International Journal of Control
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713393989

System order and structure indices of linear systems in polynomial form

Paul Van Den Hof [a]
[a] Mechanical Engineering Systems and Control Group, Delft University of Technology, The Netherlands

Online Publication Date: 01 June 1992

To cite this Article Hof, Paul Van Den(1992)'System order and structure indices of linear systems in polynomial form',International Journal of Control,55:6,1471 — 1490

To link to this Article: DOI: 10.1080/00207179208934295

URL: http://dx.doi.org/10.1080/00207179208934295

INT. J. CONTROL, 1992, VOL. 55, NO. 6, 1471–1490

System order and structure indices of linear systems in polynomial form

PAUL VAN DEN HOF†

Linear time-invariant, finite-dimensional discrete time systems are very often specified in a polynomial form representation in either a forward or backward shift operator (sometimes called MFD and ARMA form). In this paper it is shown that there exists a unifying theory for the determination of McMillan degree and Kronecker observability indices of systems represented by polynomial matrices in the two shift operators, considering the MFD and ARMA forms as special cases. Treating dynamical systems in terms of their behaviour, i.e. the set of admissible signal trajectories, the notions of McMillan degree and observability indices can be generalized to non-controllable as well as to non-causal systems.

1. Introduction

Linear, time-invariant, finite-dimensional, discrete time systems can be represented in many different ways. Descriptions in terms of difference equations are quite common, formulating direct relations between the input and output variables of the system and their time shifts. For a multivariable system with input $u(t)$ and output $y(t)$, two different types of polynomial representations are frequently used:

MFD-form:

$$P(\sigma)y(t) = Q(\sigma)u(t) \quad \forall\, t \in \mathbb{Z} \tag{1}$$

with $y(t) \in \mathbb{R}^p$, $u(t) \in \mathbb{R}^m$, and P, Q polynomial matrices of size $(p \times p)$, $(p \times m)$ in one indeterminate, and σ the (forward) shift operator: $\sigma w(t) = w(t+1)$.

ARMA-form:

$$P^*(\sigma^{-1})y(t) = Q^*(\sigma^{-1})u(t) \quad \forall t \in \mathbb{Z} \tag{2}$$

with P^*, Q^* polynomial matrices of similar size as P, Q, and σ^{-1} the (backward) shift operator: $\sigma^{-1}w(t) = w(t-1)$.

The notions of matrix fraction description (MFD) and autoregressive moving average (ARMA) would actually be more consistently denoted by forward/backward difference equation forms. However, in order to keep up with most of the literature on this subject the notions of MFD- and ARMA-forms will be used throughout this paper. Both forms are frequently applied in problems of systems and control theory, while the choice between the two is more or less dependent on the type of problem that is discussed. In identification, and especially in prediction error identification, ARMA forms are popular because of their natural way of treating output predictions: the output signal $y(t)$ is written as a linear combination

Received 1 October 1990. Revised 15 May 1991.

† Mechanical Engineering Systems and Control Group, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.

of previous output and input signals $y(t-i)$, $u(t-j)$, $i>0$, $j\geqslant 0$; see e.g. Ljung (1987). In this respect the so-called monic ARMA forms are very suitable, having $P^*(\sigma^{-1}) = I + P_1\sigma^{-1} + \cdots + P_r\sigma^{-r}$, with $P_i \in \mathbb{R}^{p\times p}$, $i=1,\ldots,r$. In the area of control and in parametrization problems the MFD forms are applied more frequently. In parametrization issues, and especially in the construction of canonical forms, this is due to the fact that in MFD forms the McMillan degree and the (Kronecker) observability indices of systems can be simply related to determinantal degrees and row degrees of specific MFD forms (Guidorzi 1975, 1981), whereas for ARMA forms this is much more complex. These simple expressions for McMillan degree and observability indices create the possibility for construction of identifiable continuous parametrizations of model sets containing all models with a prespecified set of observability indices.

In terms of ARMA-forms, identifiable continuous parametrizations can be constructed showing the general property that neither McMillan degree nor observability indices are prescribed and consequently these structural properties will vary over the parametrized model set; see for example Deistler et al. (1978), Deistler (1983), Hannan and Kavalieris (1984). An alternative parametrization in which the McMillan degree is prescribed but the transfer function $P^{-1}Q$ of the models is restricted to have no poles in the origin ($z=0$), is discussed by Bokor and Keviczky (1987). It has been shown by Gevers (1986) that monic ARMA models will generically represent systems whose McMillan degree is a multiple of the output dimension p.

Obviously there exist straightforward relations between system representations in either MFD or ARMA form, which can be specified by considering the transfer functions $H(z) = P(z)^{-1}Q(z)$ and $H^*(z) = P^*(z^{-1})Q^*(z^{-1})$ as the basic system characteristics. Using the equality of transfer functions as a notion of system equivalence, transformations between MFD and ARMA forms have been shown, see, for example Wolovich and Elliot (1983). Gevers (1986) studied the McMillan degree and observability indices of ARMA models, showing that also for ARMA models expressions for these phenomena can be given, although resulting in more complex expressions than for MFD models. These results have been extended and generalized by Janssen (1988 *a*, *b*), where parallels between the situations of MFD and ARMA forms have been clearly indicated.

