in the following way:

$$\mathscr{B} = \{e_{11}, \ldots, e_{1\nu_1}, \ldots, e_{p1}, \ldots, e_{p\nu_p}\}, \tag{66}$$

where

$$e_{ij} = \begin{pmatrix} \delta_{i1}\delta_{j1} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{i1}\delta_{j\nu_1} \\ \cdot \\ \cdot \\ \delta_{ip}\delta_{j1} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{ip}\delta_{j\nu_p} \end{pmatrix}.$$

The standard basis in F^p will be denoted by $\{e_1, \ldots, e_p\}$. Then:

1. *The following set of vectors*

$$\{f_{kl}(z) = z^{l-1}e_k \mid k = 1, \ldots, p,\ l = 1, \ldots, \nu_k\} \tag{67}$$

is a basis of X_T with the basis matrix given by

$$\begin{pmatrix} 1 & z & \cdot & \cdot & z^{\nu_1-1} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & z & \cdot & \cdot & z^{\nu_p-1} \end{pmatrix}. \tag{68}$$

We call this the standard basis of X_T.

2. *Given the standard basis $\{e_1, \ldots, e_p\}$ of F^p, we define k row vector polynomials by*

$$\Phi_{ij} = \pi_+ z^{-j}\tilde{e}_i T(z), \quad i = 1, \ldots, p \ \text{ and } \ j = 1, \ldots, \nu_i. \tag{69}$$

The set of vectors $\{\tilde{T}^{-1}\tilde{\Phi}_{11}, \ldots, \tilde{T}^{-1}\tilde{\Phi}_{p\nu_p}\}$ is a basis of $X^{\tilde{T}}$ and it is the dual basis to the standard basis of X_T defined in (68). *We refer to* (69) *as the control basis of $X^{\tilde{T}}$.*

3. *Let the $k \times p$ polynomial matrix be defined by*

$$\Phi = \begin{pmatrix} \Phi_{11} \\ \cdot \\ \cdot \\ \cdot \\ \Phi_{1\nu_1} \\ \cdot \\ \cdot \\ \cdot \\ \Phi_{p1} \\ \cdot \\ \cdot \\ \cdot \\ \Phi_{p\nu_p} \end{pmatrix}. \tag{70}$$

Then we have

$$(zI - A)^{-1}B = \Phi(z)T(z)^{-1}, \tag{71}$$

where

$$A = [S_T]^{\mathrm{st}}_{\mathrm{st}}, \quad B = [\pi_T \cdot]^{\mathrm{st}}_{\mathrm{st}}. \tag{72}$$

4. *With* (A, B) *defined by* (72), *the* μ*-partial reachability matrix* $R_\mu(A, B)$ *is given by*

$$\begin{aligned} &R_\mu(A, B) \\ &= \left[\pi_T \begin{pmatrix} 1 & z & \cdot & \cdot & z^{\mu_1 - 1} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & z & \cdot & \cdot & z^{\mu_p - 1} \end{pmatrix} \right]^{\mathrm{st}}_{\mathrm{st}} \\ &= [\pi_T \mid X_D]^{\mathrm{st}}_{\mathrm{st}}. \end{aligned} \tag{73}$$

Equivalently, the following diagram is commutative.

$$\begin{tikzcd} X_D \arrow[r, "\pi_T|X_D"] \arrow[d, "{[\cdot]^{st}}"'] & X_T \arrow[d, "{[\cdot]^{st}}"] \\ F^n \arrow[r, "R_\mu(A,B)"] & F^k \end{tikzcd}$$

5. *The dual diagram is given by*

$$\begin{tikzcd} X^{\tilde{T}} \arrow[r, "\pi^{\tilde{D}}|X^{\tilde{T}}"] & X^{\tilde{D}} \\ F^k \arrow[u, "J^{\tilde{T}}"] \arrow[r, "\mathcal{O}_\mu(\tilde{B},\tilde{A})"] & F^n \arrow[u, "J^{\tilde{D}}"'] \end{tikzcd}$$

The inverse of the map $J^{\tilde{T}}$ *is given by*

$$(J^{\tilde{T}})^{-1}h = [h]^{\mathrm{ob}} \quad \textit{for } h \in X^{\tilde{T}}, \tag{74}$$

and so the commutativity of the previous diagram is equivalent to

$$\mathcal{O}_\mu(\tilde{B}, \tilde{A}) = [\pi^{\tilde{D}} \mid X^{\tilde{T}}]^{\mathrm{ob}}_{\mathrm{ob}}. \tag{75}$$

Proof. (1) This follows from Proposition 2.1, noting that T is row proper with row degrees $\nu_1, \ldots, \nu_p$.

(2) First we note that the vectors $\tilde{\Phi}_{ij}$ are elements of $X_{\tilde{T}}$. We cannot apply Proposition 2.1 as $\tilde{T}$ is column proper but not row proper. We proceed as in [10]. Note that

$$\tilde{\Phi}_{ij}(z) = \pi_+ z^{-j}\tilde{T}(z)e_i = \left[\frac{\tilde{T}(z) - \sum_{k=0}^{j-1}\tilde{T}_k z^k}{z^j}\right]e_i.$$

Therefore

$$\tilde{T}^{-1}\tilde{\Phi}_{ij}(z) = \frac{e_i}{z^j} - \tilde{T}^{-1}\frac{\sum_{k=0}^{j-1}T_{k-1}z^k}{z^j}e_i.$$

Now, since $\tilde{T}$ is column proper, its inverse is proper and this shows that $\tilde{\Phi}_{ij} \in X_{\tilde{T}}$.

To show the duality of the two bases, we compute, using the computational rules proved in [11,12]

$$\begin{aligned}[\tilde{T}^{-1}\tilde{\Phi}_{ij}, f_{kl}] &= [\tilde{T}^{-1}\pi_+ z^{-j}\tilde{T}e_i, z^{l-1}e_k] = [\pi_-\tilde{T}^{-1}\pi_+\tilde{T}z^{-j}e_i, z^{l-1}e_k] \\ &= [\pi^{\tilde{T}}z^{-j}e_i, z^{l-1}e_k] = [z^{-j}e_i, \pi_T z^{l-1}e_k] = [z^{-j}e_i, z^{l-1}e_k] \\ &= \tilde{e}_i e_k \cdot (1, z^{l-j-1})_{-1} = \delta_{ik}\delta_{jl}.\end{aligned}$$

(3) Follows by taking the matrix representation, with respect to the standard basis of X_T, of the shift realization associated to $\Phi(z)T(z)^{-1}$. Note that, by part 2, the output map $f \mapsto (\Phi T^{-1}f)_{-1}$ has the identity as matrix representation.

(4) We compute

$$\pi_T z^{l-1}e_k = \pi_T z^{l-1}\pi_T e_k = S_T^{l-1}\pi_T e_k.$$

So, taking matrix representations,

$$[\pi_T z^{l-1}e_k] = [S_T^{l-1}][\pi_T]e_k = ([S_T])^{l-1}[\pi_T]e_k = A^{l-1}Be_k = A^{l-1}B_k.$$

(5) The commutativity of the diagram follows from the commutativity of the previous diagram by duality considerations. It may be of interest to give a direct proof. To this end we compute, for $\xi \in F^k$:

$$\begin{aligned}\pi^{\tilde{D}}J^{\tilde{T}}\xi &= \pi^{\tilde{D}}\tilde{B}(zI - \tilde{A})^{-1}\xi \\ &= \pi^{\tilde{D}}\begin{pmatrix}\tilde{B}_1(zI - \tilde{A})^{-1}\xi \\ \cdot \\ \cdot \\ \cdot \\ \tilde{B}_p(zI - \tilde{A})^{-1}\xi\end{pmatrix} = \begin{pmatrix}\frac{\tilde{B}_1}{z} + \cdots + \frac{\tilde{B}_1\tilde{A}^{\mu_1 - 1}}{z^{\mu_1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\tilde{B}_p}{z} + \cdots + \frac{\tilde{B}_p\tilde{A}^{\mu_p - 1}}{z^{\mu_p}}\end{pmatrix} \\ &= J^{\tilde{D}}(\tilde{B}_1\xi, \ldots, \tilde{B}_1\tilde{A}^{\mu_1 - 1}\xi, \ldots, \tilde{B}_p\xi, \ldots, \tilde{B}_p\tilde{A}^{\mu_p - 1}\xi) \\ &= J^{\tilde{D}}\mathcal{O}_\mu(\tilde{B}, \tilde{A})\xi\end{aligned}$$

We have, for $\xi \in F^k$, $\xi = \sum_{i=1}^{p} \sum_{j=1}^{\nu_i} \xi_{ij} e_{ij}$. For the map $J^{\tilde{T}} : F^k \longrightarrow X^{\tilde{T}}$, we have

$$\begin{aligned} J^{\tilde{T}}\xi &= \tilde{T}(z)^{-1}\tilde{\Phi}(z)\xi = \tilde{T}(z)^{-1}\tilde{\Phi}(z)\sum_{i=1}^{p}\sum_{j=1}^{\nu_i}\xi_{ij}e_{ij} \\ &= \sum_{i=1}^{p}\sum_{j=1}^{\nu_i}\xi_{ij}\tilde{T}(z)^{-1}\tilde{\Phi}_{ij}(z) \end{aligned}$$

Therefore we have

$$(J^{\tilde{T}})^{-1}h = [h]^{\mathrm{ob}} \quad \text{for } h \in X^{\tilde{T}}. \tag{76}$$

Equality (75) follows by duality from (73) as we take matrix representations of dual maps with respect to dual bases. □

The vector polynomials introduced in (69) are generalizations of the control basis in the scalar case, see [11,12]. In the scalar case they have been attributed by Kalman to Tschirnhausen.

Theorem 4.1. *Let $D(z) = \mathrm{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$ and $\mu_1 + \cdots + \mu_p = n$. Assume $(C, A) \in F^{p\times k} \times F^{k\times k}$ is an observable pair. Let $T(z)^{-1}H(z)$ be a left coprime factorization of $C(zI - A)^{-1}$. Then*:

1. *Define the map $J^T : F^k \longrightarrow X^T$ defined by*

$$\begin{aligned} &J^T : F^k \longrightarrow X^T, \\ &\xi \mapsto C(zI - A)^{-1}\xi. \end{aligned} \tag{77}$$

Let (C^T, A^T) be defined by the shift realization associated with T, i.e. for $h \in X^T$:

$$\begin{aligned} &A^T h = S^T = \pi_- zh, \\ &C^T h = h_{-1}. \end{aligned} \tag{78}$$

Then J^T is an isomorphism that intertwines the pairs (C, A) and (C^T, A^T), i.e. we have

$$\begin{aligned} &J^T A = A^T J^T, \\ &C = C^T J^T. \end{aligned} \tag{79}$$

2. *The map $J^D : F^n \longrightarrow X^D$ defined by* (60) *intertwines the pairs $(\mathscr{A}^\mu, \mathscr{B}^\mu)$ and (A^D, B^D), i.e. we have*

$$\begin{aligned} &J^D\mathscr{B}^\mu = \mathscr{B}^D, \\ &J^D\mathscr{A}^\mu = \mathscr{A}^D J^D. \end{aligned} \tag{80}$$

3. *Let $\mathscr{O}_\mu(C, A)$ be defined by* (51). *Then the following is a commutative diagram*:

$$\begin{tikzcd} F^k \arrow[r, "J^T"] \arrow[d, "\mathcal{O}_\mu(C,A)"'] & X^T \arrow[d, "\phi = \pi^D|X^T"] \\ F^n \arrow[r, "J^D"] & X^D \end{tikzcd}$$

i.e. we have

$$J^D \mathcal{O}_\mu(C, A) = (\pi^D \mid X^T) J^T. \tag{81}$$

4. *The dual diagram is given by*

$$\begin{tikzcd} F^k & X_T \arrow[l, "(J^T)^*"'] \\ F^n \arrow[u, "\mathcal{O}_\mu(C,A)^*"] & X_D \arrow[l, "(J^D)^*"'] \arrow[u, "\phi^* = \pi_T|X_D"'] \end{tikzcd}$$

5. *The partial observability matrix $\mathcal{O}_\mu(C, A)$ is the matrix representation of $\pi^D \mid X^T$ with respect to the bases $\mathscr{B}_{ob}$ in X^T and $\mathscr{B}_{BR}$ in X^D, i.e.*

$$\mathcal{O}_\mu(C, A) = [\pi^D \mid X^T]_{\mathrm{ob}}^{\mathrm{BR}}. \tag{82}$$

6. *The partial observability matrix $\mathcal{O}_\mu(C, A)$ has full column rank if and only if $\pi^D \mid X^T$ is injective.*
7. *There exists a feedback $\mathscr{K}^\mu$ for which*

$$\mathcal{O}_\mu(C, A)A = \mathscr{A}\mathcal{O}_\mu(C, A) = (\mathscr{A}^\mu + \mathscr{B}^\mu \mathscr{K}^\mu)\mathcal{O}_\mu. \tag{83}$$

8. *A subspace $\mathscr{W} \subset X^D$ is a controlled invariant subspace for the reachable pair $(\mathscr{A}^D, \mathscr{B}^D) = (S^D, \pi_- D^{-1}\cdot)$ if and only if*

$$\mathscr{W} = \pi^D(X^T) \tag{84}$$

for some nonsingular polynomial matrix T.
9. *A subspace $\mathbf{W} \subset F^n$ is a controlled invariant subspace for the Brunovsky pair $(\mathscr{A}^\mu, \mathscr{B}^\mu)$ if and only if*

$$\mathbf{W} = \operatorname{Im} \mathcal{O}_\mu(C, A) \tag{85}$$

for some observable pair $(C, A) \in F^{p\times k} \times F^{k\times k}$.

10. *A subspace* $\mathscr{V} \subset X_D$ *is a conditioned invariant subspace for the observable pair* $(\mathscr{A}_D, \mathscr{C}_D) = (S_D, (D^{-1}\cdot)_{-1})$ *if and only if*

$$\mathscr{V} = \operatorname{Ker} \pi_T \mid X_D = X_D \cap TF^p[z] \tag{86}$$

for some nonsingular polynomial matrix T.

11. *A subspace* $\mathbf{V} \subset F^n$ *is a conditioned invariant subspace for the Brunovsky pair* $(\mathscr{C}_\mu, \mathscr{A}_\mu)$ *if and only if*

$$\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B) \tag{87}$$

for some reachable pair $(A, B) \in F^{k\times k} \times F^{k\times p}$. *We refer to* (86) *and* (87) *as the kernel representations of* $\mathscr{V}$ *and* $\mathbf{V}$, *respectively.*

12. *A subspace* $\mathbf{V} \subset F^n$ *is a codimension k conditioned invariant subspace for the Brunovsky pair* $(\mathscr{C}_\mu, \mathscr{A}_\mu)$ *if and only if*

$$\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B) \tag{88}$$

for some μ*-regular pair* $(A, B) \in F^{k\times k} \times F^{k\times p}$.

13. *A subspace* $\mathbf{V} \subset F^n$ *is a tight conditioned invariant subspace of codimension k for the Brunovsky pair* $(\mathscr{C}_\mu, \mathscr{A}_\mu)$ *if and only if*

$$\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B) \tag{89}$$

for some μ*-tight pair* $(A, B) \in F^{k\times k} \times F^{k\times p}$.

Proof.

1. We compute, for $\xi \in F^k$,

$$\begin{aligned} A^T J^T \xi &= \pi_- z C(zI - A)^{-1}\xi = C(zI - A)^{-1}A\xi = J^T A\xi \\ C^T J^T \xi &= (C(zI - A)^{-1}\xi)_{-1} = C\xi. \end{aligned}$$

2. Let $\xi \in F^p$, with $\xi^T = (\xi_1, \ldots, \xi_p)$. Then

$$\mathscr{B}^\mu \xi = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \xi_1 \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline 0 \\ \cdot \\ \cdot \\ \xi_p \end{pmatrix}$$

and hence

$$\mathscr{B}^{\mu}\xi = \begin{pmatrix} \frac{\xi_1}{z^{\mu_1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\xi_p}{z^{\mu_p}} \end{pmatrix}.$$

On the other hand,

$$\pi_- D^{-1} \begin{pmatrix} \xi_1 \\ \cdot \\ \cdot \\ \cdot \\ \xi_p \end{pmatrix} = \begin{pmatrix} \frac{\xi_1}{z^{\mu_1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\xi_p}{z^{\mu_p}} \end{pmatrix}.$$

This proves the first equality. For the second equality, we compute

$$J^D \mathscr{A}^{\mu} \begin{pmatrix} \xi_{11} \\ \cdot \\ \cdot \\ \xi_{1\mu_1} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{p1} \\ \cdot \\ \cdot \\ \xi_{p\mu_p} \end{pmatrix} = J^D \mathscr{A}^{\mu} \begin{pmatrix} \xi_{12} \\ \cdot \\ \cdot \\ \xi_{1\mu_1} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \xi_{p2} \\ \cdot \\ \cdot \\ \xi_{p\mu_p} \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\xi_{12}}{z} + \cdots + \frac{\xi_{1\mu_1}}{z^{\mu_1 - 1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\xi_{p2}}{z} + \cdots + \frac{\xi_{p\mu_p}}{z^{\mu_p - 1}} \end{pmatrix} = S^D J^D \begin{pmatrix} \xi_{11} \\ \cdot \\ \cdot \\ \xi_{1\mu_1} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{p1} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{p\mu_p} \end{pmatrix}.$$

3. We compute, for $\xi \in F^k$,

$$\pi^D J^T \xi = \pi^D C(zI - A)^{-1}\xi = \pi^D \begin{pmatrix} \sum_{i=1}^{\infty} \frac{C_1 A^{i-1}\xi}{z^i} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=1}^{\infty} \frac{C_p A^{i-1}\xi}{z^i} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{\mu_1} \frac{C_1 A^{i-1}\xi}{z^i} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=1}^{\mu_p} \frac{C_p A^{i-1}\xi}{z^i} \end{pmatrix} = J_D \begin{pmatrix} C_1\xi \\ \cdot \\ \cdot \\ \cdot \\ C_1 A^{\mu_1-1}\xi \\ \cdot \\ \cdot \\ \cdot \\ C_p\xi \\ \cdot \\ \cdot \\ \cdot \\ C_p A^{\mu_p-1}\xi \end{pmatrix} = J_D \mathcal{O}_\mu(C, A)\xi.$$

4. Use our identification of dual spaces.
5. By (81) we have $J^D \mathcal{O}_\mu = \pi^D J^T$ and so

   $$[J^D]_{\mathrm{br}}^{\mathrm{BR}} [\mathcal{O}_\mu]_{\mathrm{st}}^{\mathrm{br}} = [\pi^D \mid X^T]_{\mathrm{ob}}^{\mathrm{BR}} [J^T]_{\mathrm{st}}^{\mathrm{ob}}.$$

   But $[J^D]_{\mathrm{br}}^{\mathrm{BR}} = I_n$, $[J^T]_{\mathrm{st}}^{\mathrm{ob}} = I_k$ and so $[\mathcal{O}_\mu]_{\mathrm{st}}^{\mathrm{br}} = \mathcal{O}_\mu$, which proves the claim.
6. Follows from the previous part.
7. We compute, for $\xi \in F^k$,

$$\mathcal{O}_\mu(C, A)A\xi = \begin{pmatrix} C_1 A\xi \\ \cdot \\ \cdot \\ \cdot \\ C_1 A^{\mu_1}\xi \\ \cdot \\ \cdot \\ \cdot \\ C_p A\xi \\ \cdot \\ \cdot \\ C_p A^{\mu_p}\xi \end{pmatrix}$$

$$= \left(\begin{array}{ccccc|ccc|ccccc}
0 & 1 & & & & & & & & & & & \\
 & \cdot & \cdot & & & & & & & & & & \\
 & & \cdot & \cdot & & & & & & & & & \\
 & & & \cdot & 1 & & & & & & & & \\
 & * & * & * & * & * & * & * & * & * & * & * & * \\
\hline
 & & & & & \cdot & & & & & & & \\
 & & & & & & \cdot & & & & & & \\
 & & & & & & & \cdot & & & & & \\
\hline
 & & & & & & & & 0 & 1 & & & \\
 & & & & & & & & & \cdot & \cdot & & \\
 & & & & & & & & & & \cdot & \cdot & \\
 & & & & & & & & & & & \cdot & 1 \\
 & * & * & * & * & * & * & * & * & * & * & * & *
\end{array}\right)$$

$$\times \begin{pmatrix} C_1\xi \\ \cdot \\ \cdot \\ \cdot \\ C_1A^{\mu_1-1}\xi \\ \cdot \\ \cdot \\ \cdot \\ C_p\xi \\ \cdot \\ \cdot \\ \cdot \\ C_pA^{\mu_p-1}\xi \end{pmatrix} = \mathscr{A}\mathscr{O}_\mu\xi = (\mathscr{A}^\mu + \mathscr{B}^\mu\mathscr{K}^\mu)\mathscr{O}_\mu\xi.$$