In this contribution it will be shown that there exists a unifying theory for the determination of McMillan degree and (Kronecker) observability indices of models in polynomial form, using polynomial representations that contain both σ and σ^{-1} as indeterminates. In these generalized forms, MFD and ARMA forms will reduce to special cases. This generalization is actually due to Willems (1986, 1988). In this paper the theory of model behaviour, as initiated by Willems (1986, 1988), will be applied and interpreted for dynamical systems having a prespecified set of input and output signals. Applying in this situation, the basic system characteristic of system behaviour as a more general notion for system equivalence than the transfer function, the concepts of McMillan degree and Kronecker observability indices will be generalized also to non-controllable as well as to non-causal systems. The generalized concepts will be denoted by the system order and observability indices of dynamical systems. For a detailed discussion on the concept of system behaviour the reader is referred to the papers of Willems. The use of the behavioural description of dynamical input–output systems in view of the system identification problem has also been discussed by Van den Hof (1989 *a*, *b*). This paper is an extended version of Van den Hof (1990).

We first introduce the notion of system behaviour as advocated by Willems (1986). In § 3 some additional definitions and notation related to polynomial matrices will be summarized. Subsequently the concepts of system order and structural indices will be discussed and related explicit expressions will be presented on the basis of a polynomial representation.

Some notational conventions: $\mathbb{R}^{p \times m}(z)$ is the field of $(p \times m)$ rational matrices; $\mathbb{R}[z]$ is the ring of polynomials in the indeterminate z; $\mathbb{R}[z^{-1}]$ is the ring of polynomials in the indeterminate z^{-1}; $\mathbb{R}^{p \times m}[z, z^{-1}]$ is the ring of $(p \times m)$ (bi)polynomial matrices in the indeterminates z and z^{-1}; $\det_{\mathbb{R}(z)}(\cdot)$ is the determinant over the field of rational functions; $\text{rank}_{\mathbb{C}}(\cdot)$ is the (ordinary) rank over the field of complex numbers; I_p is the $p \times p$ identity matrix; $\mathbb{Z}$ the set of (positive and negative) integers; $\mathbb{R}\backslash\{0\}$ denotes the set of real numbers excluding 0; $[P_1 \mid P_2]$ denotes the composite matrix, composed of P_1 and P_2 having the same number of rows.

2. System representations and system behaviour

There are many ways of representing linear, time-invariant, finite-dimensional discrete time systems. Apart from the MFD and ARMA forms as mentioned in the introduction, and apart from the transfer function representation, the state space form is also a commonly used description. Obviously these different representations are all representations of the same phenomena: dynamical systems. The transfer function of a dynamical system is very often referred to as its basic characteristic; i.e. two dynamical systems (irrespective of their representation) are equal if and only if they have the same transfer function. However, a basic definition of a dynamical system in terms of its transfer function shows the disadvantage that non-controllable modes in the system will simply not be recognized as part of the system, whereas the effect of non-controllable modes might explicitly appear in the signals that are extracted from the system. This is illustrated in the following example.

Example 2.1: Consider a dynamical single-input single-output (SISO) system described by the following difference equation:

$$(\sigma - a)(\sigma - c)y(t) = (\sigma - b)(\sigma - c)u(t) \quad \forall t \in \mathbb{Z} \tag{3}$$

with $a, b, c \in \mathbb{R}\backslash 0$.

Let $\{u(t)\}_{t=-\infty,\ldots,\infty}$ be given and let $\{\hat{y}(t)\}_{t=-\infty,\ldots,\infty}$ satisfy the difference equation (3); then any $\{y(t)\}_{t=-\infty,\ldots,\infty}$ that can be written as $y(t) = \hat{y}(t) + \tilde{y}(t)$, $t \in \mathbb{Z}$, with $(\sigma - a)(\sigma - c)\tilde{y}(t) = 0$ will satisfy the difference equation (3). Thus the signals that can be extracted from the system can be influenced by the coefficient c, whereas the transfer function of the system is $H(z) = (z - b)/(z - a)$ which is independent of c.

A generalized and very natural definition of a dynamical system is introduced by Willems (1986), employing the notion of system behaviour as the basic characteristic of a dynamical system. Consider a dynamical system with input $u(t)$ and output $y(t)$, $t \in \mathbb{Z}$; then the behaviour $\mathscr{B}$ is defined as the set of all signal trajectories $(y^{\mathrm{T}}\ u^{\mathrm{T}})^{\mathrm{T}}$: $\mathbb{Z} \to \mathbb{R}^{p+m}$ that are admissible, i.e. that satisfy the restrictions (equations) of the dynamical system. Note that a signal trajectory $\{(y^{\mathrm{T}}(t)\ u^{\mathrm{T}}(t))^{\mathrm{T}}, t = -\infty, \ldots, \infty\}$

is actually denoted as a mapping from $\mathbb{Z}$ to $\mathbb{R}^{p+m}$, also expressed by the notation $(y^{\mathrm{T}}\ u^{\mathrm{T}})^{\mathrm{T}} \in (\mathbb{R}^{p+m})^{\mathbb{Z}}$. The behaviour is defined on the level of signals and not by coefficient matrices in specific equations. Therefore, the system behaviour is representation independent. It will be illustrated that this behaviour can serve as a proper tool for describing dynamical systems in more detail than, for example, a transfer function representation, by proper dealing with the uncontrollable modes as discussed above.

Let us first consider the formal definition of a dynamical system according to Willems (1986).