8. Follows from Theorem 3.1.
9. From the equality $\mathscr{O}_\mu A = (\mathscr{A}^\mu + \mathscr{B}^\mu\mathscr{K}^\mu)\mathscr{O}_\mu$ it follows that $\mathscr{W} = \operatorname{Im}\mathscr{O}_\mu(C, A)$ is indeed a controlled invariant subspace for the Brunovsky pair $(\mathscr{A}^\mu, \mathscr{B}^\mu)$. Conversely, if $\mathscr{W} \subset F^n$ is a controlled invariant subspace for the Brunovsky pair $(\mathscr{A}^\mu, \mathscr{B}^\mu)$, then from (80) it follows that $J^D\mathscr{W}$ is controlled invariant for $(\mathscr{A}^D, \mathscr{B}^D)$. By Theorem 3.1, there exists a nonsingular polynomial matrix T for which $\pi^D \mid X^T$ is injective and $J^D\mathscr{W} = \pi^D X^T$. Now, by (27),

   $$X^T = \{C(zI - A)^{-1}\xi \mid \xi \in F^k\}$$

   for some observable pair (C, A) which is uniquely determined up to a state isomorphism. Clearly, $\operatorname{Im}\mathscr{O}_\mu(C, A)$ is controlled invariant and, since $J^D\mathscr{O}_\mu = \pi^D J^T$ and J^D is invertible, we conclude $\mathscr{W} = \operatorname{Im}\mathscr{O}_\mu(C, A)$.
10. Follows from the fact that $\operatorname{Ker}\pi_T = T(z)F^p[z]$.
11. $\mathbf{V} \subset F^n$ is a conditioned invariant subspace for the Brunovsky pair $(\mathscr{C}_\mu, \mathscr{A}_\mu)$ if and only if $\mathbf{V} = J_D\mathscr{V}$, where $\mathscr{V} \subset X_D$ is a conditioned invariant subspace for $(\mathscr{C}_D, \mathscr{A}_D)$, i.e. if and only if $\mathscr{V} = X_D \cap T(z)F^p[z]$ for some nonsingular $T \in F^{p\times p}[z]$. Now it is easily computed that, with the duality pairing introduced in Section 2, we have

    $$(X_D \cap T(z)F^p[z])^\perp = \pi^{\tilde{D}}X^{\tilde{T}} = \operatorname{Im}(\pi^{\tilde{D}} \mid X^{\tilde{T}}).$$

    Applying (82), with $\tilde{T}$ and $\tilde{D} = D$, we have

    $$\mathbf{V} = (\operatorname{Im}\mathscr{O}_\mu(\tilde{B}, \tilde{A}))^\perp = \operatorname{Ker}\mathscr{O}_\mu(\tilde{B}, \tilde{A})^* = \operatorname{Ker}\mathscr{R}_\mu(A, B).$$

12. $\operatorname{Ker}\mathscr{R}_\mu(A, B)$ is conditioned invariant if and only if

    $$\operatorname{Ker}\mathscr{R}_\mu(A, B)^\perp = \operatorname{Im}\mathscr{O}_\mu(B^T, A^T)$$

    is $(\mathscr{A}^T, \mathscr{C}^T)$-invariant. Here $\mathscr{O}_\mu(H, F)$ denotes the analogously defined partial observability matrix. For

$$\mathscr{A}^T = \left(\begin{array}{cccc|c|cccc}
0 & 1 & & & & & & & \\
 & \ddots & \ddots & & & & & & \\
 & & \ddots & 1 & & & & & \\
 & & & 0 & & & & & \\
\hline
 & & & & \ddots & & & & \\
\hline
 & & & & & 0 & 1 & & \\
 & & & & & & \ddots & \ddots & \\
 & & & & & & & \ddots & 1 \\
 & & & & & & & & 0
\end{array}\right),$$

$$\mathscr{C}^T = \left(\begin{array}{c|c|c}
0 & & \\
\vdots & & \\
1 & & \\
\hline
 & \ddots & \\
\hline
 & & 0 \\
 & & \vdots \\
 & & 1
\end{array}\right).$$

and $X \in F^{n\times k}$ partitioned as

$$X = \begin{pmatrix} X_1^0 \\ \vdots \\ X_1^{\mu_1-1} \\ \vdots \\ X_p^0 \\ \vdots \\ X_p^{\mu_p-1} \end{pmatrix},$$

it follows that span X is $(\mathscr{A}^T, \mathscr{C}^T)$-invariant if and only if there exists $T \in F^{k\times k}$ with

$$X_j^1 = X_j^0 T, \ldots, X_j^{\mu_j-1} = X_j^{\mu_j-2} T, \quad j = 1, \ldots, p.$$

Equivalently, for $B := (X_1^{0T}, \ldots, X_p^{0T})$, $A^T := T$ we obtain that

$$X = \mathscr{O}_\mu(B^T, A^T).$$

13. By Lemma 4.1 the subspace $\operatorname{Ker} \mathscr{R}_\mu(A, B)$ is tight if and only if (A, B) is $(\mu - 1)$-regular, i.e. tight. Suppose (A_1, B_1), (A_2, B_2) are tight pairs with

$$\operatorname{Ker} \mathscr{R}_\mu(A_1, B_1) = \operatorname{Ker} \mathscr{R}_\mu(A_2, B_2).$$

Then, up to a similarity transformation on (A_i, B_i), we can assume that

$$\mathscr{R}_\mu(A_1, B_1) = \mathscr{R}_\mu(A_2, B_2).$$

Since $\mu_p \geqslant 1$ we have $B_1 = B_2$. Moreover, it is easily seen that $A_2 = A_1 + X$ for an arbitrary $X \in F^{k\times k}$ with

$$XR_{\mu-1}(A_1, B_1) = 0.$$

Since (A_1, B_1) is tight we conclude that $X = 0$ and hence $A_1 = A_2$. This completes the proof. □

We note that the subspace **W** defined in (85) is equally defined on the similarity class of the observable pair (C, A). Similarly, the subspace **V** defined in (87) is equally defined on the similarity class of the reachable pair (A, B). In the same vein, given a unimodular polynomial matrix U, we have $X^{UT} = X^T$ and hence $\mathscr{W} = \pi^D(X^T) = \pi^D(X^{UT})$. Analogously, we have $\mathscr{V} = \operatorname{Ker}\pi_T \mid X_D = \operatorname{Ker}\pi_{TU} \mid X_D$.

Corollary 3.2 can be interpreted from the state space point of view.

Corollary 4.1. *Let the polynomial matrices D and T be defined as in Theorem* 4.1. *A sufficient condition for the injectivity of $\pi^D \mid X^T$ is $\delta = \deg\det T \leqslant \mu_p$.*

Proof. Note that the rows of the observability matrix

$$\begin{pmatrix} C_1 \\ C_1A \\ \cdot \\ \cdot \\ C_1A^{\delta-1} \\ \cdot \\ \cdot \\ \cdot \\ C_p \\ \cdot \\ \cdot \\ C_pA^{\delta-1} \end{pmatrix},$$

which has full column rank by the assumption that the pair (C, A) is observable, are included in the partial observability matrix $\mathscr{O}_\mu(C, A)$ defined in (51). Thus $\mathscr{O}_\mu(C, A)$ has full column rank and this proves the injectivity of $\pi^D \mid X^T$. □

For a subspace V of a linear space X we define the *cokernel* of V by

$$\operatorname{coker} V = X/V.$$

Clearly, we have

$$\dim X/V = \dim(X/V)^* = \dim V^\perp,$$

and hence

$$\operatorname{codim} V = \dim V^{\perp} = \dim \operatorname{coker} V.$$

Similarly, if $\phi : X \longrightarrow Y$ is a linear map we define

$$\operatorname{coker} \phi = Y/\operatorname{Im} \phi.$$

Since $(X/V)^*$ can be identified with $V^{\perp} \subset X^*$, we have

$$(\operatorname{coker} \phi)^* \simeq (Y/\operatorname{Im} \phi)^* = (\operatorname{Im} \phi)^{\perp} = \operatorname{Ker} \phi^*.$$

Next, we consider the map $\phi : X^T \longrightarrow X^D$ defined by $\phi = \pi^D \mid X^T$. As the duals of X^D, X^T are X_D, X_T, respectively, it follows that $\phi^* : X_D \longrightarrow X_T$ is given by

$$\phi^* = \pi_T \mid X_D. \tag{90}$$

Indeed, if $g \in X_D$ and $h \in X^T$, we have

$$\begin{aligned}[g, \phi h] &= [g, \pi^D h] = [\pi_D g, h] = [g, h] \\ &= [g, \pi^T h] = [\pi_T g, h] = [\phi^* g, h].\end{aligned}$$

As a consequence, we obtain

$$\operatorname{Ker} \phi^* = \operatorname{Ker} \pi_T \mid X_D = X_D \cap F^p[z]T(z). \tag{91}$$

In the sequel we will need the following.

Proposition 4.2.

1. *Let $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ with $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$ and $\mu_1 + \cdots + \mu_p = n$. Let $(\mathscr{A}_D, \mathscr{C}_D)$ be the coobservable pair associated with D. Then there exists a bijection between conditioned invariant subspaces of X_D of codimension k and (duals) of cokernels of the projection map π^D restricted to submodules X^T with* $\deg \det T = k$ *and TD^{-1} having nonpositive right factorization indices.*
2. *The map $(C, A) \mapsto (\operatorname{Im} \mathscr{O}_\mu(C, A))^{\perp}$ defines a bijection between codimension k, conditioned invariant subspaces of the Brunovsky pair $(\mathscr{A}_\mu, \mathscr{C}_\mu)$ and (duals) of cokernels of full column rank partial observability matrices of similarity classes of observable pairs $(C, A) \in F^{p\times k} \times F^{k\times k}$.*
3. *The map $(A, B) \mapsto \operatorname{Ker} \mathscr{R}_\mu(A, B)$ defines a bijection between similarity orbits of μ-tight pairs $(A, B) \in F^{k\times k} \times F^{k\times p}$ and codimension k, conditioned invariant subspaces of $(\mathscr{A}_\mu, \mathscr{C}_\mu)$.*

Proof.

1. By Proposition 3.1, a subspace $V \subset X_D$ is a conditioned invariant subspace for $(\mathscr{A}_D, \mathscr{C}_D)$ if and only if it has a representation of the form $V = X_D \cap F^p[z]T(z)$. For coker V we have
$$(\operatorname{coker} V)^* = (X_D/X_D \cap F^p[z]T(z))^* \simeq (X_D \cap F^p[z]T(z))^{\perp} = \pi^D(X^T).$$
2. $\mathscr{V}$ is conditioned invariant for $(\mathscr{A}_\mu, \mathscr{B}_\mu)$ if and only if $\mathscr{V}^{\perp}$ is controlled invariant for $(\mathscr{A}^\mu, \mathscr{B}^\mu)$, i.e. if and only if $\mathscr{V}^{\perp} = \operatorname{Im} \mathscr{O}_\mu(C, A)$, or equivalently if and only if $\mathscr{V} = (\operatorname{Im} \mathscr{O}_\mu(C, A))^{\perp}$.

3. By Theorem 4.1, part 13 any tight conditioned invariant subspace $\mathbf{V} \subset F^n$ of codimension k has the representation $\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B)$ for some μ-tight pair (A, B). By μ-tightness the pair (A, B) is uniquely determined from $\mathscr{V}$ up to similarity.

Conversely, any μ-tight pair (A, B) generates a tight conditioned invariant subspace $\operatorname{Ker} \mathscr{R}_\mu(A, B)$ of codimension k. The result follows. □

5. On observers and the partial realization problem

The characterization of regular and tight condition invariant subspaces is closely related to partial realizations and this in turn to observer theory. This section is devoted to the elucidation of these connections. We begin by defining preobservers with observers in mind. The theory of preobservers can be developed over an arbitrary field, whereas observers need asymptotic analysis and hence we use the real or complex field. To prepare the ground for this analysis we introduce some notations.

We assume, without loss of generality, that the pair $(\mathscr{C}, \mathscr{A})$ is given in Brunovsky form with observability indices $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$. The fact that all observability indices are positive is equivalent to our standard assumption that the rows of $\mathscr{C}$ are linearly independent. In particular, we have the coprime factorizations

$$\mathscr{C}(zI - \mathscr{A})^{-1} = D(z)^{-1}\Theta(z),$$

with

$$D(z) = \begin{pmatrix} z^{\mu_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z^{\mu_p} \end{pmatrix}, \quad \mu_1 \geqslant \cdots \geqslant \mu_p > 0.$$

This implies $X_D = X_{z^{\mu_1}} \oplus \cdots \oplus X_{z^{\mu_p}}$ with $\dim X_D = \sum_{i=1}^p \mu_i = n$. An element $f \in X_D$ can be written as

$$f(z) = \begin{pmatrix} f_1(z) \\ \cdot \\ \cdot \\ \cdot \\ f_p(z) \end{pmatrix},$$

with $f_j(z) = \sum_{\nu=1}^{\mu_j} f_{j\nu} z^{\nu-1} \in X_{z^{\mu_j}}$ polynomials of degree $\leqslant \mu_j - 1$.

A map $K : X_{z^{\mu_1}} \oplus \cdots \oplus X_{z^{\mu_p}} \longrightarrow F^l$ can be written as $K = \begin{pmatrix} K_1 & \cdot & \cdot & \cdot & K_p \end{pmatrix}$, with $K_j : X_{z^{\mu_j}} \longrightarrow F^l$. We can write

$$K_j = \begin{pmatrix} K_{1j} \\ \cdot \\ \cdot \\ \cdot \\ K_{lj} \end{pmatrix},$$

where $K_{ij} \in (X_{z^{\mu_j}})^* = X^{z^{\mu_j}}$. Thus K_{ij} has a representation of the form

$$K_{ij} = \sum_{\nu=1}^{\mu_j} \frac{K_{ij}^{(\nu)}}{z^\nu},$$

and hence

$$K_{ij} f_j = \sum_{\nu=1}^{\mu_j} K_{ij}^{(\nu)} f_{j\nu}.$$

So K_j defines an $l \times \mu_j$ matrix whose elements are $K_{ij}^{(\nu)}$. Clearly, for $\xi \in F^l$, we have

$$[K^*\xi, f] = \sum_{j=1}^{p} [K_j^*\xi, f_j] = \sum_{i=1}^{l} \sum_{j=1}^{p} \sum_{\nu=1}^{\mu_j} K_{ij}^{(\nu)} f_{j\nu} \xi_i. \tag{92}$$

As a result $K^* : F^l \longrightarrow (X_{z^{\mu_1}} \oplus \cdots \oplus X_{z^{\mu_p}})^* = X^{z^{\mu_1}} \oplus \cdots \oplus X^{z^{\mu_p}}$ is given by

$$K^*\xi = \begin{pmatrix} \sum_{i=1}^{l} \sum_{\nu=1}^{\mu_1} \frac{K_{i1}^{(\nu)} \xi_i}{z^\nu} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=1}^{l} \sum_{\nu=1}^{\mu_p} \frac{K_{ip}^{(\nu)} \xi_i}{z^\nu} \end{pmatrix} \tag{93}$$

Next we introduce the *dual indices* to the μ_i. They are defined by

$$\lambda_k = \sharp\{\mu_j \mid \mu_j \geqslant k\}. \tag{94}$$

Obviously, we have

$$\lambda_1 \geqslant \cdots \geqslant \lambda_{\mu_1} \geqslant 0, \tag{95}$$

and $\sum_{k=1}^{\mu_1} \lambda_k = \sum_{j=1}^{p} \mu_j = n$.

For fixed indices $1 \leqslant \nu \leqslant \mu_1$ and $1 \leqslant i \leqslant l$, we have $1 \leqslant j \leqslant \lambda_i$. Eq. (92) can be rewritten as

$$[K^*\xi, f] = \sum_{\nu=1}^{\mu_j} \sum_{i=1}^{l} \sum_{j=1}^{\lambda_i} K_{ij}^{(\nu)} f_{j\nu} \xi_i. \tag{96}$$

So an index ν defines an $l \times \lambda_j$ matrix by

$$K^{(\nu)} = \left(K_{ij}^{(\nu)}\right). \tag{97}$$

We will say that a sequence of $l \times \lambda_\nu$ matrices $K^{(\nu)}$ is a *nice sequence* if $\lambda_1 \geqslant \cdots \geqslant \lambda_m$. We will say that a system

$$\left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right)$$

in the state space F^k is a *partial realization* of the sequence $\{K^{(\nu)}\}$ if

$$K_{ij}^{(\nu)} = C_i A^{\nu-1} B_j, \quad i = 1, \ldots, l; \quad j = 1, \ldots, \lambda_\nu. \tag{98}$$

It is a minimal partial realization if there exists no partial realization of smaller McMillan degree.

In particular, (95) implies that the sequence $K^{(1)}, \ldots, K^{(\mu_1)}$ is a *nice sequence*. Moreover, by reordering the columns $CA^i B_j$ in the following way:

$$\begin{aligned} K &= (K_{.1}^{(1)}, \ldots, K_{.\mu_1}^{(1)}, \ldots, K_{.1}^{(p)}, \ldots, K_{.\mu_p}^{(p)}) \\ &= (CB_1, \ldots, CA^{\mu_1-1}B_1, \ldots, CB_p, \ldots, CA^{\mu_p-1}B_p), \end{aligned}$$

the partial realization condition can be written in the form

$$K = C\mathscr{R}_\mu(A, B). \tag{99}$$

Much of the preceeding discussion has great relevance to observer theory. In the construction of observers stability considerations are paramount. However, for clarity, it seems to us beneficial to decouple the algebraic analysis of the observer construction from stability considerations. Thus we make the following definition.

Definition 5.1. Given a minimal linear system

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \\ y_t &= Cx_t, \\ r_t &= Kx_t \end{aligned} \tag{100}$$

with A, B, C, and K in $F^{n\times n}$, $F^{n\times m}$, $F^{p\times n}$, and $F^{l\times n}$, respectively. We assume C and K have both full row rank. A *preobserver* of order q for K is a system

$$\begin{aligned} z_{t+1} &= Fz_t + Gy_t + Hu_t, \\ w_t &= Lz_t \end{aligned} \tag{101}$$

with F, G, H, and L in $F^{q\times q}$, $F^{q\times p}$, $F^{q\times m}$, and $F^{l\times q}$, respectively, such that there exists a surjective linear transformations, Z for which

$$\begin{aligned} ZA - FZ &= GC, \\ H &= ZB, \\ K &= LZ. \end{aligned} \tag{102}$$

The richness of the equation $\mathscr{V} = X_D \cap TF^p[z]$ comes to light again in its use for the construction of preobservers.