Definition 2.2: A dynamical system S is defined as a *triple*: $S = (T, W, B)$, with $T \subset \mathbb{R}$ the time set, W the signal set, i.e. the space in which the (input and output) variables that are related to the system take on their values, and $\mathscr{B} \subset W^{T}$ the behaviour of the system, i.e. the space of all signal trajectories $w: T \to W$ that are compatible with the system. □

The behaviour $\mathscr{B}$ of a dynamical system S will generally be denoted by $\mathscr{B}(S)$. In this paper we will deal with discrete-time systems having a time set $T = \mathbb{Z}$. The restriction to linearity, time-invariance and finite dimensionality of the dynamical systems is established by the additional requirements that W is a vector space and that $\mathscr{B}$ is a linear subspace of W^{T} that is closed in the topology of pointwise convergence, satisfying the shift-invariance property: $w \in \mathscr{B}(S) \Leftrightarrow \sigma w \in \mathscr{B}(S) \Leftrightarrow \sigma^{-1} w \in \mathscr{B}(S)$. The basic Definition 2.2 does not distinguish between input and output components of a dynamical system but only refers to system variables (signals). When considering these system variables as inputs and outputs (causes and effects), a special structure is laid upon the dynamical system, referring to the intuitive appeal that is connected to these notions. The character of an input signal is that it can be chosen freely, it is not bound by the system; the character of an output signal is that it is caused by the input, the system dynamics and possibly a finite number of (initial) conditions. These notions have been formalized by Willems (1986, 1988). In this paper we consider systems with a fixed number of output and input signals, in such a way that the signals related to the system can be denoted by $w = (y, u)$, with $w \in (\mathbb{R}^{p+m})^{\mathbb{Z}}$, and consequently $\mathscr{B}(S) \subset (\mathbb{R}^{p+m})^{\mathbb{Z}}$. This class of input–output system will be denoted by $\Sigma_{p,m}$.

Since the time set ($T = \mathbb{Z}$) and the signal set ($W = \mathbb{R}^{p+m}$) are fixed within the class, two dynamical systems will be equal if and only if they have the same behaviour. In the following example we will briefly illustrate the concept of system behaviour.

Example 2.3: Consider a dynamical system S described by a polynomial representation

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = -b_0 u(t) - b_1 u(t-1) - b_2 u(t-2) \tag{4}$$

and $a_1, a_2, b_0, b_1, b_2 \in \mathbb{R}$. With $y(t) \in \mathbb{R}$, $u(t) \in \mathbb{R}$ it follows that $W = Y \times U$, $Y = U = \mathbb{R}$; $y, u \in \mathbb{R}^{\mathbb{Z}}$ and $w(t) = (y(t), u(t)) \in \mathbb{R}^2$, leading to $w \in (\mathbb{R}^2)^{\mathbb{Z}}$.

The behaviour $\mathscr{B}(S) \subset (\mathbb{R}^2)^{\mathbb{Z}}$ contains those trajectories $w = (y, u)$ that satisfy (4) for all $t \in \mathbb{Z}$; this can be written as

$$\mathscr{B}(S) := \{w \in (\mathbb{R}^2)^{\mathbb{Z}} \mid T(\sigma, \sigma^{-1}) w = 0,$$
$$\text{with} \quad T(\sigma, \sigma^{-1}) = [1 + a_1\sigma^{-1} + a_2\sigma^{-2} \mid b_0 + b_1\sigma^{-1} + b_2\sigma^{-2}]\} \tag{5}$$

□

Now let us consider the consequences of this concept of system behaviour for polynomial representations of systems. The results as presented in this section are due to Willems (1986, 1988).

Proposition 2.4: *For any system* $S \in \Sigma_{p,m}$ *there exists a polynomial matrix* $T = [P \mid -Q]$, *with* $P \in \mathbb{R}^{p \times p}[z, z^{-1}]$, $\det_{\mathbb{R}(z)} P \neq 0$, $Q \in \mathbb{R}^{p \times m}[z, z^{-1}]$, *such that* $\mathscr{B}(S) = \{(y, u) \mid P(\sigma, \sigma^{-1})y - Q(\sigma, \sigma^{-1})u = 0\}$.

The system S will be said to be induced by the polynomial matrix T, denoted by $S = M_p(T)$. Using polynomial matrices in two indeterminates, a generalization has been obtained for the MFD and ARMA forms. Note that the condition of P being invertible as a rational matrix is implied by the fact that y is an output variable, which means that there exists a rational transfer function between the input u and the output y. The system behaviour will be used for deciding whether two dynamical systems are equal or not.

Proposition 2.5: *Let* $S_1, S_2 \in \Sigma_{p,m}$ *be induced by polynomial matrices* T_1, T_2 *respectively, as in Proposition* 2.4. *Then* $\mathscr{B}(S_1) = \mathscr{B}(S_2)$ *if and only if* $T_1 = UT_2$ *with* U *a polynomial matrix that is unimodular over the ring* $\mathbb{R}[z, z^{-1}]$, *i.e.* $\det_{\mathbb{R}(z)} U = cz^d$ *with* $c \in \mathbb{R}\backslash\{0\}$ *and* $d \in \mathbb{Z}$.