Theorem 5.1. *Given the minimal linear system* (100). *The following conditions are equivalent*:

1. *There exists a preobserver for K.*
2. *There exists a conditioned invariant subspace* $\mathscr{V} \subset \mathscr{X}$ *satisfying*

$$\mathscr{V} \subset \operatorname{Ker} K. \tag{103}$$

Proof. (1) $\Rightarrow$ (2) We assume a preobserver exists and define $\mathscr{V} = \operatorname{Ker} Z$. We show now that $\mathscr{V}$ is conditioned invariant. To this end we assume $f \in \operatorname{Ker} Z \cap \operatorname{Ker} C$. We have therefore $ZAf = FZf + GCf = 0$, and so $Af \in \operatorname{Ker} Z = \mathscr{V}$. Since $K = LZ$, it follows that $\mathscr{V} \subset \operatorname{Ker} K$.

(2) $\Rightarrow$ (1) We assume a conditioned invariant subspace $\mathscr{V} \subset \operatorname{Ker} K$ is given. We also assume the coprime factorizations $C(zI - A)^{-1} = D(z)^{-1}\Phi(z)$. So without loss of generality we can assume that the pair (C, A) is given by means of the shift realization (24). So in this case $X = X_D$ and $\mathscr{V} = X_D \cap TF^p[z]$, where $D^{-1}T$ has nonpositive factorization indices. We take X_T as the state space of the preobserver and define a map $Z : X_D \longrightarrow X_T$ by $Z = \pi_T \mid X_D$. By Theorem 3.3, Z is surjective. Next we define

$$F = S_T, \quad G\xi = \pi_T \xi. \tag{104}$$

Then

$$\begin{aligned}(ZA - FZ)f &= (\pi_T S_D - S_T \pi_T)f = \pi_T \pi_D zf - \pi_T z \pi_T f \\ &= \pi_T(zf - D(z)\xi_f) - \pi_T zf = -\pi_T D(z)\xi_f\end{aligned}$$

Recalling that $Cf = (D^{-1}f)_{-1} = \xi_f$, the condition $f \in \operatorname{Ker} C$ implies $(\pi_T S_D - S_T \pi_T)f = 0$ or $\operatorname{Ker}(\pi_T S_D - S_T \pi_T) \supset \operatorname{Ker} C$. Thus there exists a map G such that $ZA - FZ = GC$. Since C is assumed surjective, it follows that $\operatorname{Im} G \subset \operatorname{Im} Z$ and hence there exists a J such that $G = ZJ$. The map J is a friend of $\mathscr{V}$ as the equality $ZA - FZ = GC$ implies $Z(A - JC) = FZ$. In particular, it follows that $f \in \operatorname{Ker} Z$ implies $(A - JC)f \in \operatorname{Ker} Z$. Of course J is not uniquely determined. By assumption we have $\mathscr{V} = \operatorname{Ker} Z \subset \operatorname{Ker} K$, and so there exists an L such that $K = LZ$. Finally, we define $H = ZB$ and we have obtained the preobserver equations. □

Existence of preobservers is related to partial realizations. We explore this connection next.

Theorem 5.2. *Let* $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ *with* $\mu_1 \geqslant \cdots \geqslant \mu_p \geqslant 0$ *and* $\mu_1 + \cdots + \mu_p = n$. *Assume* $(A, B) \in F^{k\times k} \times F^{k\times p}$ *is a reachable pair. Let* $H(z)T(z)^{-1}$ *be a right coprime factorization of* $(zI - A)^{-1}B$. *Then*:

1. *There exists a codimension k conditioned invariant subspace* $\mathscr{V} \subset \operatorname{Ker} K$ *if and only if the nice sequence* $K^{(1)}, \ldots, K^{(\mu_1)}$ *has a minimal, McMillan degree k partial realization.*
2. *The solution to the above partial realization problem is unique, up to state space isomorphisms, if and only if* $\mathscr{V}$ *is a* μ*-tight conditioned invariant subspace.*

Proof.

1. Assume there exists a codimension k conditioned invariant subspace $\mathbf{V}$ satisfying $\mathbf{V} \subset \operatorname{Ker} K$. Then, by Theorem 4.1, there exists a μ-regular pair $(A, B) \in F^{k\times k} \times F^{k\times p}$ such that $\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B)$. Since $\operatorname{Ker} \mathscr{R}_\mu(A, B) \subset \operatorname{Ker} K$, there exists a map $C \in F^{l\times k}$ for which
$$K = C\mathscr{R}_\mu(A, B). \tag{105}$$
Since the pair (A, B) is μ-regular, the map C is uniquely determined. Eq. (105) is a construction of a McMillan degree k partial realization. The partial realization is necessarily minimal, for otherwise there exists a partial realization of smaller McMillan degree and hence the codimension of $\mathbf{V}$ would be greater than k, contrary to our assumption.
Conversely, assume we are given the system (100) and a minimal, McMillan degree k partial realization
$$\left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right).$$
Thus (105) holds. We claim that the minimality of the partial realization (105) necessarily implies that the pair (A, B) is μ-regular, i.e. that $\mathscr{R}_\mu(A, B)$ has full row rank k. For otherwise we could find a state space transformation R such that
$$\mathscr{R}_\mu(A, B) = \mathscr{R}_\mu\left(\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}, \begin{pmatrix} B_1 \\ 0 \end{pmatrix}\right)$$
with $(A, B) \in F^{k_1\times k_1} \times F^{k_1\times p}$ and $k_1 < k$. Writing $C = \begin{pmatrix} C_1 & C_2 \end{pmatrix}$ with $C_1 \in F^{l\times k_1}$, it would follow that
$$\left(\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & 0 \end{array}\right)$$
is a partial realization of lower McMillan degree, contradicting our assumption of minimality. Defining now $\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B)$, it follows from Theorem 4.1 that $\mathbf{V}$ is a codimension k conditioned invariant subspace which, by (105), satisfies $\mathbf{V} \subset \operatorname{Ker} K$.
2. Assume the pair (A, B) is a μ-tight pair and that $\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B) \subset \operatorname{Ker} K$. By Theorem 4.1, $\mathbf{V}$ is a tight subspace. By Theorem 3.5, the pair (A, B) is uniquely determined up to similarity. Moreover, since $\mathscr{R}_\mu(A, B)$ has full row rank, C determined by (105) is uniquely determined. Thus this partial realization is minimal and unique up to similarity.
Conversely, assume the partial realization problem has a unique, McMillan degree k, solution given by (105). Define $\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B)$. Clearly, by Theorem 4.1, $\mathbf{V}$ is a codimension k conditioned invariant subspace. Applying Theorem 3.5 once more, the uniqueness of the partial realization implies that the pair (A, B) is μ-tight or, equivalently, that $\mathbf{V} = \operatorname{Ker} \mathscr{R}_\mu(A, B)$ is a tight conditioned invariant subspace. $\square$

We proceed now with the study of asymptotic observers. Observer theory has a long history starting with the work of Luenberger [31].

Definition 5.2. Given a minimal linear system

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx, \\ r &= Kx \end{aligned} \tag{106}$$

with A, B, C, and K in $\mathbf{R}^{n\times n}$, $\mathbf{R}^{n\times m}$, $\mathbf{R}^{p\times n}$, and $\mathbf{R}^{l\times n}$, respectively. An observer of order q for K is a system

$$\begin{aligned} \dot{z} &= Fz + Gy + Hu, \\ w &= Dz \end{aligned} \tag{107}$$

with F, G, H, and D in $\mathbf{R}^{q\times q}$, $\mathbf{R}^{q\times p}$, $\mathbf{R}^{q\times m}$, and $\mathbf{R}^{l\times q}$, respectively, such that, for all controls u and all initial conditions of the state, we have $\lim_{t\to\infty}(r(t) - w(t)) = 0$.

So an observer for a linear function of the state is a linear system that uses the input and outputs of the original system to give an asymptotic estimate of the linear function. A special case would be of a state observer. Throughout the rest of the paper we will assume that both C and K have full row rank.

We will need the following several auxiliary results. We begin with a simple lemma.

Lemma 5.1. *Let A be a linear transformation in a linear space $\mathscr{X}$. $T : \mathscr{X} \longrightarrow \mathscr{Y}$. Then*

$$A \operatorname{Ker} T \subset \operatorname{Ker} T \tag{108}$$

if and only if

$$\operatorname{Ker} T \subset \operatorname{Ker} TA. \tag{109}$$

The second condition implies the existence of a linear transformation F such that $TA = FT$.

Proof. Assume (108) and let $f \in \operatorname{Ker} T$. Then $Af \in \operatorname{Ker} T$ or $TAf = 0$, i.e. (109) holds.

Conversely, assume (109). This implies that, for some F, $TA = FT$. We show now that if $f \in \operatorname{Ker} T$ then also $Af \in \operatorname{Ker} T$. We compute

$$T(Af) = F(Tf) = 0,$$

i.e. $Af \in \operatorname{Ker} T$. $\square$

Our next lemma plays a critical role in the proof of Theorem 5.3.

Lemma 5.2.

1. *Let $VT^{-1}U = W$, where T, U, V, W are polynomial matrices of appropriate size and T, U are left coprime. Then there exists a polynomial matrix V_1 for which $V = V_1T$.*
2. *Let (D, F) be an observable pair and let (A, B) be a reachable pair. Then for any linear polynomial $Mz + N$ we have*

$$D(zI - F)^{-1}(Mz + N)(zI - A)^{-1}B = 0 \tag{110}$$

if and only if there exists constant matrices X, Y, with $DY = 0$ and $XB = 0$, for which

$$Mz + N = X(zI - A) - (zI - F)Y. \tag{111}$$

Proof.

1. Since T, U are left coprime, there exists a polynomial solution to the Bezout equation $TX + UY = I$. This implies

$$VT^{-1} = VX + VT^{-1}UY = VX + WY = V_1,$$

and hence $V = V_1T$.

2. If $Mz + N$ has representation (111) with the conditions $DY = 0$ and $XB = 0$ satisfied, then clearly (110) holds.

To prove the converse, we observe that using the identity

$$D(zI - F - JD)^{-1} = (I - D(zI - F)^{-1}J)^{-1}D(zI - F)^{-1}$$

and the pole assignment property for output injection, we can assume without loss of generality that F and A have disjoint spectra. Next, note that we have the partial fraction decomposition

$$(zI - F)^{-1}(Mz + N)(zI - A)^{-1} = (zI - F)^{-1}X - Y(zI - A)^{-1},$$

where X and Y satisfy (111). In turn this is equivalent to

$$X - Y = M,$$
$$FY - XA = N.$$

Now, by the disjointness of the spectra, we have

$$\begin{aligned} &D(zI - F)^{-1}(Mz + N)(zI - A)^{-1}B \\ &\quad = D(zI - F)^{-1}XB - DY(zI - A)^{-1}B = 0 \end{aligned}$$

if and only if

$$D(zI - F)^{-1}XB = DY(zI - A)^{-1}B = 0.$$

By our assumptions on reachability and observability, this is equivalent to $XB = 0$ and $DY = 0$. □

Our starting point is the following geometric characterization of the existence of observers for linear functions of the state. To state the theorem, we need a new concept. We say that a controlled invariant subspace $\mathscr{V}$ for the pair (A, B) is stabilizable,

or *inner stabilizable* if there exists a feedback F such that $\mathscr{V}$ is $A + BF$-invariant and $A + BF|_{\mathscr{V}}$ is stable. A subspace $\mathscr{V}$ is *outer detectable* for the pair (C, A) if there exists an output injection map H such that $\mathscr{V}$ is $A + HC$-invariant and $A + HC|_{X/\mathscr{V}}$ is stable.

Next we give a characterization of outer detectable and inner stabilizable, in the context of the shift realizations, that are generalizations of Proposition 3.1 in the sense that they incorporate stability. We omit the proof and refer to a forthcoming paper [15].

Proposition 5.1. *Let G be a $p \times m$ transfer function having the polynomial coprime factorizations*

$$G = E_r D_r^{-1} = D_l^{-1} E_l. \tag{112}$$

Then:

1. *With respect to the realization*

$$\begin{aligned} A &= S_{D_l}, \\ Cf &= (D_l^{-1} f)_{-1} \end{aligned} \tag{113}$$

in the state space X_{D_l}, a subspace $\mathscr{V}$ is outer antidetectable if and only if

$$\mathscr{V} = X_{D_l} \cap F_+ F^m[s] \tag{114}$$

with F_+ an antistable polynomial matrix F_+ for which all the left factorization indices of $D_l^{-1} F_+$ are nonpositive.

2. *With respect to realization* (113) *in the state space X_{D_l}, a subspace $\mathscr{V}$ is outer detectable if and only if*

$$\mathscr{V} = X_{D_l} \cap F_- F^m[s] \tag{115}$$

for some stable polynomial matrix F_- for which all the left factorization indices of $D_l^{-1} F_+$ are nonpositive.

3. *With respect to the realization*

$$\begin{aligned} A &= S^{D_r}, \\ B\xi &= \pi_- D_r^{-1} \xi \end{aligned} \tag{116}$$

in the state space X^{D_r}, a subspace $\mathscr{V} \subset X^{D_r}$ is inner stabilizable if and only if there exists a stable polynomial matrix E_- such that

$$\mathscr{V} = \pi^{D_r} X^{E_-} \tag{117}$$

and for which all the right factorization indices of $E_- D_r^{-1}$ are nonpositive.

4. *With respect to realization* (116) *in the state space X^{D_r}, a subspace $\mathscr{V} \subset X^{D_r}$ is inner antistabilizable if and only if there exists an antistable polynomial matrix E_+ such that*

$$\mathscr{V} = \pi^{D_r} X^{E_+} \tag{118}$$

and for which all the right factorization indices of $E_- D_r^{-1}$ are nonpositive.

Theorem 5.3. *Given the minimal linear system* (106). *The following conditions are equivalent*:

1. *There exists an observer for K.*
2. *There exist linear transformations*, T, F, G, H, D (*with T surjective*) *and F stable*, *such that*

$$\begin{aligned} TA - FT &= GC, \\ H &= TB, \\ K &= DT. \end{aligned} \tag{119}$$

3. *There exist proper stable rational functions* M, N *that solve*

$$\begin{pmatrix} M & N \end{pmatrix} \begin{pmatrix} zI - A \\ C \end{pmatrix} = K. \tag{120}$$

4. *With* $Z_K(z) = K(zI - A)^{-1}B$ *and* $Z_C(z) = C(zI - A)^{-1}B$, *there exist proper stable rational functions* Z_1, Z_2 *that solve*

$$Z_K = Z_1 Z_C + Z_2. \tag{121}$$

5. *There exists an outer detectable subspace* $\mathscr{V} \subset \mathscr{X}$ *satisfying*

$$\mathscr{V} \subset \operatorname{Ker} K. \tag{122}$$

Proof. (1) $\Rightarrow$ (2) Assume an observer for K is given by Eqs. (107). The transfer function of the observer be given by

$$\begin{pmatrix} Z_1 & Z_2 \end{pmatrix} = \left(\begin{array}{c|cc} F & G & H \\ \hline D & 0 & 0 \end{array}\right), \tag{123}$$

i.e.

$$\begin{aligned} Z_1(z) &= D(zI - F)^{-1}G, \\ Z_2(z) &= D(zI - F)^{-1}H. \end{aligned}$$

Solving (106) with initial condition $x(0) = \xi$, we get for the Laplace transforms

$$x = (zI - A)^{-1}Bu + (zI - A)^{-1}\xi.$$

We compute now

$$\begin{aligned} w &= Z_1(z)y + Z_2(z)u = Z_1(z)Cx + Z_2(z)u \\ &= Z_2(z)u + Z_1(z)C[(zI - A)^{-1}Bu + (zI - A)^{-1}\xi] \end{aligned}$$

So

$$\begin{aligned} w - Kx = {} & [Z_2(z) + (Z_1(z)C - K)(zI - A)^{-1}B]u \\ & + (Z_1(z)C - K)(zI - A)^{-1}\xi. \end{aligned}$$

Since, for all controls u and all initial conditions ξ we have $\lim_{t\to\infty} w(t) - Kx(t) = 0$, it follows that

$$Z_2(z) + (Z_1(z)C - K)(zI - A)^{-1}B = 0. \tag{124}$$

and that $-(Z_1(z)C - K)(zI - A)^{-1}$ is stable. In turn this implies that $Z_2(z) = -(Z_1(z)C - K)(zI - A)^{-1}B$ is stable. Thus, Eq. (124) can be rewritten as

$$D(zI - F)^{-1}H = -[D(zI - F)^{-1}GC - K](zI - A)^{-1}B, \tag{125}$$

or

$$K(zI - A)^{-1}B = D[(zI - F)^{-1}H + (zI - F)^{-1}GC(zI - A)^{-1}B].$$

By the controllability of the pair (A, B), we conclude that $\operatorname{Im} K \subset \operatorname{Im} D$. Since K is assumed surjective, we have $\operatorname{Im} K = \operatorname{Im} D$ and so $K = DT$ for some surjective linear transformation T. Similarly, from (125) we immediately see that $\operatorname{Ker} B \subset \operatorname{Ker} H$. Thus $H = T_2B$ for some linear transformation T_2. Eq. (125) can therefore be rewritten as

$$D(zI - F)^{-1}[(zI - F)T_1 - T_2(zI - A) - GC](zI - A)^{-1}B = 0.$$

By applying Lemma 5.2, we can write

$$(zI - F)T_1 - T_2(zI - A) - GC = X(zI - A) - (zI - F)Y$$

with $DY = 0$ and $XB = 0$. Comparing terms, we obtain

$$\begin{aligned} &T_1 - T_2 - X + Y = 0, \\ &-FT_1 + T_2A - GC + XA - FY = 0. \end{aligned}$$

Defining $T := T_1 + Y = T_2 + X$, we have $TA - FT = GC$. Furthermore, we note that

$$\begin{aligned} K &= DT_1 = D(T_1 + Y) = DT, \\ H &= T_2B = (T_2 + X)B = TB. \end{aligned}$$

The stability of $-(Z_1(z)C - K)(zI - A)^{-1}$ implies the stability of F. To see this note that if $(Z_1(z)C - K)(zI - A)^{-1}$ is stable so is $Z_1(z)C$. As C is assumed to be of full row rank, the stability of $Z_1(z)$ follows. Now (123) is a minimal realization of a stable transfer function, so necessarily, F is stable.

(2) $\Rightarrow$ (3) Defining Z_1, Z_2 by (123), and setting

$$\begin{aligned} M(z) &= -(Z_1(z)C - K)(zI - A)^{-1}, \\ N(z) &= Z_1 \end{aligned}$$

we obtain (120).

(3) $\Rightarrow$ (4) Rewrite (120) as

$$\begin{pmatrix} M & N \end{pmatrix} \begin{pmatrix} I \\ C(zI - A)^{-1} \end{pmatrix} = K(zI - A)^{-1},$$

which implies

$$\begin{pmatrix} M & N \end{pmatrix} \begin{pmatrix} B \\ C(zI-A)^{-1}B \end{pmatrix} = K(zI-A)^{-1}B.$$

Thus (121) is solvable with

$$\begin{aligned} Z_1(z) &= N(z), \\ Z_2(z) &= M(z)B. \end{aligned} \tag{126}$$

(4) $\Rightarrow$ (3) Assume (121) is solvable with Z_1, Z_2 stable and proper. Thus

$$K(zI-A)^{-1}B = \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} \begin{pmatrix} C(zI-A)^{-1}B \\ B \end{pmatrix}$$

Setting $Z_2 = K(zI-A)^{-1}B - Z_1C(zI-A)^{-1}B$, we have $Z_2 = Z_2'B$. This implies

$$K(zI-A)^{-1}B = \begin{pmatrix} Z_1 & Z_2' \end{pmatrix} \begin{pmatrix} C \\ zI-A \end{pmatrix} (zI-A)^{-1}B$$

Hence, by the controllability of (A, B) and Lemma 5.2, we conclude (120) with

$$\begin{aligned} M(z) &= Z_2'(z), \\ N(z) &= Z_1(z). \end{aligned} \tag{127}$$

(2) $\Rightarrow$ (4) Assume transformations T, F, G, H, D exist and satisfy (119). We show now that Eq. (120) is solvable with Z_1, Z_2 defined by

$$\begin{pmatrix} Z_1 & Z_2 \end{pmatrix} = \left(\begin{array}{c|cc} F & G & H \\ \hline D & 0 & 0 \end{array}\right). \tag{128}$$

We show now that Eq. (120) is solvable. To this end we compute

$$\begin{aligned} &K(zI-A)^{-1}B - D(zI-F)^{-1}GC(zI-A)^{-1}B - D(zI-F)^{-1}H \\ &\quad = DT(zI-A)^{-1}B - D(zI-F)^{-1}(TA-FT)(zI-A)^{-1}B \\ &\qquad -D(zI-F)^{-1}TB \\ &\quad = DT(zI-A)^{-1}B - D(zI-F)^{-1}[(zI-F)T - T(zI-A)] \\ &\qquad (zI-A)^{-1}B - D(zI-F)^{-1}TB \\ &\quad = DT(zI-A)^{-1}B - DT(zI-A)^{-1}B + D(zI-F)^{-1}TB \\ &\qquad -D(zI-F)^{-1}TB = 0. \end{aligned}$$

Note that by the stability of F, Z_1, Z_2 are stable and strictly proper.