With respect to the unimodularity of U we denote U as unimodular over $\mathbb{R}[z, z^{-1}]$ if its determinant is a unit in $\mathbb{R}[z, z^{-1}]$, i.e. the matrix U is invertible in $\mathbb{R}^{p \times p}[z, z^{-1}]$, which implies the condition on its determinant as formulated in the proposition. Note that in the situation of MFD or ARMA forms, where the rings $\mathbb{R}[z]$ or $\mathbb{R}[z^{-1}]$ are considered, the determinant of a unimodular matrix has to be a constant $c \in \mathbb{R}\backslash\{0\}$. As a result of this proposition it follows, for example, that premultiplication of each polynomial equation with powers of z and z^{-1} does not change the behaviour of the dynamical system, but only comes down to shifting the equations. Several well-known system properties can be described in terms of the system's polynomial representation.

(*a*) The system's transfer function is denoted by

$$H(z) = P(z, z^{-1})^{-1}Q(z, z^{-1}) \quad \text{with} \quad H(z) \in \mathbb{R}^{p \times m}(z)$$

(*b*) This transfer function describes the dynamical system completely if and only if the system is controllable, i.e. a corresponding polynomial representation $T = [P \mid -Q]$ has the property that P and Q are left coprime over $\mathbb{R}[z, z^{-1}]$ (all left common factors are unimodular over $\mathbb{R}[z, z^{-1}]$), or equivalently $\operatorname{rank}_{\mathbb{C}} T(\lambda, \lambda^{-1}) = p$ for all $\lambda \in \mathbb{C}\backslash\{0\}$.

(*c*) The causality of the system is reflected by the properness of the corresponding transfer function.

It has to be stressed that the notions of controllability and causality are defined as properties of the system behaviour, i.e. on the signal level (Willems 1986). In this paper we only refer to the consequences of these properties in specific representations of the behaviour as in the polynomial forms discussed. The notion of observability will be discussed later.

Structural properties of these polynomial representations will be elaborated in the rest of this paper. We briefly pay attention to a second representation of system behaviour in terms of a state space description.

Proposition 2.6: *For any system $S \in \Sigma_{p,m}$ that is causal there exists an integer n and matrices A, B, C, D with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$, such that*

$$\mathscr{B}(S) = \{(y, u) \mid \exists\, x \in (\mathbb{R}^n)^{\mathbb{Z}},\quad \sigma x = Ax + Bu, y = Cx + Du\}$$

In this representation x is called the state, and n the state space dimension. Similar to the situation of polynomial representations, the dynamical systems S will be said to be induced by the realization (A, B, C, D), denoted by $S = M_{ss}(A, B, C, D)$. A realization of S which has a minimal state space dimension n is called a minimal realization. This situation seems to be very similar to the 'classical' way of dealing with state space models. However, there is a difference with respect to the minimality of the realization. In 'classical' terms a realization is minimal if it is observable and reachable, i.e. (C, A) is an observable pair and (A, B) is a reachable pair. In the context discussed here, the state space parametrization acts as a representation of a model behaviour and a model behaviour can essentially contain non-controllable parts, represented in a state space form by the non-reachability of the pair (A, B). Consequently, the minimality of a state space representation is now mainly due to observability of the pair (C, A) as reflected in the following proposition.

Proposition 2.7: *Let $S \in \Sigma_{p,m}$ with $S = M_{ss}(A, B, C, D)$, and (A, B, C, D) a realization of dimension n. Then (A, B, C, D) is a minimal realization of S if and only if*

(*a*) *(C, A) is an observable pair, i.e.*

$$\operatorname{rank}_{\mathbb{C}}\begin{pmatrix}\lambda I - A\\ C\end{pmatrix} = n \quad \textit{for all } \lambda \in \mathbb{C}$$

and

(*b*) $\operatorname{rank}\,[A \mid B] = n$.

Condition (*b*) of Proposition 2.7 can be given the interpretation that (A, B) should not have any unreachable poles in 0. Unreachable poles in 0 lead to non-minimality of the realization because a corresponding state component cannot contribute to the behaviour $\mathscr{B}(S)$. This can be visualized by considering that any value of this state component $x_i(t)$ for some $t = t_1$ would become 0 for $t = t_1 + 1$. Since this mode is not affected by any input signal and the considered time set is $\mathbb{Z}$, this $x_i(t_1)$ could never have become $\neq 0$.

Note that in the presented framework a non-symmetrical treatment is created of the 'classical' concepts of reachability/controllability and observability. Non-reachable parts in a state space realization are expressed in the model behaviour, whereas non-observable parts are not; from a behavioural point of view non-observability is actually just a matter of redundant representation. Additional system properties in terms of state space representations are:

(*a*) The system's transfer function equals $H(z) = C(zI - A)^{-1}B + D$;

(*b*) A system S is controllable if and only if a realization (A, B, C, D) of S with dimension n satisfies

$$\operatorname{rank}_{\mathbb{C}}\,[\lambda I - A \mid B] = n \quad \text{for all } \lambda \in \mathbb{C}\backslash\{0\} \tag{6}$$

The minimal state space dimension of a dynamical system will be used in the sequel of this paper when discussing the McMillan degree and the order of dynamical systems.

3. Definitions and notation on polynomial matrices

Before we can come to the main goal of this paper—a discussion on the system order and observability indices of dynamical systems in terms of their polynomial representation—we have to introduce some additional notation and definitions on polynomial matrices.