(3) $\Rightarrow$ (2) Let

$$\begin{pmatrix} N & M \end{pmatrix} = \left(\begin{array}{c|cc} F & G & H_1 \\ \hline D & 0 & 0 \end{array}\right),$$

with the realization minimal. From (120) we get

$$D(zI-F)^{-1}H_1(zI-A) + D(zI-F)^{-1}GC = K.$$

This implies $\operatorname{Im} K \subset \operatorname{Im} D$ and hence there exists a linear transformation T for which $K = DT$. The previous equation implies now

$$D(zI-F)^{-1}H_1B + D(zI-F)^{-1}GC(zI-A)^{-1}B = DT(zI-A)^{-1}B,$$

which can be rewritten as

$$D(zI-F)^{-1}\left[H_1(zI-A) + GC - (zI-F)T\right](zI-A)^{-1}B = 0.$$

Applying Lemma 5.2, we conclude $H_1 = T$ and $TA = FT + GC$. The stability of F is implied by the stability of M, N and the minimality of the realization.

(4) $\Rightarrow$ (2) Assume (121) is solvable with Z_1, Z_2 strictly proper and stable. Let

$$\left(\begin{array}{c|cc} F & G & H \\ \hline D & 0 & 0 \end{array}\right)$$

be a minimal realization of $\begin{pmatrix} Z_1 & Z_2 \end{pmatrix}$. By our assumption F is necessarily stable. Moreover, by our assumption of minimality, the pair (F, D) is observable. So $Z_1 = D(zI-F)^{-1}G$ and $Z_2 = D(zI-F)^{-1}H$, and we have

$$K(zI-A)^{-1}B - D(zI-F)^{-1}GC(zI-A)^{-1}B - D(zI-F)^{-1}H = 0. \tag{129}$$

By the controllability of the pair (A, B), we necessarily have $\operatorname{Im} K \subset \operatorname{Im} D$. Similarly by the observability of the pair (D, F), we necessarily have $\operatorname{Ker} B \subset \operatorname{Ker} H$. Thus there exist linear transformations T_1 and H_1 such that $K = DT_1$ and $H = H_1B$. We can rewrite (129) as

$$D(zI-F)^{-1}[GC - (zI-F)T_1 + H_1(zI-A)](zI-A)^{-1}B = 0.$$

Applying Lemma 5.2, we conclude that

$$GC - (zI-F)T_1 + H_1(zI-A) = X(zI-A) - (zI-F), \tag{130}$$

with $DY = 0$ and $XB = 0$. Setting $T = T_1 - Y = H_1 - X$, we still have $K = DT$ and $H = TB$ and moreover $GC + FT - TA = 0$. So (119) follows.

(2) $\Rightarrow$ (1) Assume there exist matrices such that (119) is satisfied. We define the observer by Eqs. (107). Clearly $H = TB$ and $K = DT$, taken together, imply $DH = KB$. Set $\epsilon(t) = Tx(t) - z(t)$. Differentiating, we get

$$\begin{aligned}
\dot{\epsilon} &= T\dot{x} - \dot{z} \\
&= T(Ax + Bu) - (Fz + Gy + Hu) \\
&= (TAx - Fz - GCx) + (TB - H)u \\
&= (TA - FT - GC)x + F(Tx - z) = F\epsilon.
\end{aligned}$$

Since F is stable, we have $\lim_{t\to\infty} \epsilon(t) = 0$. In turn, this implies that $e = r - w = Kx - Dz = D(Tx - z) \to 0$.

(5) $\Rightarrow$ (2) Assume there exists an outer detectable subspace $\mathscr{V} \subset \operatorname{Ker} K \subset \mathscr{X}$. Thus there exists an output injection map L for which $(A - LC)\mathscr{V} \subset \mathscr{V}$ and the induced map $A - LC \mid \mathscr{X}/\mathscr{V}$ is stable. Set $\mathscr{V} = \operatorname{Ker} T$ for some surjective linear transformation T. By Lemma 5.1, there exists a map F such that $T(A - LC) = FT$, or $TA - FT = GC$ with $G = TL$.

Since $\operatorname{Ker} T \subset \operatorname{Ker} K$ we have $K = DT$ for some D. The stability of the induced map $(A - LC) \mid \mathscr{X}/\mathscr{V}$ implies the stability of F. Finally, we define $H = TB$. Thus Eqs. (119) hold.

(2) $\Rightarrow$ (5) We show first that $\operatorname{Ker} T$ is a conditioned invariant subspace. We use Eqs. (119) to conclude that $TA = FT + GC$. Assume $f \in \operatorname{Ker} T \cap \operatorname{Ker} C$, then $TAf = FTf + GCf = 0$, i.e. $Af \in \operatorname{Ker} T$ or $A(\operatorname{Ker} T \cap \operatorname{Ker} C) \subset \operatorname{Ker} T$. The surjectvity of T, the assumption that C has full row rank and the equality $TA - FT = GC$ show that $\operatorname{Im} G \subset \operatorname{Im} T$, or $G = TL$ for some L, which in turn can be used to write $T(A - LC) = FT$. This incidentally shows that $(A - LC)\operatorname{Ker} T \subset \operatorname{Ker} T$. Moreover, it implies the commutativity of the following diagram.

$$\begin{tikzcd}
\mathcal{X} \arrow[r, "T"] \arrow[d, "A - LC"'] & \operatorname{Im} T \arrow[d, "F"] \\
\mathcal{X} \arrow[r, "T"] & \operatorname{Im} T
\end{tikzcd}$$

This induces a quotient diagram with the induced map $\overline{T}$ an isomorphism.

$$\begin{tikzcd}
\mathcal{X}/\mathcal{V} \arrow[r, "\overline{T}"] \arrow[d, "A - LC|_{\mathcal{X}/\mathcal{V}}"'] & \operatorname{Im} T \arrow[d, "F"] \\
\mathcal{X}/\mathcal{V} \arrow[r, "\overline{T}"] & \operatorname{Im} T
\end{tikzcd}$$

Since F is stable, $\mathscr{V}$ is outer detectable. $\square$

Some remarks are in order. The equivalence of conditions (1) and (3) was proved by Hautus and Sontag [24] and we follow their proof. Eq. (129) can be considered as a partial fraction decomposition. For more on multivariable partial fractions, see [13]. Note that no assumptions about the order of the controller were made. We can easily strengthen this result to include order considerations as follows. Most of the proof follows the line of the previous theorem and we outline only the changes made to connect order and McMillan degree.

Theorem 5.4. *Let $K \in \mathbf{R}^{l\times n}$ be of full rank. The following conditions are equivalent*:

1. *There exists a Kalman observer for K of order q.*
2. *There exists a surjective linear transformation T of rank q and F stable such that*

$$\begin{aligned} &TA - FT = GC,\\ &H = TB,\\ &K = DT. \end{aligned} \tag{131}$$

3. *There exist proper stable rational functions M, N, with the McMillan degree of $\begin{pmatrix} M & N \end{pmatrix}$ equal q that solve*

$$\begin{pmatrix} M & N \end{pmatrix} \begin{pmatrix} zI - A \\ C \end{pmatrix} = K. \tag{132}$$

4. *With $Z_K(z) = K(zI - A)^{-1}B$ and $Z_C(z) = C(zI - A)^{-1}B$, there exist proper stable rational functions Z_1, Z_2, with the McMillan degree of $\begin{pmatrix} Z_1 & Z_2 \end{pmatrix}$ equal q that solve*

$$Z_K = Z_1 Z_C + Z_2. \tag{133}$$

5. *There exists an outer detectable subspace $\mathscr{V} \subset \mathscr{X}$ of codimension q satisfying*

$$\mathscr{V} \subset \operatorname{Ker} K. \tag{134}$$

Proof. (1) $\Rightarrow$ (4) With (Z_1, Z_2) defined by (123), we obviously have $\delta(Z_1, Z_2) \leqslant q$.

(3) $\Rightarrow$ (4) We assume $\delta \begin{pmatrix} M & N \end{pmatrix} = q$. By (126), we have $\delta \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} \leqslant q$.

(4) $\Rightarrow$ (3) We assume $\delta \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} = q$. By (127), we have $\begin{pmatrix} M & N \end{pmatrix} \leqslant q$.

(2) $\Rightarrow$ (1) Clearly the system

$$\left(\begin{array}{c|cc} F & G & H \\ \hline D & 0 & 0 \end{array}\right)$$

has order q and it is an observer by the previous theorem.

(5) $\Rightarrow$ (2) We set $\mathscr{V} = \operatorname{Ker} T$. We compute

$$\operatorname{rank} T = n - \dim \operatorname{Ker} T = n - \dim \mathscr{V} = \operatorname{codim} \mathscr{V} = q.$$

So T is a map onto $\operatorname{Im} T$ is surjective and of rank q. The rest follows the lines of the previous proof.

(2) $\Rightarrow$ (5) We set as before $\mathscr{V} = \operatorname{Ker} T$. Since $\operatorname{rank} T = q$, we compute

$$\operatorname{codim} \mathscr{V} = n - \dim \operatorname{Ker} T = \operatorname{rank} T = q. \qquad \square$$

Some remarks are in order. The equivalence of conditions (1) and (5) is due to Kawaji [29].

In case we want to have a state observer, then $K = I$ and $\operatorname{Ker} K = \{0\}$. So the only conditioned invariant subspace $\mathscr{V}$ of X_D satisfying $\mathscr{V} \subset \operatorname{Ker} K$ is the zero subspace. Now $\mathscr{V} = X_D \cap TF^p[z]$ which leads to a polynomial matrix T for which all the left

factorization indices of $D^{-1}T$ are trivial. This means that $\pi_T \mid X_D$ is injective and hence invertible. Thus $\dim X_T = \dim X_D = n$ and the observer will have order n. In general, i.e. for an arbitrary K, if we look for a minimal order observer, and since the order of the observer is $\deg\det T = \operatorname{codim} \mathscr{V}$, it is equivalent to look for a maximal dimensional conditioned invariant subspace in $\operatorname{Ker} K$. Of course, since the sum of conditioned invariant subspaces is not necessarily a conditioned invariant subspace, there may be many such subspaces.

Without delving into the details, we wish to point out that Theorems 5.3 and 5.4 have a potential, natural extension in the context of Hardy spaces. In fact, using the shift realization for continuous time, stable systems, the natural state space turns out to be a coinvariant subspace of a vectorial Hardy space. In this context we have characterizations for outer detectable subspaces in the spirit of Proposition 5.1. The details can be found in [15,18].

The results of Proposition 5.1 and of Theorem 5.3 can be applied to clarify the connections between asymptotic observer theory and the stable partial realization problem. This extends the results of Theorem 5.2 by incorporating stability considerations. That these two problems are related is known for a long time. Some references are Refs. [1,33].

Theorem 5.5. *Given the minimal linear system* (106). *Assume without loss of generality that the pair* (A, C) *is in Brunovsky form with observability indices* $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$. *Then there exists an order q observer for K if and only if the sequence* $K^{(1)}, \ldots, K^{(\mu_1)}$ *of matrices, defined in* (97), *has a McMillan degree q stable partial realization.*

Proof. Assume first that the sequence $K^{(1)}, \ldots, K^{(\mu_1)}$ has a McMillan degree q stable partial realization given by

$$\left(\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & 0 \end{array}\right).$$

We need to show the existence of an outer detectable subspace $\mathscr{V} \subset X_\Delta$ of codimension q satisfying $\mathscr{V} \subset \operatorname{Ker} K$. Let $E(z)^{-1}H(z)$ be a left coprime factorization of $C_1(zI - A_1)^{-1}$. Since A is stable, $E(z)$ is a stable polynomial matrix. By the assumed realization, we have, for each $\xi \in \mathbf{R}^l$, that $K^*\xi \in \mathscr{W} = \pi^{\tilde{\Delta}} X^{\tilde{E}}$, or $\operatorname{Im} K^* \subset \mathscr{W}$. This clearly implies that $\mathscr{W}$ is, by Theorem 5.1, an inner stabilizable subspace of $X^{\tilde{\Delta}}$. Now the annihilator of this subspace in X_Δ is clearly $\mathscr{V} = X_\Delta \cap E\mathbf{R}^p[z]$, which is an outer detectable subspace. Finally, by the minimality of the realization, the map $\pi^{\tilde{\Delta}} \mid X^{\tilde{E}}$ is injective. Therefore $\dim \mathscr{W} = \deg\det \tilde{E} = \deg\det E = q$ and $\operatorname{codim} \mathscr{V} = q$. By Theorem 5.4 an order q observer for K exists.

Conversely, assume that an order q observer for K exists. By Theorem 5.4, there exists a codimension q outer detectable subspace $\mathscr{V}$ that satisfies $\mathscr{V} \subset \operatorname{Ker} K$. By the characterization in Proposition 5.1, an outer detectable subspace $\mathscr{V} \subset X_\Delta$ has

the representation $\mathscr{V} = X_\Delta \cap E\mathbf{R}^p[z]$, with E a stable polynomial matrix for which all the left Wiener–Hopf factorization indices of $\Delta^{-1}E$ are nonpositive.

Thus, using duality, we have $\mathscr{W} = \mathscr{V}^\perp = \pi^{\tilde{\Delta}} X^{\tilde{E}} \supset \operatorname{Im} K^*$. Since, by Theorem 3.4, the map $\pi^{\tilde{\Delta}} \mid X^E$ is injective, it follows that $\dim \mathscr{W} = \dim X^{\tilde{E}} = \deg\det \tilde{E} = \deg\det E = q$.

Now, for each vector $\xi = \begin{pmatrix}\xi_1 & . & . & . & \xi_l\end{pmatrix}^{\sim} \in \mathbf{R}^l$, we have

$$\begin{aligned} K^*\xi &= \begin{pmatrix} K_1^*\xi & \cdot & \cdot & \cdot & K_p^*\xi \end{pmatrix} \\ &= \left(\sum_{\nu=1}^{l} \sum_{j=1}^{\mu_1} \xi_\nu \frac{K_{1j}^{(\nu)}}{z^j} \quad . \quad . \quad . \quad \sum_{\nu=1}^{l} \sum_{j=1}^{\mu_p} \xi_\nu \frac{K_{pj}^{(\nu)}}{z^j} \right) \in \pi^{\tilde{\Delta}} X^{\tilde{E}}. \end{aligned}$$

Choosing succesively $\xi = e_\nu$ where e_ν is the νth unit vector in $\mathbf{R}^l$, we conclude the existence of elements $\tilde{\Theta}_\nu \in X^{\tilde{E}}$ for which $K^* e_\nu = \pi^{\tilde{\Delta}} \tilde{\Theta}_\nu$. Thus there exist polynomial vectors $\tilde{H}_\nu \in X_{\tilde{\Delta}}$ for which $\tilde{\Theta}_\nu = \tilde{E}^{-1}\tilde{H}_\nu$. Let $\tilde{H}(z)$ be the $l \times p$ polynomial matrix whose νth column is $\tilde{H}_\nu(z)$ and $\tilde{\Theta}$ the $l \times p$ stricly proper rational matrix whose columns are $\tilde{\Theta}_\nu$. Thus we have the, not necessarily left coprime, matrix fraction representation

$$\tilde{\Theta}(z) = \tilde{E}(z)^{-1}\tilde{H}(z). \tag{135}$$

Using the shift realization, with $X^{\tilde{E}}$ as the state space, it is clear that the McMillan degree of $\tilde{\Theta}$ is at most q. Taking this realization as

$$\left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right),$$

it is clear that

$$K_{ij}^{(\nu)} = C_\nu A^{j-1} B_i, \quad 1 \leqslant \nu \leqslant l, \quad 1 \leqslant i \leqslant \lambda_j,$$

i.e. the sequence $K^{(1)}, \ldots, K^{(\mu_1)}$ has an order q stable partial realization. □

6. On the Kronecker–Hermite canonical form

The main object of this section is the derivation of a canonical form for rectangular polynomial matrices under the action of right multiplication by unimodular polynomial matrices, i.e. under elementary column operations. Given a $p \times l$ polynomial matrix H, we want to exhibit the minimal number of nonintegral parameters. This means reducing the degrees of elements of H as much as possible. So reduction to column proper form is a natural step. However, after such a reduction, there is still some freedom of applying elementary column operation. This would allow us to choose a pivot element in each column and reduce elements in its row. This leads to a canonical form, introduced in [3,28], where it is called the Kronecker–Hermite form. This canonical form, and a variation of it, turns out to be a central tool in the parametrization of conditioned invariant subspaces. Moreover, in the case of square

nonsingular $p \times p$ polynomial matrices, modulo right multiplication by unimodular matrices, it leads to a parametrization of the set of reachable pairs $(A, B) \in F^{n\times n} \times F^{n\times p}$ modulo state space isomorphisms.

We recall the definition of the Kronecker–Hermite form.

Definition 6.1. Let $H(z)$ be a $p \times l$, full column rank polynomial matrix. Let $h_1, \ldots, h_l$ denote the columns of H and let h_{ij} denote the ith element of h_j. We say that H is in Kronecker–Hermite canonical form if there exists a uniquely determined set of indices $1 \leqslant i_1 < \cdots < i_l \leqslant p$, such that:
1. $h_{i_j j}$ is monic with $\delta_j := \deg h_{i_j j} = \deg h_j$.
2. We have $\deg h_{i_j k} < \delta_j$ for $1 \leqslant k \leqslant l, k \neq j$.
3. We have if $i > i_j$ then $\deg h_{ij} < \delta_j$.

Note that a matrix H in the Kronecker–Hermite canonical form is column proper and the $l \times l$ submatrix $\overline{H}$ consisting of the $i_1, \ldots, i_l$ rows is row proper with $[H]_l$, the matrix of leading row coefficients of H, equal to I_l.

It will be useful for our purposes to modify somewhat the definition of the Kronecker–Hermite canonical form.

Definition 6.2. Let $H(z)$ be a $p \times l$, full column rank polynomial matrix. We say that H is in *modified Kronecker–Hermite canonical form* if there exists a two, uniquely determined set of indices $\{\nu_1 < \cdots < \nu_s\}$ and $\{0 < k_1, \ldots, k_s\}$ with $\sum_{i=1}^{s} k_i = l$, and disjoint sets of row indices $R_i = \{1 \leqslant \rho_1^{(i)} < \cdots < \rho_{k_i}^{(i)} \leqslant p\}$ such that:
1. We have the partitioning

$$H(z) = (H_1(z); \ldots; H_s(z)). \tag{136}$$

2. $H^{(i)}$ is a $p \times k_i$ column proper matrix with all column degrees equal to ν_i.
3. Denoting the λ, μ entry of $H^{(i)}$ by $h_{\lambda,\mu}^{(i)}$, we have
   (a) $h_{\rho_j,j}^{(i)}$ is monic of degree ν_i.
   (b) For $\lambda > \rho_j$, we have $\deg h_{\lambda,j}^{(i)} < \deg h_{\rho_j,j}^{(i)}$.
   (c) For all $i = 1, \ldots, s$, we have

$$\deg h_{\rho_j,\mu}^{(i)} < \deg h_{\rho_j,j}^{(i)} \quad \text{for} \begin{cases} j = 1, \ldots, k_i, \\ \mu \neq j. \end{cases} \tag{137}$$

   (d) For all $i = 1, \ldots, s$, we have

$$\deg h_{\rho_j,\mu}^{(t)} < \deg h_{\rho_j,j}^{(i)} \quad \text{for} \begin{cases} t \neq i, \\ j = 1, \ldots, k_i, \\ \mu = 1, \ldots, k_t. \end{cases} \tag{138}$$

Partition (136) will be referred to as the *canonical partition* of H.