Definition 3.1: Consider a polynomial matrix $T \in \mathbb{R}^{p \times q}[z, z^{-1}]$. Denote the following:

$\delta^{(u)}(T) :=$ the *maximum power* of z in $T(z, z^{-1})$

$\delta^{(l)}(T) :=$ the *minimum power* of z in $T(z, z^{-1})$

$\delta(T) := \delta^{(u)}(T) - \delta^{(l)}(T) \geqslant 0$

$T_{i*} :=$ the ith *row* of $T(z, z^{-1})$

$\nu_i^{(u)}(T) :=$ the *maximum power* of z in $T_{i*}(z, z^{-1})$

$\nu_i^{(l)}(T) :=$ the *minimum power* of z in $T_{i*}(z, z^{-1})$

$\nu_i(T) := \nu_i^{(u)}(T) - \nu_i^{(l)}(T) \geqslant 0$

$\pi_i^{(u)}(T) :=$ the *maximum power* of z in any $i \times i$-minor of $T(z, z^{-1})$

$\pi_i^{(l)}(T) :=$ the *minimum power* of z in any $i \times i$-minor of $T(z, z^{-1})$ □

Note that for MFD polynomial representations $T \in \mathbb{R}^{p \times (p+m)}[z]$ the integer indices $\delta^{(u)}, \delta^{(l)}, \nu_i^{(u)}, \nu_i^{(l)}, \pi_i^{(u)}, \pi_i^{(l)}$ will all be $\geqslant 0$, whereas for ARMA representations $T \in \mathbb{R}^{p \times (p+m)}[z^{-1}]$ they will all be $\leqslant 0$.

Definition 3.2: Consider a polynomial matrix $T \in \mathbb{R}^{p \times q}[z, z^{-1}]$ with row degrees $\nu_i^{(u)}$ and $\nu_i^{(l)}$ and let T be written as

$$T(z, z^{-1}) = \operatorname{diag}(z^{\nu_1^{(u)}}, \ldots, z^{\nu_p^{(u)}})\Gamma_{hr} + T_1(z, z^{-1}) \tag{7}$$

and

$$T(z, z^{-1}) = T_2(z, z^{-1}) + \operatorname{diag}(z^{\nu_1^{(l)}}, \ldots, z^{\nu_p^{(l)}})\Gamma_{lr} \tag{8}$$

with $\Gamma_{hr}, \Gamma_{lr} \in \mathbb{R}^{p \times q}$ and $T_1, T_2 \in \mathbb{R}^{p \times q}[z, z^{-1}]$ satisfying $\nu_i^{(u)}(T_1) < \nu_i^{(u)}(T)$ and $\nu_i^{(l)}(T_2) > \nu_i^{(l)}(T)$ for $i = 1, \ldots, p$, then

$\Gamma_{hr}(T) :=$ the *leading row coefficient matrix* of T, i.e. the coefficient matrix related to the highest row degree terms in T;

$\Gamma_{lr}(T) :=$ the *trailing row coefficient matrix* of T, i.e. the coefficient matrix related to the lowest row degree terms in T; □

The well-known property of row properness of polynomial matrices (Wolovich 1974) has to be extended to the situation of polynomial matrices in two indeterminates.

Definition 3.3: A polynomial matrix $T \in \mathbb{R}^{p \times q}[z, z^{-1}]$ is called *row proper over* $\mathbb{R}[z]$ if rank $\Gamma_{hr}(T) = p$; it is called *row proper over* $\mathbb{R}[z^{-1}]$ if rank $\Gamma_{lr}(T) = p$. T is called *bilaterally row proper* if it is row proper over $\mathbb{R}[z, z^{-1}]$, i.e.

$$\operatorname{rank} \Gamma_{hr}(T) = \operatorname{rank} \Gamma_{lr}(T) = p$$ □

The notion of bilateral row properness was first presented by Willems (1986). Parametrizations in polynomial matrix form are generally not used in a representation with both shift operators σ and σ^{-1}. It is common to apply either one of the shift operators. In accordance with the nomenclature often used in the literature, we refer to an MFD-model in the case of a polynomial representation $T_f \in \mathbb{R}^{p \times (p+m)}[z]$, and to an ARMA-model for a representation $T_b \in \mathbb{R}^{p \times (p+m)}[z^{-1}]$. Note that for any system $S \in \Sigma_{p,m}$ a representation in either MFD or ARMA form always exists. Both representations can be transformed into one another by premultiplication of the rows in the polynomial matrices by (positive or negative) powers of z. This shift operation will not change the system behaviour.

However, the two representations do make a difference when we consider model sets that are defined by specific restrictions on the degrees of polynomial entries in the matrices T_b or T_f.

4. McMillan degree, system order and minimal state space dimension

We start this section by considering the McMillan degree of a rational matrix which is interpreted as the transfer function of a dynamical input–output system $S \in \Sigma_{p,m}$.

Definition 4.1: The *McMillan degree of a rational matrix* $H(z) \in \mathbb{R}^{p \times m}(z)$ is defined as

$$\delta_M(H) := \text{the sum of polar degrees at all poles of } H(z) \qquad \square$$

For finite poles this sum of polar degrees can be determined by the sum of degrees of the denominator polynomials in the Smith–McMillan form of $H(z)$. The polar degree at infinity has to be added to arrive at the McMillan degree (see e.g. Kailath 1980). For systems that are controllable and causal there exists a close connection between the McMillan degree of the transfer function and the minimal state space dimension that is required for representing the corresponding input–output system. This is formulated in the following proposition.

Proposition 4.2: *Let S be a dynamical system $S \in \Sigma_{p,m}$ that is controllable and causal, and let $H(z) \in \mathbb{R}^{p \times m}(z)$ be the corresponding proper transfer function of S. Then $\delta_M(H)$ equals the state space dimension of any minimal realization (A, B, C, D) of S.*

The proposition is a classical result from the theory of dynamical systems, see e.g. Rosenbrock (1970) and Kailath (1980). However note that the classical minimal realization of a transfer function also requires controllability in order to match the McMillan degree. The McMillan degree is considered to be a measure of complexity that is related to the transfer function of an input–output dynamical system. Therefore it takes account only of the controllable part of a dynamical system and it neglects any non-controllability. The dimension of a minimal state space realization is only defined for causal systems, related to proper transfer functions $H(z)$. For this reason it is appropriate to consider a closely related measure of complexity that takes account of both non-controllable as well as non-causal parts in a system. To this end we will use a notion of system's order, formulated in terms of the model behaviour (Willems 1986), which will be shown to be a generalization of the McMillan degree of a dynamical input/output system as discussed above.

Definition 4.3 (Willems 1986): Let $S = (\mathbb{Z}, W, \mathscr{B})$ be a dynamical system $S \in \Sigma_{p,m}$. Then define the *order* of S as:

$$n(S) := \sum_{N=1}^{\infty} (\dim(\mathscr{B}^N) - \dim(\mathscr{B}^{N-1}) - m) \tag{9}$$

with $\mathscr{B}^N := \mathscr{B}|_{\mathbb{Z} \cap [0, N-1]}$, i.e. the set of finite time trajectories $\{(y(t)\ u(t)),\ t = 0, \ldots, N-1\}$ that are admissible by the system, and $\dim(\mathscr{B}^0) := 0$. □

For a detailed interpretation of (9) we refer to Willems (1986). A brief illustration of the concept is given in the following example.

Example 4.4: Consider a dynamical system $S \in \Sigma_{1,1}$ described by the difference equation

$$y(t) = ay(t-1) + bu(t-1) \tag{10}$$

with $a, b \in \mathbb{R}\backslash\{0\}$. When we evaluate $\mathscr{B}^1$ it follows that

$$\mathscr{B}^1 := \{y(0), u(0) \mid y(t) - ay(t-1) - bu(t-1) = 0 \quad \forall\, t \in \mathbb{Z}\} \tag{11}$$

Note that the set of equations (10) for $t \in \mathbb{Z}$ does not put any restriction on the set of admissible values $\{y(0), u(0)\}$; $\{y(0), u(0)\}$ can be given arbitrary values without conflicting with the difference equations; hence $\dim \mathscr{B}^1 = 2$.

For $\mathscr{B}^2$ it follows:

$$\mathscr{B}^2 := \{y(1), u(1), y(0), u(0) \mid y(t) - ay(t-1) - bu(t-1) = 0, \quad \forall\, t \in \mathbb{Z}\} \tag{12}$$

Having four signal values to be chosen, there is one equation from (10) ($t = 1$) that is restricting the admissible space of these four signal values. Consequently, $\dim \mathscr{B}^2 = 4 - 1 = 3$. Analysing $\mathscr{B}^N$, $N > 2$, shows that every additional step in N adds two signals and one restricting equation, so the corresponding dimension increases by one, i.e. the dimension of the input space.

Applying (9) shows that $n(S) = \dim \mathscr{B}^1 - 0 - 1 = 1$. □

The relation of $n(S)$ with the McMillan degree and the minimal state space dimension is presented in the following theorem, which also gives an expression for the determination of $n(S)$ based on a polynomial representation of the system.

Theorem 4.5: *Let S be a dynamical system $S \in \Sigma_{p,m}$ being induced by a polynomial matrix $T \in \mathbb{R}^{p \times (p+m)}[z, z^{-1}]$, with $T = [P \mid -Q]$. Then*

(*a*) $n(S) = \pi_p^{(u)}(T) - \pi_p^{(l)}(T)$;

(*b*) *there exists a permutation matrix $X \in \mathbb{R}^{(p+m)\times(p+m)}$ such that TX induces a system $S^* \in \Sigma_{p,m}$ that is a causal dynamical system with m inputs and p outputs; any such system S^* has a minimal realization with dimension $n(S)$.*

(*c*) $n(S) \geqslant \delta_{\mathrm{M}}(P^{-1}Q)$ *with equality if and only if* $\mathrm{rank}_{\mathbb{C}}\, T(\lambda, \lambda^{-1}) = p$ *for all* $0 \neq \lambda \in \mathbb{C}$, *or equivalently S is controllable.*

(*d*) $n(S) \leqslant \sum_{i=1}^{p} v_i(T)$ *with equality if and only if T is bilaterally row proper.*

Proof: Parts (*b*) and (*d*) are proved in Willems (1986); the proofs of parts (*a*) and (*c*) are given in the Appendix. □