Note that the Kronecker–Hermite canonical form is obtained from the modified Kronecker–Hermite canonical form by a permutation of columns that arranges the set indices $\bigcup_{i=1}^{s} R_i$ in increasing order.

Before proceeding with the statement of the principal reduction result, we prove the following simple lemmas we shall use in the sequel.

Lemma 6.1. *Let $T(z) = T_0 + \cdots + T_\nu z^\nu \in F^{m\times m}[z]$ with T_ν nonsingular. Let $f \in F^m[z]$ with $\deg f = n$. Then f can be uniquely represented as*

$$f = Tg + r \tag{139}$$

with $r \in F^m[z]$ and $\deg r < \nu$. Moreover, we have $\deg g = \deg f - \nu$.

Proof. Clearly, the nonsingularity of T_ν implies the nonsingularity of T. We apply the projection π_T to f and write $r = \pi_T f$. Now $h = T^{-1}r$ is strictly proper, so

$$\begin{aligned} r(z) &= T(z)h(z) = (T_0 + \cdots + T_\nu z^\nu)\left(\frac{h_1}{z} + \frac{h_2}{z^2} + \cdots\right) \\ &= T_\nu h_1 z^{\nu-1} + \cdots, \end{aligned}$$

which shows that $\deg r \leqslant \nu - 1$. Assume $\deg g = \mu$. If $\deg f < \nu$, we have $g = 0$. Otherwise $\deg f = \deg T + \deg g$ for the highest term in Tg is $T_\nu g_\mu z^{\nu+\mu}$ and the coefficient is necessarily nonzero by the assumed nonsingularity of T_ν. Hence $\deg g = \deg f - \nu$. $\square$

The previous lemma can be extended to the case that $T(z)$ is a row proper matrix, see [3].

Lemma 6.2. *Let $T(z) \in F^{m\times m}[z]$ be row proper with row degrees $\nu_1, \ldots, \nu_m$. Let $f \in F^m[z]$. The f can be uniquely represented in the form*

$$f = Tg + r \tag{140}$$

with

$$g \in F^m[z], \quad r = \begin{pmatrix} r_1 \\ \cdot \\ \cdot \\ \cdot \\ r_m \end{pmatrix} \in F^m[z] \ \textit{and} \ \deg r_i < \nu_i.$$

Moreover, if $\gamma = \max_{1\leqslant i\leqslant m}(\deg f_i - \nu_i)$, then we have $\deg g = \gamma$ when $\gamma \geqslant 0$, otherwise $g = 0$.

Proof. Since T is row proper and square, it is necessarily nonsingular. As before, we apply the projection π_T to f and let $r = \pi_T f$. Since $f - \pi_T f \in \operatorname{Ker}\pi_T = TF^m[z]$ we get the representation $f = Tg + r$. Now $r \in X_T$ is, by

Proposition 2.1, is equivalent to $\deg r_i < \nu_i$. Any representation $f = Tg + r$ with $\deg r_i < \nu_i$ satisfied implies $r = \pi_T f$ which establishes uniqueness.

We have $g = 0$ if and only if $f \in X_T$ which is the case if and only if $\deg f_i - \nu_i < 0$. It follows that $g \neq 0$ if and only if for some index i we have $\deg f_i \geqslant \nu_i$. Since we have $T^{-1} f = g + T^{-1} r$ and $T^{-1} r$ is strictly proper, the row degrees of g are equal to the row degrees of $T^{-1} f$ which are $\deg f_i - \nu_i$. Hence $\deg g = max_{1 \leqslant i \leqslant m}(\deg f_i - \nu_i)$. $\square$

We have the following theorem due to Eckberg [3] and Hinrichsen and Prätzel-Wolters [29].

Theorem 6.1. *Every full column rank, $p \times l$ polynomial matrix $H(z)$ can be reduced to a unique modified Kronecker–Hermite canonical form by right multiplication by a unimodular matrix.*

The number of free parameters in the set determined by the integral parameters, namely the row index lists $R_i = \{\rho_1^{(i)}, \ldots, \rho_{k_i}^{(i)}\}$ and the degree lists ν_i, $i = 1, \ldots, s$ is given by

$$N = p(n+l) - \sum_{i=1}^{s} \sum_{j=1}^{k_i} (p - \rho_j^{(i)} + 1) - \sum_{i=1}^{s} \frac{(k_i - 1)k_i}{2} - \sum_{i=2}^{s} \sum_{j=1}^{i-1} \sum_{\mu=0}^{k_i} \omega_{j,\mu}^{(i)} [k_i(\nu_i - \nu_j + 1) - \mu]. \tag{141}$$

Proof. The basic idea of the proof is the reduction of the number of free parameters, using elementary column operations. The technical tool is a repeated application of the division rule for polynomials.

We will prove the theorem by induction. We shall show that in this case H is reducible, by elementary column operations, to the modified Kronecker–Hermite canonical form. To this end we assume without loss of generality that H is in column proper form with s distinct column indices $\nu_1 < \cdots < \nu_s$ and with k_i columns of degree ν_i. Thus $\sum_{i=1}^{s} k_i = l$ and $\sum_{i=1}^{s} k_i \nu_i = n$. Let us write

$$H(z) = (H_1(z); \ldots; H_s(z)),$$

H_i are $p \times k_i$ column proper polynomial matrices. Our proof goes by induction on the number s of distinct minimal column degrees.

If $s = 1$, all column degrees are equal to ν_1 and we can reduce the highest coefficient column matrix to reverse echelon form, using multiplication on the right by constant elementary matrices. The resulting matrix is obviously in the modified Kronecker–Hermite canonical form and this form is uniquely determined. We assume the pivot elements to be in the rows $i_1^{(1)} < \cdots < i_{k_1}^{(1)}$.

The proof will be by induction on the number of distinct column indices. So our induction hypothesis is that the $p \times (k_1 + \cdots + k_{i-1})$ matrix $(H^{(1)}(z); \ldots; H^{(i-1)}(z))$ is in modified Kronecker–Hermite form. In particular the $p \times k_j$ matrices $H^{(j)}$, $j = 1, \ldots, i-1$, are column proper with the columns of $H^{(j)}$ of degree ν_j and the leading column coefficient matrix in reverse echelon form. Furthermore, we assume that $\deg h_\sigma^{(j)} < \nu_m$ for $\sigma \in \{i_1^{(m)}, \ldots, i_{k_m}^{(m)}\}$.

We consider now the $p \times k_i$ matrix $H^{(i)}$. $(H^{(1)}(z); \ldots; H^{(i-1)}(z))$, the $k_1 + \cdots + k_{i-1}$ submatrix of $H(z)$ which is based on the rows $\bigcup_{m=1}^{i-1}\{i_1^{(m)}, \ldots, i_{k_m}^{(m)}\}$ is, by the induction hypothesis, nonsingular, column proper as well as row proper with its leading row coefficient matrix a permutation matrix.

We reduce now, using Lemma 6.2, the $(k_1 + \cdots + k_{i-1}) \times k_i$ submatrix of $H^{(i)}$, which is based on the rows $\bigcup_{m=1}^{i-1}\{i_1^{(m)}, \ldots, i_{k_m}^{(m)}\}$ of $(H^{(1)}(z); \ldots; H^{(i-1)}(z))$ with respect to the $(k_1 + \cdots + k_{i-1}) \times (k_1 + \cdots + k_{i-1})$ submatrix of $(H^{(1)}(z); \ldots; H^{(i-1)}(z))$ which is based on the rows $\bigcup_{m=1}^{i-1}\{i_1^{(m)}, \ldots, i_{k_m}^{(m)}\}$. We retain the notation $H^{(i)}$ for the reduced matrix and we note that this reduction does not increase the degrees of the columns. The reduction process did not affect the columns of the submatrices $H^{(j)}$, $j \neq i$. Thus it follows that $H^{(i)}$ is still column proper with column degrees ν_i. Moreover the elements in the rows $i_1^{(j)}, \ldots, i_{k_j}^{(j)}$, for $j = 1, \ldots, i-1$, have degrees $< \nu_j$. So the degree ν_i elements occur in the complementary rows. By constant elementary column operations on the columns of $H^{(i)}$ we reduce the leading column coefficient matrix to reverse echelon form. These elementary operations keep the degrees of the elements in rows $i_1^{(j)}, \ldots, i_{k_j}^{(j)}$, for $j = 1, \ldots, i-1$, below ν_j. On the other hand the elements of $H^{(1)}(z), \ldots, H^{(i-1)}(z))$ in rows $i_1^{(i)}, \ldots, i_{k_i}^{(i)}$, have degrees less than ν_i. Thus we have reduced $(H^{(1)}(z); \ldots; H^{(i)}(z))$ to canonical form and this holds also for $i = s$.

We note that, by construction, the $l \times l$ submatrix based on the rows $i_1^{(j)}, \ldots, i_{k_j}^{(j)}$, $j = 1, \ldots, s$, is both row and column proper and its determinant has degree ν.

We proceed now to count the number of free continuous parameters in the canonical form. Obviously, the canonical form is completely determined by the integral parameters $R_i = \{\rho_1^{(i)}, \ldots, \rho_{k_i}^{(i)}\}$, ν_i, $i = 1, \ldots, s$, where the sets of row indices R_i are disjoint and the degrees ν_i are distinct. Note that we have $\sum_{i=1}^{s} k_i = l$.

We count first the number of free parameters in $H^{(i)}$, disregarding the degree reduction due to the constraints (137) and (138). Consider first the jth column, $j = 1, \ldots, k_i$, of the submatrix $H^{(i)}$. It has a monic, degree ν_i polynomial in row $\rho_j^{(i)}$. This polynomial has ν_i free parameters as do all polynomials in the lower rows of the same column. The polynomials in the rows above it have degree ν_i and therefore have $\nu_i + 1$ free parameters. Altogether the jth column has

$$(\nu_i + 1)(\rho_j^{(i)} - 1) + \nu_i(p - \rho_j^{(i)} + 1) = (\nu_i + 1)p - (p - \rho_j^{(i)} + 1)$$

free parameters. Adding over all columns of $H^{(i)}$ and over all submatrices, we get

$$\begin{aligned}
&\sum_{i=1}^{s}\sum_{j=1}^{k_i}[(\nu_i+1)p-(p-\rho_j^{(i)}+1)]\\
&\quad=\sum_{i=1}^{s}k_i(\nu_i+1)p-\sum_{i=1}^{s}\sum_{j=1}^{k_i}(p-\rho_j^{(i)}+1)\\
&\quad=p(n+l)-\sum_{i=1}^{s}\sum_{j=1}^{k_i}(p-\rho_j^{(i)}+1).
\end{aligned}$$

From this number we have to subtract the number of parameters cancelled by the constraints (137) and (138). We treat the constraint (137) first.

In this case, each pivot element reduces the degrees of the terms to its right by 1. Thus the total number of parameters in $H^{(i)}$ reduced due to this constraint is

$$(k_i-1)+(k_i-2)+\cdots+1+0=\frac{(k_i-1)k_i}{2}.$$

Summing up over all indices $i=1,\ldots,s$, we get

$$\sum_{i=1}^{s}\frac{(k_i-1)k_i}{2}.$$

Clearly, the constraint (138) affects the entries of $H^{(i)}$ only by the submatrices to its left. Now the pivot row indices of $H^{(i)}$ are

$$\rho_1^{(i)}<\cdots<\rho_{k_i}^{(i)}$$

whereas the pivot row indices of $H^{(j)}$, $1\leqslant j<i$, are

$$\rho_1^{(j)}<\cdots<\rho_{k_j}^{(j)}$$

and the sets of indices are disjoint. The relative position of the second set with respect to the first determines the extra reduction in degrees. This is a combinatorial problem and for this purpose we need to count the number $\omega_{j,\mu}^{(i)}$ of $\rho_t^{(j)}$ which satisfy

$$\rho_\mu^{(i)}<\rho_t^{(j)}<\rho_{\mu+1}^{(i)},\quad \mu=0,\ldots,k_i. \tag{142}$$

To account for the boundary cases, we take $\rho_0^{(i)}=0$ and $\rho_{k_i+1}^{(i)}=p+1$, i.e. in the extremal cases there is only one nontrivial inequality to satisfy.

Taking, for $k\in\mathbf{Z}$, $k^+=(k+|k|)/2$, we clearly have that the number of $\rho_t^{(j)}$ satisfying (142) is given by

$$\omega_{j,\mu}^{(i)}=\sum_{t=1}^{k_j}\frac{\operatorname{sign}(\rho_{\mu+1}^{(i)}-\rho_t^{(j)})^+ + \sum_{t=1}^{k_j}\operatorname{sign}(\rho_t^{(j)}-\rho_\mu^{(i)})^+}{2}. \tag{143}$$

For each index $\rho_t^{(j)}$ satisfying (142), the extra reduction in the number of parameters of $H^{(i)}$ is

$$\mu(\nu_i - \nu_j) + (k_i - \mu)(\nu_i - \nu_j + 1) = k_i(\nu_i - \nu_j + 1) - \mu.$$

So, the total number of parameters of $H^{(i)}$ that is reduced by satisfying (142) is

$$\sum_{j=1}^{i-1} \sum_{\mu=0}^{k_i} \omega_{j,\mu}^{(i)} [k_i(\nu_i - \nu_j + 1) - \mu].$$

Finally, the number of free continuous parameters in H is given by (141).

To show uniqueness of the Kronecker–Hermite canonical form, assume that H, H' are two column equivalent matrices, both in canonical form. Clearly, by the definition, they have the same row indices $i_1, \ldots, i_l$ and column degrees $\delta_1, \ldots, \delta_l$. Let $\overline{H}, \overline{H}'$ be the two submatrices consisting of the $i_1, \ldots, i_l$, rows of H. By column equivalence, there exists a unimodular polynomial matrix such that $H = H'U$ and hence also $\overline{H} = \overline{H}'U$. Since the row degrees are the same, we can write $\overline{H} = \Delta\Gamma$ and $\overline{H}' = \Delta\Gamma'$, where $\Delta(z) = \operatorname{diag}(z^{\delta_1}, \ldots, z^{\delta_l})$. Here Γ and Γ' are biproper with leading term equal to I_l. This implies that $(\Gamma')^{-1}\Gamma = U$ and hence both terms are constant matrices, necessarily equal to I_l. Thus $H = H'$. $\square$

Given a rational function $g = p/q$, we define as usual $\deg g = \deg p - \deg q$. With this definition we can extend the definition of the Kronecker–Hermite canonical form to all rational matrix functions. It is in this form that we shall use it in Theorem 7.1.

Example 6.1. To check formula (141), we consider the following example: We let

p	l	s	ν_1	ν_2	k_1	k_2	$\rho_1^{(1)}$	$\rho_2^{(1)}$	$\rho_1^{(2)}$	$\rho_2^{(2)}$	$\rho_3^{(2)}$
11	5	2	2	5	2	3	2	8	4	7	10

The following matrix shows the degrees of the polynomials. The underlined degrees indicate monic polynomials. A polynomial of degree k has obviously $k + 1$ free parameters unless it is monic when the number is reduced by 1. The next matrix shows the degrees of the polynomials in the corresponding modified Kronecker–Hermite canonical form.

$$\begin{pmatrix} 2 & 2 & 5 & 5 & 5 \\ \boxed{2} & 1 & 1 & 1 & 1 \\ 1 & 2 & 5 & 5 & 5 \\ 1 & 2 & \boxed{5} & 4 & 4 \\ 1 & 2 & 4 & 5 & 5 \\ 1 & 2 & 4 & 5 & 5 \\ 1 & 2 & 4 & \boxed{5} & 4 \\ 1 & \boxed{2} & 1 & 1 & 1 \\ 1 & 1 & 4 & 4 & 5 \\ 1 & 1 & 4 & 4 & \boxed{5} \\ 1 & 1 & 4 & 4 & 4 \end{pmatrix}.$$

A direct count of free parameters gives the number 209. To apply (141), we compute

$$\omega_{1,1}^{(2)} = 1, \quad \omega_{1,2}^{(2)} = 0, \quad \omega_{1,3}^{(2)} = 1, \quad \omega_{1,4}^{(2)} = 0.$$

Now

$$p(n+l) = 11(19+5) = 264,$$

$$\begin{aligned}\sum_{i=1}^{s}\sum_{j=1}^{k_i}(p-\rho_j^{(i)}+1) \\ &= [(11-2+1)+(11-8+1)] \\ &\quad + [(11-4+1)+(11-7+1)+(11-10+1)] = 29,\end{aligned}$$

$$\sum_{i=1}^{s}\frac{(k_i-1)k_i}{2} = 1+3 = 4,$$

$$\begin{aligned}\sum_{i=2}^{s}\sum_{j=1}^{i-1}\sum_{\mu=0}^{k_i}\omega_{j,\mu}^{(i)}[k_i(\nu_i-\nu_j+1)-\mu] \\ &= 1[3(3+1)-0]+1[3(3+1)-2] = 22.\end{aligned}$$

This leads to

$$N = 264-29-4-22 = 209.$$

Given a reachable pair (A, B) we consider the coprime factorization $(zI-A)^{-1}B = H(z)T(z)^{-1}$. The polynomial matrix T encodes all the information of the pair (A, B) up to similarity. Thus if we reduce T to Kronecker–Hermite form by elementary column operations, by taking matrix representations of the shift realization with respect to the canonical basis, we get a canonical form for the given pair. This is described next.

Theorem 6.2. *Let $T(z) \in F^{p\times p}[z]$ be nonsingular in Kronecker–Hermite canonical form with, unordered, row degrees $\nu_1, \ldots, \nu_p$. Let the columns of T be given by*

$$t_j(z) = \begin{pmatrix} t_{ij}(z) \\ \cdot \\ \cdot \\ \cdot \\ t_{pj}(z) \end{pmatrix} \tag{144}$$

with $t_{ij}(z) = \sum_{k=0}^{\nu_j} t_{ij,k}z^k$. Then, with respect to the canonical basis of X_T we have the following block matrix representations for the pair $(S_T, \pi_T\cdot)$, namely $A = (A_{ij})$, $A_{ij} \in F^{\nu_i\times\nu_j}$, where the block A_{ij} is void if either ν_i or ν_j are equal to 0 *and*

$$A_{ii} = \begin{pmatrix} 0 & & & & -t_{ii,0} \\ 1 & & & & \cdot \\ & \cdot & & & \cdot \\ & & \cdot & & \cdot \\ 0 & & & 1 & -t_{ii,\nu_i-1} \end{pmatrix}$$

$$A_{ij} = \begin{pmatrix} 0 & \cdot & \cdot & 0 & -t_{ij,0} \\ 0 & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & 0 & -t_{ij,\nu_i-1} \end{pmatrix}$$

$B = (B_{ij})$, *with* $B_{ij} \in F^{\nu_i \times 1}$, *where for* $\nu_j > 0$ *we have*

$$B_{ij} = \begin{cases} \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, & \nu_j \neq 0,\ i \neq j \ \ or \ \ \nu_j = 0,\ i > j, \\ \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, & \nu_j \neq 0, \ \ i = j, \\ \begin{pmatrix} -t_{ij,0} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, & \nu_j = 0, \ \ i < j. \end{cases}$$

Proof. By Proposition 4.1, and the fact that $T(z)$ is row proper with row degrees $\nu_1, \ldots, \nu_p$, the standard basis of X_T is given by (68). Note that whenever $\nu_i = 0$ all elements of the ith row are 0. We enumerate this basis as $e_1^{(1)}, \ldots, e_{\nu_1}^{(1)}, \ldots, e_1^{(p)}, \ldots, e_{\nu_p}^{(p)}$. Assuming $\nu_i > 0$ and considering degrees, we have

$$S_T e_j^{(i)} = \begin{cases} e_{j+1}^{(i)}, & j = 1, \ldots, \nu_i - 1, \\ -\sum_{j=1}^{p} \sum_{k=1}^{\nu_j} t_{ij,k} e_k^{(j)}, & j = \nu_i. \end{cases}$$

Similarly, if $\eta_1, \ldots, \eta_p$ is the standard basis of F^p, then

$$B\eta_j = \pi_T \eta_j = \begin{cases} e_1^{(j)}, & \nu_j > 0, \\ -\sum_{i=1}^{j-1} t_{ij,0} e_1^{(i)}, & \nu_j = 0. \end{cases}$$

This leads to the above matrix representations. □

To illustrate this we consider the following examples.

Example 6.2. Let $T(z) \in F^{3\times 3}[z]$ be in Kronecker–Hermite canonical form with row degrees $(3, 0, 2)$. Thus

$$T(z) = \begin{pmatrix} z^3 + \alpha_2 z^2 + \alpha_1 z + \alpha_0 & \beta_0 & \gamma_2 z^2 + \gamma_1 z + \gamma_0 \\ 0 & 1 & 0 \\ \tau_1 z + \tau_0 & 0 & z^2 + \rho_1 z + \rho_0 \end{pmatrix}. \tag{145}$$

We have $\dim X_T = 5$ and with respect to the canonical basis

$$\begin{pmatrix} 1 & z & z^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & z \end{pmatrix}$$

we have the matrix representations

$$A = \left(\begin{array}{ccc|cc} 0 & 0 & -\alpha_0 & 0 & -\gamma_0 \\ 1 & 0 & -\alpha_1 & 0 & -\gamma_1 \\ 0 & 1 & -\alpha_2 & 0 & -\gamma_2 \\ \hline 0 & 0 & -\tau_0 & 0 & -\rho_0 \\ 0 & 0 & -\tau_1 & 1 & -\rho_1 \end{array}\right), \qquad B = \left(\begin{array}{c|c|c} 1 & -\beta_0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right).$$

Example 6.3. Again we assume $T(z) \in F^{3\times 3}[z]$ to be in Kronecker–Hermite canonical form with row degrees $(3, 2, 0)$. Thus

$$T(z) = \begin{pmatrix} z^3 + \alpha_2 z^2 + \alpha_1 z + \alpha_0 & \beta_2 z^2 + \beta_1 z + \beta_0 & \gamma_0 \\ \delta_1 z + \delta_0 & z^2 + \epsilon_1 z + \epsilon_0 & \eta_0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{146}$$

We have $\dim X_T = 5$ and with respect to the canonical basis

$$\begin{pmatrix} 1 & z & z^2 & 0 & 0 \\ 0 & 0 & 0 & 1 & z \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

we have the matrix representations

$$A = \left(\begin{array}{ccc|cc} 0 & 0 & -\alpha_0 & 0 & -\beta_0 \\ 1 & 0 & -\alpha_1 & 0 & -\beta_1 \\ 0 & 1 & -\alpha_2 & 0 & -\beta_2 \\ \hline 0 & 0 & -\delta_0 & 0 & -\epsilon_0 \\ 0 & 0 & -\delta_1 & 1 & -\epsilon_1 \end{array}\right), \qquad B = \left(\begin{array}{c|c|c} 1 & 0 & -\gamma_0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 1 & -\eta_0 \\ 0 & 0 & 0 \end{array}\right).$$

7. On the parametrization of conditioned invariant subspaces

We now proceed, given an observable pair $(\mathscr{C}, \mathscr{A})$, to parametrize the set of all conditioned invariant subspaces. This problem has been solved first in [23]. Unfor-

tunately, this pathbreaking paper did not get the attention it deserves and for a long time the study of this problem was effectively abandoned. A recurrence of interest in this parametrization problem occurred in [16] and in the papers arising out of Puerta's Thesis [32], namely [5,6].

In [23] a module theoretic approach was used based on the characterization of conditioned invariant subspaces obtained in [11,12], but put in the context of rational models. Our approach is inspired by that of Hinrichsen, Münzner and Prätzel-Wolters as well as the thesis of Eckberg [3]. We give a full proof of the reduction of a polynomial matrix to (modified) Kronecker–Hermite canonical form which, we believe, is more transparent than those given in [3,27]. The addition is a new count of the number of free parameters in the Kronecker–Hermite canonical form.

The first instinct might be, given a conditioned invariant subspace $\mathscr{V} \subset X_D$, to choose a basis for $\langle \mathscr{V} \rangle$, the submodule of $F^p[z]$ generated by it and reduce it to some, say column proper, canonical form. This turns out to be an unfruitful road. The main reason for this is the fact that in such a reduction the basis elements may not stay in X_D. Moreover, the corresponding column indices have nothing to do with D, so cannot have a system theoretic significance. The clue to overcoming this difficulty is provided by Lemma 3.1, Proposition 3.6 and Theorem 3.2. This shows that the significant object should be $D^{-1}T$.

In order not to make the notation overly cumbersome, we assume that $\mathscr{C}$ has full row rank. This is equivalent to the positivity of all of the observability indices, i.e.

$$\mu_1 \geqslant \cdots \geqslant \mu_p > 0. \tag{147}$$

Since the lattice of conditioned invariant subspace is invariant, up to isomorphism, under the full output injection group, it entails no loss of generality to assume that $(\mathscr{C}, \mathscr{A})$ is in dual Brunovsky form, i.e. that it corresponds to the polynomial matrix

$$D(z) = \begin{pmatrix} z^{\mu_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z^{\mu_p} \end{pmatrix}. \tag{148}$$

In order to gain some intuition, we consider the simplest case, namely when the submodule $\langle \mathscr{V} \rangle$ has a single generator, say $h \in \mathscr{V}$. This generator is uniquely determined up to a nonzero scalar factor. We can use this freedom to make one of the polynomials h_i monic. Now, there are two clear invariants for the subspace $\mathscr{V}$. The first one is $d = \dim \mathscr{V}$ which we know, by Theorem 3.2, is equal to the negative of the only one left Wiener–Hopf factorization index δ of $g = D^{-1}h$. Since $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$, we have $-\delta = \max_k \deg z^{-\mu_k} h_k$. So there exists at least one row index for which this equality holds and we assume i is the largest such index. So with each singly generated conditioned invariant subspace, there exist two, uniquely determined, integers δ, i defined as above. The set of all conditioned invariant subspaces having those invariants is denoted by $\mathscr{M}(\delta, i)$. Each singly generated

conditioned invariant subspace is contained in exactly one of these sets. Clearly, $g = D^{-1}h \in \mathscr{M}(\delta, i)$ if and only if, with

$$g = \begin{pmatrix} g_1 \\ \cdot \\ \cdot \\ \cdot \\ g_p \end{pmatrix},$$

we have

$$g_k(z) = \begin{cases} \sum_{j=0}^{\mu_k-\delta} h_{kj} z^{-\mu_k+j}, & k < i, \\ \sum_{j=0}^{\mu_i-\delta} h_{ij} z^{-\mu_i+j}, & k = i h_{ii} \text{ is monic}, \\ \sum_{j=0}^{\mu_k-\delta-1} h_{kj} z^{-\mu_k+j}, & k > i \text{ and } \mu_k > \delta, \\ 0, & k > i \text{ and } \mu_k \leqslant \delta. \end{cases} \tag{149}$$

The number of free parameters we need to parametrize $\mathscr{M}(\delta, i)$ is the number, $N(\delta, i)$, of coefficients, not counting h_{ii}. This is given by

$$\begin{aligned} N(\delta, i) &= \sum_{k=1}^{i-1} (\mu_k - \delta + 1) + (\mu_i - \delta) + \sum_{k>i} (\mu_k - \delta)^+ \\ &= \sum_{k=1}^{p} (\mu_k - \delta)^+ + (i - 1). \end{aligned} \tag{150}$$

Let G be any generator matrix. Two generating matrices for $\langle \mathscr{W} \rangle$ differ by a right unimodular factor. Let ν be any integer for which $z^\nu G = K$ is a polynomial matrix. Recall that D is in Brunovsky form (148) and hence such an integer exists. One should note that, starting from a generating matrix H for the submodule $\langle \mathscr{V} \rangle$, we have $K = H$, for some integer ν, if and only if all the left Wiener–Hopf factorization indices are equal. This corresponds to our observation in Section 3 that the main object of interest is not the generator H of $\langle \mathscr{V} \rangle$, but rather the generator $G = D^{-1}H$ of $\langle \mathscr{W} \rangle = \langle D^{-1}\mathscr{V} \rangle$. Clearly, if U is any unimodular polynomial matrix such that GU is column or row proper, the same holds for K. So, without loss of generality, we can look for a canonical form, under right unimodular equivalence, for rectangular polynomial matrices.

We can reformulate now the central result of Hinrichsen et al. [28], in the following way. For the present purpose the use of the Kronecker–Hermite canonical form is preferable to that of the modified one and it is in this form that we state the result.

We saw in Proposition 3.1 that any conditioned invariant subspace of X_D has a, not necessarily unique, representation in the form $\mathscr{V} = X_D \cap TF^p[z]$ with T nonsingular. Each such T defines, via the shift realization and choice of basis in X_T, a similarity equivalence class of reachable pairs, as in Theorem 2.1. Our aim is, in the nontight case, to parametrize all such pairs. Considering the scalar case as

a guide, corresponding to $d(z) = z^n$, the only nontight conditioned invariant subspace is $\mathscr{V} = \{0\}$. Clearly, we have the reresentation $\{0\} = X_d \cap tF[z]$ if and only if $\deg t \geqslant n$. For our purposes it will suffice to restrict the parametrization to the case $\deg t = n = \deg d$. The other cases can be treated similarly. In this case, taking the polynomial t to be monic, we have n free parameters, namely the coefficients of t. This set of choices provides maximal submodules for which the representation holds. At the same time this gives us a parametrization of the set of reachable pairs of minimal McMillan degree for which the kernel representation (88) holds. We solve the general parametrization problem this problem by looking at embeddings of G in nonsingular proper rational matrices of maximal degree. This is done in the following.

Theorem 7.1. *Given*

$$D(z) = \begin{pmatrix} z^{\mu_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z^{\mu_p} \end{pmatrix}, \tag{151}$$

with $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$ *and* $\sum_{i=1}^{p} \mu_i = n$, *let the pair* $(\mathscr{C}_D, \mathscr{A}_D)$ *be defined by*

$$\begin{aligned} \mathscr{A}_D f &= S_D f, \quad f \in X_D, \\ \mathscr{C}_D f &= (D^{-1} f)_{-1}, \quad f \in X_D. \end{aligned} \tag{152}$$

Then:

1. *There exists a bijective correspondence between the set of conditioned invariant subspaces of* $(\mathscr{C}_D, \mathscr{A}_D)$ *of codimension* k *and the disjoint union of all sets* $\mathscr{M}(\rho, \lambda)$ *of all* $p \times l$ *strictly proper matrices* G *for which* $H = DG \in F^{p\times l}[z]$ *and*, *with*

$$G(z) = (g_1(z), \ldots, g_l(z)), \tag{153}$$

are in Kronecker–Hermite canonical form, *corresponding to the two sets of indices*, *one of row indices* $\{1 \leqslant \rho_1 < \cdots < \rho_l \leqslant p\}$, *the other of reduced observability indices* $\{\lambda_1, \ldots, \lambda_l\}$ *satisfying* $0 < \lambda_i \leqslant \mu_i$ *and* $\sum_{i=1}^{l} \lambda_i = n - k$, *with the* g_i *being* $p \times 1$ *polynomial vectors in* z^{-1} *of degree* $-\lambda_i$, $i = 1, \ldots, l$. *Moreover*, *denoting the column elements of* g_j *by* g_{ij}, *they satisfy*:

   (a) $g_{\rho_j j}$ *is monic of degree* $-\lambda_j$.

   (b) *We have*

$$\deg g_{qj} \begin{cases} < \deg g_{\rho_j j}, & q > \rho_j, \\ \leqslant \deg g_{\rho_j j}, & q < \rho_j. \end{cases} \tag{154}$$

   (c) *For all* $\mu \neq j$, *we have*

$$\deg g_{\rho_j, \mu} < \deg g_{\rho_j, j}. \tag{155}$$

2. *Defining*
$$H(z) = D(z)G(z) = (h_1(z), \dots, h_l(z)) \in F^{p\times l}[z], \tag{156}$$
the $p \times (n-k) = p \times \sum_{i=1}^{l} \lambda_i$ *polynomial matrix*
$$E(z) = (h_1, \dots, z^{\lambda_1 - 1}h_1, \dots, h_l, \dots, z^{\lambda_l - 1}h_l) \tag{157}$$
is a linear basis matrix for $\mathscr{V} = X_D \cap HF^l[z]$, *i.e.*
$$\mathscr{V} = \operatorname{Im} E(z).$$
We call this the image representation *of* $\mathscr{V}$ *and call this the* canonical basis *of* $\mathscr{V}$. *We interpret* $E(z)$ *as a bijective multiplication map from* F^{n-k} *onto* $\mathscr{V} = X_D \cap H(z)F^l[z]$.
3. *There exist embeddings of* G *in a* $p \times p$ *nonsingular proper rational matrices* Γ, *so that the* jth *column of* G *appears as the* ρ_jth *column of* Γ *which is assumed in Kronecker–Hermite canonical form with* $\Gamma = (\gamma_1, \dots, \gamma_p)$, *where the additional columns of* Γ *have degree* 0 *and for which*
$$T(z) = D(z)\Gamma(z) \in F^{p\times p}[z] \tag{158}$$
is nonsingular, *D-proper*, *satisfies* $\deg\det T = k$ *and* $\mathscr{V} = X_D \cap TF^p[z]$.
4. *Defining*
$$\tau_i = \deg t_{ii}, \tag{159}$$
and two sets of indices in $\{1, \dots, p\}$
$$\begin{aligned} I_1 &= \{j \mid \mu_j > \tau_j\} = \{\rho_1, \dots, \rho_l\}, \\ I_2 &= \{j \mid \mu_j = \tau_j\}, \end{aligned} \tag{160}$$
the columns of H *are embedded in* T *as the set of columns with indices in* I_1. *For the other columns we define* $\lambda_i = 0$. *We denote the columns of* T *by* t_i, $i = 1, \dots, p$.
5. *Define maps* $\mathscr{A} : X_D \longrightarrow X_D$ *by* $\mathscr{A} = S_D$ *and* $\mathscr{B} : F^l \longrightarrow X_D$ *by*
$$\mathscr{B}\eta = \pi_D H\eta = H\eta. \tag{161}$$
Then, *with* $\lambda = (\lambda_1, \dots, \lambda_l)$, *we have*
$$E(z) = \mathscr{R}_\lambda(\mathscr{A}, \mathscr{B}). \tag{162}$$
6. *Let*
$$\mathbf{A} = [S_D]_{\mathrm{st}}^{\mathrm{st}} = \begin{pmatrix} J_{\mu_1} & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & J_{\mu_p} \end{pmatrix}_{n\times n}, \tag{163}$$
where
$$J_{\mu_i} = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & 0 \\ 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 1 & 0 \end{pmatrix}_{\mu_i\times\mu_i} \tag{164}$$

and

$$\mathbf{B} = [H]_{\mathrm{st}}^{\mathrm{st}}. \tag{165}$$

Considering the polynomial matrix $E(z)$ of (157) *as a multiplication map from F^{n-k} into $\mathscr{V}$, then its matrix representation with respect to the standard bases of F^{n-k} and X_D is*

$$Z = [E(z)]_{\mathrm{st}}^{\mathrm{st}} = [\mathscr{R}_\lambda(\mathscr{A}, \mathscr{B})]_{\mathrm{st}}^{\mathrm{st}} = \mathscr{R}_\lambda(\mathbf{A}, \mathbf{B}). \tag{166}$$

The $n \times (n-k)$ matrix $\mathscr{R}_\lambda(\mathbf{A}, \mathbf{B})$ has the property that its $(n-k) \times (n-k)$ submatrix based on the rows indexed by

$$\bigcup_{i \in I_1} \left\{ \sum_{\iota=0}^{i-1} \mu_\iota + \mu_i - \lambda_i + 1, \ldots, \sum_{\iota=0}^{i} \mu_\iota \right\} \tag{167}$$

is a nonsingular matrix.
The cardinality of the index set in (167) *is $\sum_{i=1}^{l} \lambda_i = n - k$.*

7. *Let T be defined by* (158), *then with (A, B) defined by* (72) *and $R_\mu(A, B)$ by* (73), *we have*

$$R_\mu(A, B)\mathscr{R}_\lambda(\mathbf{A}, \mathbf{B}) = 0. \tag{168}$$

The $k \times n$ partial reachability matrix $R_\mu(A, B)$ is canonical in the sense that its $k \times k$ submatrix based on the columns indexed by the set

$$\bigcup_{i \in I_1} \left\{ \left(\sum_{\iota=0}^{i-1} \mu_\iota \right) + 1, \ldots, \left(\sum_{\iota=0}^{i-1} \mu_\iota \right) + \mu_i - \lambda_i \right\}, \tag{169}$$

is the identity matrix.
The cardinality of the index set in (169) *is $\sum_{i=1}^{l} (\mu_i - \lambda_i) = n - (n-k) = k$.*

8. *The index sets in* (167) *and* (169) *are complementary with respect to $\{1, \ldots, n\}$.*
9. *Assume T is D-proper, then the following diagram is exact.*

$$\begin{tikzcd}
0 \arrow[r] & F^{n-k} \arrow[r, "H"] \arrow[d, "I"'] & X_D \arrow[r, "\pi_T|X_D"] \arrow[d, "{[\cdot]^{st}}"'] & X_T \arrow[r] \arrow[d, "{[\cdot]^{st}}"'] & 0 \\
0 \arrow[r] & F^{n-k} \arrow[r, "\mathcal{R}_\lambda(\mathbf{A},\mathbf{B})"] & F^n \arrow[r, "R_\mu(A,B)"] & F^k \arrow[r] & 0
\end{tikzcd}$$

10. *With (A, B) defined by* (72) *let $\Phi(z)$ be defined by*

$$(zI - A)^{-1}B = \Phi(z)T(z)^{-1}, \tag{170}$$

then $\tilde{\Phi}$ is the observer basis matrix for $X^{\tilde{T}}$. The rows of Φ are obtained by shifting down the rows of T and keeping only the such obtained nonzero rows.

Proof. (1) Assume we have the conditioned invariant subspace $\mathscr{V} = X_D \cap HF^l[z]$, where we assume, as in Theorem 3.2, that H is a $p \times l$ basis matrix for the free module $\langle \mathscr{V} \rangle$, the smallest submodule of $F^p[z]$ containing $\mathscr{V}$, and furthermore whose columns are in $\mathscr{V}$. We define $G = D^{-1}H$ and it is clearly strictly proper. By the structure of D, the only singularity of G is at $z = 0$. Thus there exists an integer ν for which $z^\nu G$ is a polynomial matrix and hence, by Theorem 6.1, is reducible by elementary column operations to Kronecker–Hermite form. Remultiplying by $z^{-\nu}$ we get a rational function in Kronecker–Hermite form which we still denote by G. This reduction is obviously independent of ν. Assume the left Wiener–Hopf factorization indices of G are $-\lambda_1, \ldots, -\lambda_l$.

Adapting the notation used in Definition 6.2, we can write

$$G = (g_1, \ldots, g_l),$$

where the g_j are strictly proper column vectors of degree $-\lambda_j$, $j = 1, \ldots, l$. We denote the g_j elements by g_{ij} and, by our assumption that G is in Kronecker–Hermite form, they satisfy (1a)–(1c).