Part (*a*) shows a general expression for $n(S)$ in terms of a polynomial representation. It requires the evaluation of the degrees of all $p \times p$ minors of T. Since the order $n(S)$ is defined on the level of system behaviour, it is invariant for any unimodular premultiplication of T, as, for example, for premultiplication of T with any power of z or z^{-1}. It follows from this theorem that $n(S)$ is a generalization of both the McMillan degree of a transfer function (as in c) and the minimal state space dimension of a state space realization (b). The permutation matrix X in (b) should be interpreted as an operation of reordering the signals (y, u), and thus possibly interchanging input and output signals. If S is causal then the permutation matrix X can be chosen to be the identity matrix and $n(S)$ equals the corresponding minimal state space dimension. If S is non-causal, the signal variables $w = (y, u)$ can be reordered in such a way that the reordered system S^* is causal and again allows a state space representation with minimal dimension $n(S)$. Note that $\mathscr{B}(S^*)$ is constructed according to $w \in \mathscr{B}(S^*) \Leftrightarrow Xw \in \mathscr{B}(S)$. It can simply be verified that any postmultiplication of T with a non-singular constant matrix $R \in \mathbb{R}^{(p+m)\times(p+m)}$ leaves the system order as expressed in Theorem 4.5(a) invariant.

Part (c) of the theorem shows that $n(S)$ takes account of the non-controllable modes in the system. If S is controllable then $n(S)$ is equal to the McMillan degree of the corresponding transfer function.

Bilateral row properness of a polynomial representation appears to be an important property since it allows the system order to be simply determined from the sum of row degrees in the polynomial matrix, as formulated in part (d). In the next section an algorithm will be described for bringing any polynomial representation into a bilaterally row proper form.

The results of the theorem are illustrated in the following example, which is taken from Janssen (1988 b).

Example 4.6: Consider a dynamical system $S \in \Sigma_{2,1}$, induced by the polynomial matrix $T(z, z^{-1}) = [P \mid -Q]$, with

$$P(z, z^{-1}) = \begin{bmatrix} z + z^{-2} & z^{-1} \\ z^{-1} & 1 \end{bmatrix}, \quad Q(z, z^{-1}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

It can easily be verified that rank $T(\lambda, \lambda^{-1}) = 2$ for all $\lambda \in \mathbb{C}\backslash\{0\}$, which implies that S is controllable. Its transfer function $H(z) = \begin{bmatrix} z^{-1} \\ -z^{-2} \end{bmatrix}$. The row degrees of T satisfy: $\nu_1^{(u)} = 1$, $\nu_1^{(l)} = -2$, $\nu_2^{(u)} = 0$, $\nu_2^{(l)} = -1$ leading to $\nu_1 + \nu_2 = 4$. In order to check the bilaterally row properness of T we verify that

$$\Gamma_{hr}(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \text{ and } \Gamma_{lr}(T) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Since $\Gamma_{lr}(T)$ does not have full row rank, T is not bilaterally row proper and consequently (by Theorem 4.5(d)) $n(S) < 4$. Evaluating the system order by Theorem 4.5(a) shows $\pi_2^{(u)}(T) = 1$, $\pi_2^{(l)}(T) = -1$ leading to $n(S) = 2$ which is equal to the McMillan degree of $H(z)$. Matrix T can be transformed to a bilaterally row proper form by unimodular premultiplication. Construct $T^* = UT$ with

$$U(z, z^{-1}) = \begin{bmatrix} 1 & -z^{-1} \\ 0 & 1 \end{bmatrix}$$

it follows that

$$T^* = \begin{bmatrix} z & 0 & -1 \\ z^{-1} & 1 & 0 \end{bmatrix}$$

T^* is bilaterally row proper with row degrees satisfying $\nu_1 + \nu_2 = 1 + 1 = 2 = n(S)$. □

For some specific situations the expression $\pi_p^{(u)}(T) - \pi_p^{(l)}(T)$ for the system order can be further simplified, as formulated in the following proposition.

Proposition 4.7: *Let S be a dynamical system $S \in \Sigma_{p,m}$, being induced by a polynomial matrix $T \in \mathbb{R}^{p \times (p+m)}[z, z^{-1}]$, with $T = [P \mid -Q]$. Then*

(*a*) $\pi_p^{(u)}(T) = \delta^{(u)}(\det(P))$ *if and only if $P^{-1}Q$ has no poles in $z = \infty$;*

(*b*) $\pi_p^{(l)}(T) = \delta^{(l)}(\det(P))$ *if and only if $P^{-1}Q$ has no poles in $z = 0$.*

Proof: See the Appendix for the proof. □

Note that the absence of poles in $z = \infty$ refers to the situation of a proper transfer function, and correspondingly to a causal system. The results of Theorem 4.5 and Proposition 4.7 include a generalization of the results on the McMillan degree of MFD and ARMA models as reported in Janssen (1988 *a*, *b*), where separate results were derived for models represented in one of the shift operators σ or σ^{-1}. A generalized model representation in both shift operators leads to an overall theory that covers both situations of MFD and ARMA representations. The connections with the results of Janssen (1988 *a*, *b*) are shown in the following corollary.