(2) The rational vectors g_i have degrees $-\lambda_i$ and hence the $z^\nu g_i$, for $\nu = 0, \lambda_i - 1$ are strictly proper and hence in $D^{-1}\mathscr{V}$. As they are clearly linearly independent, using the fact that codim $\mathscr{V} = k$, and that the columns of $H(z)$ in (157) are $\sum_{i=1}^s \lambda_i = n - k$ in number, it follows that the columns of $H(z)$ form a basis matrix for $\mathscr{V} = X_D \cap HF^l[z]$.

(3) The embedding as described is clearly possible. Clearly $f \in \mathscr{V}$ if and only if $f \in X_D \cap HF^l[z]$ or, equivalently, if and only if $h = D^{-1}f \in X^D \cap D^{-1}HF^l[z] = X^D \cap GF^l[z]$. We show now that

$$X^D \cap GF^l[z] = X^D \cap \Gamma F^p[z]. \tag{171}$$

The inclusion $X^D \cap GF^l[z] \subset X^D \cap \Gamma F^p[z]$ is obvious. Assume therefore that $h \in X^D \cap \Gamma F^l[z]$ and write $h = \sum_{j=1}^p f_j \gamma_j$ with $f_j \in F[z]$. Since Γ is column proper, the degree preserving property holds, i.e. we have $\deg h = \max(\deg \Gamma^{(j)} + \deg f_j)$. Hence, for h to be strictly proper we must have $f_j = 0$ for j not in $\{\rho_1, \ldots, \rho_l\}$. Thus we have obtained the reverse inclusion $X^D \cap \Gamma F^p[z] \subset X^D \cap GF^l[z]$ and hence the equality (171). In turn this implies the equality

$$X_D \cap HF^l[z] = X_D \cap TF^p[z]. \tag{172}$$

By construction, T is D-proper and hence, see the remark after Definition 3.1, we have codim $\mathscr{V} = \deg \det T$.

(4) We consider the jth column vector

$$h_j(z) = \begin{pmatrix} h_{1j}(z) \\ \cdot \\ \cdot \\ \cdot \\ h_{pj}(z) \end{pmatrix}.$$

We assume the pivot polynomial to be in row ρ_j, so $h_{\rho_j j}$ is monic of degree $\mu_{\rho_j} - \lambda_j$ and

$$\deg h_{ij} \begin{cases} < \mu_i - \lambda_j, & i > \rho_j, \\ \leqslant \mu_i - \lambda_j, & i < \rho_j. \end{cases}$$

We change notation and, as in the embedding construction, consider the column vectors h_j, $j = 1, \ldots, l$, as embedded in the polynomial matrix T, where the pivot elements are on the diagonal and the embedded columns are distinguished by the set of indices I_1. Obviously we have

$$\tau_j = \deg t_{jj} = \mu_j - \lambda_j, \quad j \in I_1.$$

(5) Since the coefficient of z^{-1} in the expansion of $D(z)^{-1} z^{\nu} H_i$ is 0 for $\nu = 0, \ldots, \lambda_i - 1$, it follows that $S_D z^{\nu} H_i = z^{\nu+1} H_i$. By the definitions of $\mathscr{B}$ and the partial reachability map in (52), representation (162) follows.

(6) For $i \in I_1$, the coefficient of $z^{\mu_i - \lambda_i}$ in t_{ii} is 1 and will appear in $[T_i]^{\mathrm{st}}$ in the $\sum_{\iota=0}^{i-1} \mu_\iota + \mu_i - \lambda_i + 1$ position. Since we can multiply the polynomial vector T_i by $1, z, \ldots, z^{\lambda_j - 1}$ and stay in $\mathscr{V}$, we have the corresponding 1 appear in the rows $\{\sum_{\iota=0}^{i-1} \mu_\iota + \mu_i - \lambda_i + 1, \ldots, \sum_{\iota=0}^{i-1} \mu_\iota + \mu_i = \sum_{\iota=0}^{i} \mu_\iota\}$. The number of elements in this set is λ_i. Taking the union over the set I_1 we get (167) which has $\sum_{i \in I_1} \lambda_i = \dim \mathscr{V} = n - k$ elements.

Clearly, the matrix representation of $(h_j, zh_j, \ldots, z^{\lambda_j - 1} h_j)$ with respect to the standard basis of X_D is an $n \times \lambda_j$ matrix having a block Toeplitz structure obtained by shifting down the coefficients. The same holds for the matrix $Z = [E(z)]_{\mathrm{st}}^{\mathrm{ca}}$.

We show now that the $(n-k) \times (n-k)$ submatrix based on the rows indexed by (167) is a nonsingular matrix. To this end we consider the basis of X_D constructed as follows. With the e_i the standard basis elements in F^p we consider

$$\bigcup_{i=1}^{p} \{e_i, ze_i, \ldots, z^{\mu_i - \lambda_i - 1} e_i, t_i, zt_i, \ldots, z^{\lambda_i - 1} t_i\}. \tag{173}$$

We interpret $\{t_i, zt_i, \ldots, z^{\lambda_i - 1} t_i\} = \emptyset$ when $\lambda_i = 0$. To see that the set of vectors in (173) is a basis of X_D we note that by Lemma 3.2, we have $X_D = X_T \oplus T X_S = X_T \oplus \mathscr{V}$, the direct sum being in the sense of linear subspaces. T is row proper with row degrees $\mu_i - \lambda_i$ hence

$$\bigcup_{i=1}^{p} \{e_i, ze_i, \ldots, z^{\mu_i - \lambda_i - 1} e_i\}$$

is a basis for X_T. On the other hand

$$\bigcup_{i=1}^{p} \{t_i, zt_i, \ldots, z^{\lambda_i - 1} t_i\}$$

is a basis for $\mathscr{V}$, so the union of both sets, i.e. the set in (173), is a basis for X_D.

Now the dual basis to the standard basis in X_D is the Brunovsky basis (64) in $X^D = X^{\tilde{D}}$. Any vector in $t \in X_D$ can be written as

$$t(z) = \sum_{i=1}^{p} t_i(z) e_i = \sum_{i=1}^{p} \sum_{k=1}^{\mu_i} t_{ik} z^{k-1} e_i.$$

Since

$$\left[\sum_{i=1}^{p} \sum_{k=1}^{\mu_i} t_{ik} z^{k-1} e_i, z^{-\nu} e_j \right] = h_{j\nu},$$

it follows that the linear functional determined by $z^{-\nu} e_j$ reads off the $\sum_{\iota=0}^{i-1} \mu_\iota + \nu$, $1 \leqslant \nu \leqslant \mu_j$ element of $[h]^{\mathrm{st}}$. The columns of $\mathscr{R}_\lambda(\mathscr{A}, \mathscr{B})$ are in X_D and applying this functional to all the columns of $\mathscr{R}_\lambda(\mathscr{A}, \mathscr{B})$ reads off the $\sum_{\iota=0}^{i-1} \mu_\iota + \nu$ row of $\mathscr{R}_\lambda(\mathbf{A}, \mathbf{B})$. We restrict ourselves now to the set of functionals

$$\{e_j z^{-\nu-1} \mid \mu_j - \lambda_j \leqslant \nu \leqslant \mu_j - 1\}.$$

This set clearly annihilates $\bigcup_i \{z^\alpha e_i \mid 0 \leqslant \alpha \leqslant \mu_i - \lambda_i - 1\}$, for

$$[z^\alpha e_i, z^{-\nu-1} e_j] = \delta_{ij} \delta_{\alpha\nu} = 0$$

as α, ν vary in disjoint sets of indices. This implies that the set of rows of $\mathscr{R}_\lambda(\mathbf{A}, \mathbf{B})$ indexed by (167) is necessarily nonsingular.

Now $\mathscr{R}_\lambda(\mathbf{A}, \mathbf{B})$ is an $n \times (n-k)$, full column rank matrix for which the submatrix based on the rows indexed by (167) is nonsingular. Hence it can be written as the kernel of a $k \times n$ full row rank matrix R and can fix R uniquely if we fix appropriately the elements in the k columns indexed by the set (169) which is complementary to the $n-k$ row indices listed in (167). By appropriately we mean that the $k \times k$ submatrix of R based on these columns is the identity.

(7) Next we will show that the canonical kernel matrix R is nothing else but the partial reachability matrix $R_\mu(A, B)$ introduced in (73).

Note that from the fact that $\mathscr{V} = \operatorname{Im} H = \operatorname{Ker} \pi_T \mid X_D$, it follows that $(\pi_T \mid X_D)$ $H = 0$, and, by taking matrix representations, we get

$$[\pi_T \mid X_D]_{\mathrm{st}}^{\mathrm{st}} [H]_{\mathrm{st}}^{\mathrm{st}} = R_\mu(A, B) \mathscr{R}_\lambda(\mathbf{A}, \mathbf{B}) = 0.$$

It only remains to show that the $k \times k$ submatrix of $R_\mu(A, B)$ consisting of the columns indexed by (169) is the identity.

Let e_i, $i = 1, \ldots, p$, be the unit vectors in F^p. In the standard basis matrix

$$\begin{pmatrix} 1 & z & \cdot & \cdot & z^{\mu_1 - 1} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & z & \cdot & \cdot & z^{\mu_p - 1} \end{pmatrix} \tag{174}$$

of X_D they appear in columns $1, \mu_1 + 1, \ldots, \mu_1 + \cdots + \mu_{p-1} + 1$. We note that $Be_i = e_i$ if and only if $e_i \in X_T$ which is equivalent to $\deg t_{ii} = \mu_i - \lambda_i > 0$. Obviously, we have in this case that $e_i, ze_i, \ldots, z^{\mu_i - \lambda_i - 1} e_i \in X_T$. Thus the set $\bigcup_{i \in I_1} \{(\sum_{j=0}^{i-1} \mu_j) + 1, \ldots, (\sum_{j=0}^{i-1} \mu_j) + \mu_i - \lambda_i\}$ has $\sum_i (\mu_i - \lambda_i) = n - (n-k)$

$= k$ elements and they parametrize the set of k columns of $R_\mu(A, B)$ for which the corresponding $k \times k$ submatrix is the identity.

(8) The sets of indices in (167) and (169) are clearly disjoint and contain $n - k$ and k, respectively, which shows the complementarity.

(9) Exactness follows from the fact that $\operatorname{Ker} \pi_T \mid X_D = X_D \cap TF^p[z]$, the injectivity of the basis matrix $H(z)$, the surjectivity of $\pi_T \mid X_D$ proved in Theorem 3.3 as well as the matrix representations (73) and (166).

(10) Follows from Proposition 4.1. □

Some remarks are in order.

The real beauty of the previous theorem is that any D-proper polynomial matrix T not only determines a unique conditioned invariant subspace, of codimension $\deg \det T$, but at the same time parametrizes all its minimal McMillan degree kernel representations.

It is clear from Theorem 7.1.7 that the canonical kernel representations for a conditioned invariant subspace $\mathscr{V}$ can be determined in two distinct ways. One by choosing a canonical, but parameter dependent, basis for the solution of a system of linear equations, and the other by computing the realization of T by the pair (A, B) and forming $R_\mu(A, B)$.

We note that while Γ is both row and column proper, T as defined in (158) is row proper but not necessarily column proper. Obviously, in the generic case, when the μ_i are all equal, T is also column proper.

The block Toeplitz structure of the matrix Z has been previously proved in [5]. In this connection see also [17].

Actually we do not have to do our analysis of the parametrization of conditioned invariant subspaces by resorting to the use of rational matrix functions, and can do it wholly in the polynomial domain. As pointed out in the introduction to this section, the naive approach of reducing the polynomial matrix T in the representation $\mathscr{V} = X_D \cap TF^p[z]$ to canonical Kronecker–Hermite form is not sufficient since it does not incorporate the information that couples D and T. To overcome this difficulty, we introduce a new canonical form which is D-dependent.

Definition 7.1.

$$D(z) = \begin{pmatrix} z^{\mu_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z^{\mu_p} \end{pmatrix}, \tag{175}$$

with $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$ and let $T \in F^{p \times l}[z]$ be of full column rank. We will say that T in *D-Kronecker–Hermite form* if:

1. T is row proper with row degrees $\tau_i \leqslant \mu_i$ for $i = 1, \ldots, p$.
2. The highest row coefficient matrix is in column echelon form with the pivot elements in rows $\rho_1, \ldots, \rho_l$. Equivalently, we have $t_{\rho_j j}$ is monic and
$$\tau_j = \deg t_{\rho_j j} > \deg t_{\rho_j k} \quad \text{for } k \neq j. \tag{176}$$
3. We have
$$\deg t_{kj} \begin{cases} \leqslant & \deg t_{\rho_j j} + \mu_k - \mu_{\rho_j}, \quad k \leqslant \rho_j, \\ < & \deg t_{\rho_j j} + \mu_k - \mu_{\rho_j}, \quad k > \rho_j. \end{cases} \tag{177}$$

The degree constraints of the previous definition have alternative representations which are given as follows. The new representation allows us to write down formulas for the number of free parameters, i.e. the topological dimension, of both the set $\mathscr{M}(\rho, \lambda)$ as well as the number of free parameters in the kernel representations $\mathbf{V} = \ker R_\mu(A, B)$ given in (88).

Theorem 7.2. *Let $D(z) = \operatorname{diag}(z^{\mu_1}, \ldots, z^{\mu_p})$ and let $T \in F^{p\times l}[z]$. Then*:
1. *T is in D-*Kronecker–Hermite form *if and only if the following degree constraints hold.*
If $k \notin \{\rho_1, \ldots, \rho_l\}$, then
$$\deg t_{kj} < \begin{cases} \deg t_{\rho_j j} + \mu_k - \mu_{\rho_j} + 1, & k < \rho_j, \\ \deg t_{\rho_j j} + \mu_k - \mu_{\rho_j}, & k \geqslant \rho_j \end{cases} \tag{178}$$
and, with $k = \rho_i$,
$$\deg t_{\rho_i j} < \begin{cases} \min\{\deg t_{\rho_i i}, \deg t_{\rho_j j} + \mu_{\rho_i} - \mu_{\rho_j} + 1\}, & \rho_i < \rho_j, \\ \min\{\deg t_{\rho_i i}, \deg t_{\rho_j j} + \mu_{\rho_i} - \mu_{\rho_j}\}, & \rho_i \geqslant \rho_j. \end{cases} \tag{179}$$
2. *If T is square and nonsingular the degree constraints simplify to*
$$\deg t_{ij} < \begin{cases} \min\{\deg t_{ii}, \deg t_{jj} + \mu_i - \mu_j + 1\}, & i < j, \\ \min\{\deg t_{ii}, \deg t_{jj} + \mu_i - \mu_j\}, & i \geqslant j. \end{cases} \tag{180}$$
3. *Assuming T is in D-Kronecker–Hermite form, then the number of free parameters in the parametrization of $\mathscr{V} = X_D \cup TF^p[z]$ is*
$$N = \sum_{j \in I_1} \sum_{\deg t_{ij} \geqslant 0} (\deg t_{ij} + 1) - \#(I_1). \tag{181}$$
4. *Assuming T is in D-Kronecker–Hermite form, then the number of free parameters in the parametrization of the kernel representation is*
$$P = \sum_{j \in I_2} \sum_{i=1}^{p} (\deg t_{ij} + 1) - \#(I_2). \tag{182}$$

Proof. The statement follows by putting the degree constraints (176) and (177) together. □

Definition 7.1 is tailored to the following obvious statement, the proof of which is omitted.

Theorem 7.3. *Given*

$$D(z) = \begin{pmatrix} z^{\mu_1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z^{\mu_p} \end{pmatrix}, \tag{183}$$

with $\mu_1 \geqslant \cdots \geqslant \mu_p > 0$ *and* $T \in F^{p\times l}[z]$ *of full column rank, then*:

1. *T is in D-Kronecker–Hermite form if and only if $D^{-1}T$ is in Kronecker–Hermite form.*
2. *Every D-proper, full column rank polynomial matrix $T \in F^{p\times l}[z]$ can be reduced to D-Kronecker–Hermite form by elementary column operations.*

Let us consider a few examples.

Example 7.1. Assume

$$D(z) = \begin{pmatrix} z & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & z \end{pmatrix} \in F[z]^{p\times p}.$$

This corresponds, in state space terms, to the Brunovsky form $(\mathscr{C}, \mathscr{A}) = (I, 0) \in F^{p\times p} \times F^{p\times p}$. In this case it is obvious that every subspace is not only conditioned invariant but actually invariant. Let $\mathscr{V}$ be a one-dimensional subspace. $\mathscr{V}$ is generated by one vector $h \in X_D$ which is necessarily constant. Let $1 \leqslant i \leqslant p$ be the largest index for which $h_i \neq 0$ and we use our freedom to normalize it to 1. Thus $\mathscr{V}$ is spanned by

$$h^{(i)} = \begin{pmatrix} \alpha_{1i} \\ \cdot \\ \cdot \\ \alpha_{i-1i} \\ 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

So the set of all one-dimensional (conditioned invariant) subspaces is therefore parametrized by

$$h^{(1)} = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, h^{(k)} = \begin{pmatrix} \alpha_{1k} \\ \cdot \\ \cdot \\ \alpha_{k-1k} \\ 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, h^{(p)} = \begin{pmatrix} \alpha_{1p} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{p-1p} \\ 1 \end{pmatrix}$$

We look at the case $\rho = i, \lambda = 1$ in somewhat more detail. Clearly,

$$g(z) = z^{-1}h^{(i)} = \begin{pmatrix} \alpha_{1i}/z \\ \cdot \\ \cdot \\ \alpha_{i-1i}/z \\ 1/z \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

The number of free parameters is $i - 1$ and the canonical $(p-1) \times p$ kernel matrix is

$$\begin{pmatrix} 1 & & & & & -\alpha_{1i} & & & & \\ & \cdot & & & & \cdot & & & & \\ & & \cdot & & & \cdot & & & & \\ & & & \cdot & & \cdot & & & & \\ & & & & 1 & -\alpha_{(i-1)i} & & & & \\ & & & & & 0 & 1 & & & \\ & & & & & & & \cdot & & \\ & & & & & & & & \cdot & \\ & & & & & & & & & \cdot \\ & & & & & & & & & 1 \end{pmatrix}.$$

The embedding matrix is given by

$$\Gamma(z) = \begin{pmatrix} 1 + \frac{\alpha_{11}}{z} & \cdot & \cdot & \cdot & \frac{\alpha_{1i}}{z} & \cdot & \cdot & \cdot & \frac{\alpha_{1p}}{z} \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & \frac{1}{z} & 0 & 0 & 0 & 0 \\ & & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & & & \cdot & \cdot & \cdot \\ & & & & & & & \cdot & \cdot \\ & & & & & & & & \cdot \\ & & & & & & & & 1 + \frac{\alpha_{pp}}{z} \end{pmatrix}$$

This leads to

$$T(z) = \begin{pmatrix} z+\alpha_{11} & \cdot & \cdot & \cdot & \alpha_{1i} & \cdot & \cdot & \cdot & \alpha_{1p} \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & 1 & 0 & 0 & 0 & 0 \\ & & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & & & \cdot & \cdot & \cdot \\ & & & & & & & \cdot & \cdot \\ & & & & & & & & \cdot \\ & & & & & & & & z+\alpha_{pp} \end{pmatrix}$$

and

$$\Phi(z) = \begin{pmatrix} 1 & & & & & & & & \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cdot & & & & & \\ & & & & 1 & 0 & & & \\ & & & & & 0 & 1 & & \\ & & & & & & & \cdot & \\ & & & & & & & & \cdot \\ & & & & & & & & & \cdot \\ & & & & & & & & & 1 \end{pmatrix}$$

The standard $p \times p-1$ basis matrix for X_T is given by

$$\begin{pmatrix} 1 & & & & & & & & \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cdot & & & & & \\ & & & & 1 & & & & \\ & & & & & 0 & & & \\ & & & & & 1 & & & \\ & & & & & & \cdot & & \\ & & & & & & & \cdot & \\ & & & & & & & & \cdot \\ & & & & & & & & & 1 \end{pmatrix}$$

Since T is row proper, it is easy to compute the matrix representations of the shift realization $(S_T, \pi_T \cdot)$ with respect to the standard basis. In fact, we have

$$A = \begin{pmatrix} -\alpha_{11} & \cdot & \cdot & -\alpha_{1(i-1)} & -\alpha_{1(i+1)} & \cdot & -\alpha_{1p} \\ & -\alpha_{22} & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot \\ & & & & & \cdot & \cdot \\ & & & & & & -\alpha_{pp} \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & & & & -\alpha_{1i} & & & & \\ & \cdot & & & \cdot & & & & \\ & & \cdot & & \cdot & & & & \\ & & & \cdot & \cdot & & & & \\ & & & 1 & -\alpha_{(i-1)i} & & & & \\ & & & & 0 & 1 & & & \\ & & & & & & \cdot & & \\ & & & & & & & \cdot & \\ & & & & & & & & \cdot \\ & & & & & & & & 1 \end{pmatrix}$$

In this case we have also $R_\mu(A, B) = B$. P the number of free parameters in the choice of the similarity class (A, B) is easily computed to be $P = (p(p+1)/2) - p = p(p-1)/2$, which is actually independent of i.

We analyze now from our point of view some cases of the main example in [28].

Example 7.2. We assume $\mu_1 = 3$, $\mu_2 = 2$, $\mu_3 = 1$.

Consider the case that the conditioned invariant subspace $\mathscr{V}$ is determined by the row index list $\rho = (1, 3)$ and the corresponding degree index list $\lambda = (3, 1)$. In particular we get $\dim \mathscr{V} = 4$. The generating matrix in Kronecker–Hermite canonical form is given by

$$G(z) = \begin{pmatrix} \frac{1}{z^3} & 0 \\ 0 & \frac{\eta_1}{z} + \frac{\eta_0}{z^2} \\ 0 & \frac{1}{z} \end{pmatrix}.$$

To get the polynomial generating matrix, we multiply by $D(z) = \operatorname{diag}(z^3, z^2, z)$ to get

$$H(z) = \begin{pmatrix} 1 & 0 \\ 0 & \eta_1 z + \eta_0 \\ 0 & 1 \end{pmatrix}.$$

In particular, the number of free parameters is $N = 2$. The state space representation of the generating matrix, with respect to the standard basis, is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \eta_0 \\ 0 & \eta_1 \\ 0 & 1 \end{pmatrix}.$$

More interesting for us is the basis matrix of $\mathscr{V}$ given polynomially by

$$\begin{pmatrix} 1 & z & z^2 & 0 \\ 0 & 0 & 0 & \eta_1 z + \eta_0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Taking its matrix representation with respect to the standard basis, we get

$$Z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \eta_0 \\ 0 & 0 & 0 & \eta_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

A direct computation of the kernel matrix in canonical form gives

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & -\eta_0 \\ 0 & 0 & 0 & 0 & 1 & -\eta_1 \end{pmatrix}.$$

Next, we embed $G(z)$ in a nonsingular canonical proper matrix taking the extra row degrees to be 0. Thus we get

$$\Gamma(z) = \begin{pmatrix} \frac{1}{z^3} & 0 & 0 \\ 0 & 1 + \frac{\epsilon_1}{z} + \frac{\epsilon_0}{z^2} & \frac{\eta_1}{z} + \frac{\eta_0}{z^2} \\ 0 & 0 & \frac{1}{z} \end{pmatrix}.$$

Multiplying $\Gamma(z)$ by $D(z) = \operatorname{diag}(z^3, z^2, z)$ we get

$$T(z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & z^2 + \epsilon_1 z + \epsilon_0 & \eta_1 z + \eta_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Clearly $\deg\det T(z) = 2$, so $\dim X_T = 2$. The standard basis matrix for X_T is given by

$$\begin{pmatrix} 0 & 0 \\ 1 & z \\ 0 & 0 \end{pmatrix}.$$

This leads to the matrix representations

$$A = \begin{pmatrix} 0 & -\epsilon_0 \\ 1 & -\epsilon_1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 & -\eta_0 \\ 0 & 0 & -\eta_1 \end{pmatrix}.$$

From the polynomial matrix $T(z)$ we get immediately, by shifting down rows and retaining the nonzero ones, that

$$\Phi = \begin{pmatrix} 0 & z+\epsilon_1 & \eta_1 \\ 0 & 1 & 0 \end{pmatrix},$$

and so the observer basis matrix for $X_{\tilde{T}}$ is given by

$$\tilde{\Phi} = \begin{pmatrix} 0 & 0 \\ z+\epsilon_1 & 1 \\ \eta_1 & 0 \end{pmatrix}.$$

Finally, the free parameters in the parametrization of the similarity class of (A, B) are ϵ_0, ϵ_1, so $P = 2$. We put this information in Table 1.

Example 7.3. We assume $\mu_1 = 3, \mu_2 = 2, \mu_3 = 1$.

Here we consider the case that the conditioned invariant subspace $\mathscr{V}$ is determined by the row index list $\rho = (1, 2, 3)$ and the corresponding degree index list $\lambda = (1, 1, 1)$. In particular we get $\dim \mathscr{V} = 3$. The generating matrix in Kronecker–Hermite canonical form is given by

$$G(z) = \begin{pmatrix} \frac{1}{z} + \frac{\alpha_1}{z^2} + \frac{\alpha_0}{z^3} & \frac{\beta_1}{z^2} + \frac{\beta_0}{z^3} & \frac{\gamma_1}{z^2} + \frac{\gamma_0}{z^3} \\ \frac{\delta_0}{z^2} & \frac{1}{z} + \frac{\epsilon_0}{z^2} & \frac{\eta_0}{z^2} \\ 0 & 0 & \frac{1}{z} \end{pmatrix}.$$

In particular, the number of free parameters is $N = 9$. To get the polynomial generating matrix, we multiply by $D(z) = \operatorname{diag}(z^3, z^2, z)$ to get

$$H(z) = \begin{pmatrix} z^2 + \alpha_1 z + \alpha_0 & \beta_0 + \beta_1 z & \gamma_1 z + \gamma_0 \\ \delta_0 & z + \epsilon_0 & \eta_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In this case we have $\Gamma(z) = G(z)$ and $T(z) = H(z)$.

The state space representation of the generating matrix, with respect to the standard basis, is given by

$$\begin{pmatrix} \alpha_0 & \beta_0 & \gamma_0 \\ \alpha_1 & \beta_1 & \gamma_1 \\ 1 & 0 & 0 \\ \delta_0 & \epsilon_0 & \eta_0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The basis matrix of $\mathscr{V}$, given polynomially, is equal in this case to $T(z)$ and for its matrix representation with respect to the standard basis, we get

Table 1

$\dim \mathscr{V}$	4
ρ	(1,3)
λ	(3,1)
$G(z)$	$\begin{pmatrix} \frac{1}{z^3} & 0 \\ 0 & \frac{\eta_1}{z} + \frac{\eta_0}{z^2} \\ 0 & \frac{1}{z} \end{pmatrix}$
$H(z)$	$\begin{pmatrix} 1 & 0 \\ 0 & \eta_1 z + \eta_0 \\ 0 & 1 \end{pmatrix}$
$\mathbf{B} = [H]_{\mathrm{st}}^{\mathrm{st}} =$ Generating matrix	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \eta_0 \\ 0 & \eta_1 \\ 0 & 1 \end{pmatrix}$
N	2
$\mathscr{V}$ Basis matrix $= \mathscr{R}_\lambda(\mathbf{A}, \mathbf{B})$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \eta_0 \\ 0 & 0 & 0 & \eta_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
Kernel matrix $= R_\mu(A, B)$	$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & -\eta_0 \\ 0 & 0 & 0 & 0 & 1 & -\eta_1 \end{pmatrix}$
$\Gamma(z)$	$\begin{pmatrix} \frac{1}{z^3} & 0 & 0 \\ 0 & 1 + \frac{\epsilon_1}{z} + \frac{\epsilon_0}{z^2} & \frac{\eta_1}{z} + \frac{\eta_0}{z^2} \\ 0 & 0 & 1 \end{pmatrix}$
$T(z)$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & z^2 + \epsilon_1 z + \epsilon_0 & \eta_1 z + \eta_0 \\ 0 & 0 & 1 \end{pmatrix}$
X_T Basis matrix standard basis	$\begin{pmatrix} 0 & 0 \\ 1 & z \\ 0 & 0 \end{pmatrix}$
$A = [S_T]_{\mathrm{st}}^{\mathrm{st}}$	$\begin{pmatrix} 0 & -\epsilon_0 \\ 1 & -\epsilon_1 \end{pmatrix}$
$B = [\pi_T \cdot]_{\mathrm{st}}^{\mathrm{st}}$	$\begin{pmatrix} 0 & 1 & -\eta_0 \\ 0 & 0 & -\eta_1 \end{pmatrix}$
$\Phi(z)$	$\begin{pmatrix} 0 & z + \epsilon_1 & \eta_1 \\ 0 & 1 & 0 \end{pmatrix}$
$X_{\tilde{T}}$ Basis matrix $= \tilde{\Phi}$ Observer basis	$\begin{pmatrix} 0 & 0 \\ z + \epsilon_1 & 1 \\ \eta_1 & 0 \end{pmatrix}$
P	2

Table 2

$\dim \mathscr{V}$	3
ρ	(1,2,3)
λ	(1,1,1)
$G(z)$	$\begin{pmatrix} \frac{1}{z}+\frac{\alpha_1}{z^2}+\frac{\alpha_0}{z^3} & \frac{\beta_1}{z^2}+\frac{\beta_0}{z^3} & \frac{\gamma_1}{z^2}+\frac{\gamma_0}{z^3} \\ \frac{\delta_0}{z^2} & \frac{1}{z}+\frac{\epsilon_0}{z^2} & \frac{\eta_0}{z^2} \\ 0 & 0 & \frac{1}{z} \end{pmatrix}$
$H(z)$	$\begin{pmatrix} z^2+\alpha_1 z+\alpha_0 & \beta_0+\beta_1 z & \gamma_1 z+\gamma_0 \\ \delta_0 & z+\epsilon_0 & \eta_0 \\ 0 & 0 & 1 \end{pmatrix}$
$\mathbf{B}=[H]^{\mathrm{st}}_{\mathrm{st}}=$ Generating matrix	$\begin{pmatrix} \alpha_0 & \beta_0 & \gamma_0 \\ \alpha_1 & \beta_1 & \gamma_1 \\ 1 & 0 & 0 \\ \delta_0 & \epsilon_0 & \eta_0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
N	9
$\mathscr{V}$ Basis matrix $=\mathscr{R}_\lambda(\mathbf{A},\mathbf{B})$	$\begin{pmatrix} \alpha_0 & \beta_0 & \gamma_0 \\ \alpha_1 & \beta_1 & \gamma_1 \\ 1 & 0 & 0 \\ \delta_0 & \epsilon_0 & \eta_0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Kernel matrix $=R_\mu(A,B)$	$\begin{pmatrix} 1 & 0 & -\alpha_0 & 0 & -\beta_0 & -\gamma_0 \\ 0 & 1 & -\alpha_1 & 0 & -\beta_1 & -\gamma_1 \\ 0 & 0 & -\delta_0 & 1 & -\epsilon_0 & -\eta_0 \end{pmatrix}$
$\Gamma(z)$	$\begin{pmatrix} \frac{1}{z}+\frac{\alpha_1}{z^2}+\frac{\alpha_0}{z^3} & \frac{\beta_1}{z^2}+\frac{\beta_0}{z^3} & \frac{\gamma_1}{z^2}+\frac{\gamma_0}{z^3} \\ \frac{\delta_0}{z^2} & \frac{1}{z}+\frac{\epsilon_0}{z^2} & \frac{\eta_0}{z^2} \\ 0 & 0 & \frac{1}{z} \end{pmatrix}$
$T(z)$	$\begin{pmatrix} z^2+\alpha_1 z+\alpha_0 & \beta_0+\beta_1 z & \gamma_1 z+\gamma_0 \\ \delta_0 & z+\epsilon_0 & \eta_0 \\ 0 & 0 & 1 \end{pmatrix}$
X_T Basis matrix	$\begin{pmatrix} 1 & z & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$
A	$\begin{pmatrix} 0 & -\alpha_0 & -\beta_0 \\ 1 & -\alpha_1 & -\beta_1 \\ 0 & -\delta_0 & -\epsilon_0 \end{pmatrix}$
B	$\begin{pmatrix} 1 & 0 & -\gamma_0 \\ 0 & 0 & -\gamma_1 \\ 0 & 1 & -\eta_0 \end{pmatrix}$

Table 2 (Continued)

$\Phi(z)$	$\begin{pmatrix} z+\alpha_1 & \beta_1 & \gamma_1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$
$X_{\tilde{T}}$ Basis matrix $= \tilde{\Phi}$ Observer basis	$\begin{pmatrix} z+\alpha_1 & 1 & 0 \\ \beta_1 & 0 & 1 \\ \gamma_1 & 0 & 0 \end{pmatrix}$
P	0

$$Z = \begin{pmatrix} \alpha_0 & \beta_0 & \gamma_0 \\ \alpha_1 & \beta_1 & \gamma_1 \\ 1 & 0 & 0 \\ \delta_0 & \epsilon_0 & \eta_0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A direct computation of the kernel matrix in canonical form gives

$$\begin{pmatrix} 1 & 0 & -\alpha_0 & 0 & -\beta_0 & -\gamma_0 \\ 0 & 1 & -\alpha_1 & 0 & -\beta_1 & -\gamma_1 \\ 0 & 0 & -\delta_0 & 1 & -\epsilon_0 & -\eta_0 \end{pmatrix}.$$

Clearly $\deg\det T(z) = 3$, so $\dim X_T = 3$. The standard basis matrix for X_T is given by

$$\begin{pmatrix} 1 & z & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

This leads to the matrix representations

$$A = \begin{pmatrix} 0 & -\alpha_0 & -\beta_0 \\ 1 & -\alpha_1 & -\beta_1 \\ 0 & -\delta_0 & -\epsilon_0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & -\gamma_0 \\ 0 & 0 & -\gamma_1 \\ 0 & 1 & -\eta_0 \end{pmatrix}.$$

From the polynomial matrix $T(z)$ we get immediately, by shifting down rows and retaining the nonzero ones, that

$$\Phi = \begin{pmatrix} z+\alpha_1 & \beta_1 & \gamma_1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

and so the observer basis matrix for $X_{\tilde{T}}$ is given by

$$\tilde{\Phi} = \begin{pmatrix} z+\alpha_1 & 1 & 0 \\ \beta_1 & 0 & 1 \\ \gamma_1 & 0 & 0 \end{pmatrix}.$$

Finally, since we are in the tight case, T is uniquely determined and hence there are no free parameters in the parametrization of the similarity class of (A, B), i.e. $P = 0$. Again, we put this information in Table 2.

Acknowledgements

The first author would like to acknowledge the support, in the early stages of this work, by the Fundação Calouste Gulbenkian during a visit to CIM in Coimbra.

References

[1] A.C. Antoulas, New results on the algebraic theory of linear systems: the solution of the cover problem, Linear Algebra Appl. 50 (1983) 1–43.
[2] G. Basile, G. Marro, Controlled and conditioned invariant subspaces in linear system theory, J. Optim. Theory Appl. 3 (1973) 306–316.
[3] A.E. Eckberg, A characterization of linear systems via polynomial matrices and module theory, Ph.D. Thesis, MIT, Cambridge, MA, 1974.
[4] E. Emre, M.L.J. Hautus, A polynomialcharacterization of (A, B)-invariant and reachability subspaces, SIAM J. Control Optim. 18 (1980) 420–436.
[5] J. Ferrer, F. Puerta, X. Puerta, Differentiable structure of the set of controllable $(A, B)^{\mathrm{t}}$-invariant subspaces, Linear Algebra Appl. 275/276 (1998) 161–177.
[6] J. Ferrer, F. Puerta, X. Puerta, Stratification of the set of general (A, B)-invariant subspaces (1999).
[7] G.D. Forney, Minimal bases of rational vector spaces, with applications to multivariable linear systems, SIAM J. Control Optim. 13 (1973) 493–520.
[8] P.A. Fuhrmann, Algebraic system theory: ananalyst's point of view, J. Franklin Inst. 301 (1976) 521–540.
[9] P.A. Fuhrmann, On strict system equivalence and similarity, Int. J. Control 25 (1977) 5–10.
[10] P.A. Fuhrmann, Linear feedback via polynomial models, Int. J. Control 30 (1979) 363–377.
[11] P.A. Fuhrmann, Linear Operators and Systems in Hilbert Space, McGraw-Hill, New York, 1981.
[12] P.A. Fuhrmann, Duality in polynomial models with some applications to geometric control theory, IEEE Trans. Automat. Control AC26 (1981) 284–295.
[13] P.A. Fuhrmann, On symmetric rational matrix functions, Linear Control Theory (special issue), Linear Algebra Appl. 50 (1983) 167–250.
[14] P.A. Fuhrmann, Geometric control revisited, 2001 in preparation.
[15] P.A. Fuhrmann, U. Helmke, A homeomorphism between observable pairs and conditioned invariant subspaces, Systems Control Lett. 30 (1997) 217–223.
[16] P.A. Fuhrmann, U. Helmke, On the parametrization of conditioned invariant subspaces, Borkum Proceedings, 1999.
[17] P.A. Fuhrmann, A. Gombani, On a Hardy space approach to the analysis of spectral factors, Int. J. Control 71 (1998) 277–357.
[18] P.A. Fuhrmann, J.C. Willems, Factorization indices at infinity for rational matrix functions, Integral Eqn. Oper. Theory 2 (1979) 287–301.
[19] P.A. Fuhrmann, J.C. Willems, A study of (A, B)-invariant subspaces via polynomial models, Int. J. Control 31 (1980) 467–494.
[20] I.C. Gohberg, I.A. Feldman, Convolution Equations and Projection Methods for their Solution, Nauka, Moscow, 1971 (English translation: AMS, Providence, RI, 1974).
[21] M.L.J. Hautus, M. Heymann, Feedback—an algebraic approach, SIAM J. Control 16 (1978) 83–105.
[22] M.L.J. Hautus, E.D. Sontag, An approach to detectability and observers, in: C.I. Byrnes, C.F. Martin (Eds.), Algebraic and Geometric Methods in Linear Systems Theory, AMS, Providence, RI, 1980, pp. 99–135.

[23] U. Helmke, The Cohomology of Moduli Spaces of Linear Dynamical Systems, Regensburger Mathematische Schriften 24, Department of Mathematics, 1993.
[24] D. Hinrichsen, D. Prätzel-Wolters, Generalized Hermite matrices and complete invariants of strict system equivalence, SIAM J. Control Optim. 21 (1983) 289–305.
[25] D. Hinrichsen, H.F. Münzner, D. Prätzel-Wolters, Parametrization of (C, A)-invariant subspaces, Systems Control Lett. 1 (1981) 192–199.
[26] S. Kawaji, Design procedures of observer for linear function of the state, Int. J. Control 32 (1980) 381–394.
[27] J.J. Loiseau, Structural necessary conditions for model matching, in: Proceedings of ECC, 1991, pp. 703–708.
[28] D.G. Luenberger, Observers for multivariable systems, IEEE Trans. Automat. Control AC11 (1966) 190–197.
[29] X. Puerta, Contributions to the geometric study of multivariable linear systems, Ph.D. Thesis, Universitat Politècnica de Catalunya, Barcelona, 1998.
[30] J.R. Roman, T.E. Bullock, Design of minimal order stable observers for linear functions of the state via realization theory, IEEE Trans. Automat. Control AC 20 (1975) 613–622.
[31] J.C. Willems, Almost $A(\mathrm{mod}\, B)$-invariant subspaces, Asterisque 75/76 (1980) 239–248.
[32] H. Wimmer, A Jordan factorization theorem for polynomial matrices, IEEE Trans. Automat. Control AC 16 (1979) 621–634.
[33] W.A. Wolovich, Linear Multivariable Systems, Springer, New York, 1974.
[34] W.M. Wonham, Linear Multivariable Control, second ed., Springer, New York, 1979.