Corollary 4.8 *Let S be a dynamical system $S \in \Sigma_{p,m}$, represented in MFD or ARMA form by*

$$S = M_p(T_f), \quad T_f = T_f(z) = [P_f \mid -Q_f] \in \mathbb{R}^{p \times (p+m)}[z]$$

and

$$S = M_p(T_b), \quad T_b = T_b(z^{-1}) = [P_b \mid -Q_b] \in \mathbb{R}^{p \times (p+m)}[z^{-1}]$$

Then the following statements hold true

(*a*) *If* $\operatorname{rank}_{\mathbb{C}} T_f(\lambda) = p$ *for* $\lambda = 0$ *then* $\pi_p^{(l)}(T_f) = 0$;

(*b*) *If* $\operatorname{rank}_{\mathbb{C}} T_b(\lambda) = p$ *for* $\lambda = 0$ *then* $\pi_p^{(u)}(T_b) = 0$;

(*c*) *If P_f, Q_f are left coprime over $\mathbb{R}[z]$, and $P_f^{-1}Q_f$ has no poles in $z = \infty$ then* $\delta_M(P_f^{-1}Q_f) = n(S) = \delta^{(u)}(\det(P_f))$;

(*d*) *If P_b, Q_b are left coprime over $\mathbb{R}[z^{-1}]$ and* $P_b^{-1}Q_b$ has no poles in $z = 0$ then $\delta_M(P_b^{-1}Q_b) = n(S) = -\delta^{(l)}(\det(P_b))$.

Proof:

(*a*) If $\pi_p^{(l)}(T_f) \neq 0$ then $\pi_p^{(l)}(T_f) > 0$ and consequently $T_f(\lambda)$ has a zero in $\lambda = 0$, which conflicts with the condition that $\operatorname{rank}_{\mathbb{C}} T_f(\lambda) = p$ for $\lambda = 0$.

(*b*) If $\pi_p^{(u)}(T_b) \neq 0$ then $\pi_p^{(u)}(T_b) < 0$ and consequently $T_b(\lambda)$ has a zero in $\lambda = 0$, which conflicts with the condition that $\operatorname{rank}_{\mathbb{C}} T_b(\lambda) = p$ for $\lambda = 0$.

(*c*) The statement follows directly from Theorem 4.5(*c*), Proposition 4.7, and part (*a*) of this corollary, taking into account that left coprimeness over $\mathbb{R}[z]$ is equivalent with the condition $\operatorname{rank}_{\mathbb{C}} T(\lambda) = p$ for all $\lambda \in \mathbb{C}$.

(*d*) The proof is similar to the proof of (*c*), applying Theorem 4.5(*c*), Proposition 4.7 and part (*b*) of this corollary. □

There is a certain symmetry in the expressions for MFD and ARMA representations of systems. However, there are still some important differences. Note that in situations (*c*) and (*d*) as formulated in the corollary, the order of the system can be extracted from the matrices P_b, respectively P_f only, irrespective of the matrices Q_b, Q_f. However this situation requires different restrictions on the dynamical systems dependent on the question whether MFD or ARMA representations are involved. In the case of an MFD representation (situation (*c*)), this restriction is the quite natural condition of causality of the system (the system should not contain anticipations). In the case of ARMA representations (situation (*d*)) the condition $H(z)$ having no poles in $z = 0$ (the system should not contain any delays) is much more restrictive. This is exactly the reason why model sets of causal models with specified model order n are much easier to parametrize in MFD form than in ARMA form. Absence of poles in the origin, is exactly the restriction that is made by Bokor and Keviczky (1987) in order to come up with a canonical ARMA parametrization for causal systems having a fixed McMillan degree.

5. Algorithm for obtaining bilaterally row proper polynomial matrices

As shown in the previous section, and as will be further specified in the next, bilaterally row properness is an important property of a polynomial representation of dynamical systems. In this section we will sketch an algorithm that brings a general polynomial representation to a bilaterally row proper form, through unimodular premultiplication.

Algorithm 5.1

Consider $T \in \mathbb{R}^{p \times q}[z, z^{-1}]$, $q \geqslant p$, with

$$\begin{bmatrix} T_{1*} \\ \vdots \\ T_{p*} \end{bmatrix} \quad \text{and} \quad T_{i*} = t_{i,v_i^{(u)}} z^{v_i^{(u)}} + \cdots + t_{i,v_i^{(l)}} z^{v_i^{(l)}}$$

Step 1. Reorder the rows of T such that $v_1 \leqslant v_2 \leqslant \cdots \leqslant v_p$.

Step 2. Set $k = 1$.

Step 3. Evaluate the rank of the matrix

$$L_k = \begin{bmatrix} t_{1,v_1^{(u)}} \\ \vdots \\ t_{k-1,v_{k-1}^{(u)}} \\ t_{k,v_k^{(u)}} \end{bmatrix}$$

(*a*) If $k = p$ and L_k has full row rank, go to Step 4;

(*b*) If $k \leqslant p$ and L_k has full row rank, set $k = k + 1$ and return to Step 3.

(*c*) If rank $(L_k) < k$ then $t_{k,v_k^{(u)}}$ is linearly dependent on the previous rows in L_k and there exist real numbers α_i, $i = 1, \ldots, k-1$, such that

$$t_{k,v_k^{(u)}} = \sum_{i=1}^{k-1} \alpha_i t_{i,v_i^{(u)}} \tag{13}$$

Replace T_{k*} by

$$\tilde{T}_{k*} = T_{k*} - \sum_{i=1}^{k-1} \alpha_i z^{v_k^{(u)} - v_i^{(u)}} T_{i*} \tag{14}$$

When, after replacement of the kth row, $v_k(T) < v_{k-1}(T)$, then reorder the rows of T again such that $v_1 \leqslant v_2 \leqslant \ldots \leqslant v_p$.
After reordering, the kth row of T becomes the jth row of T, ($j \leqslant k$); set $k = j$ and return to Step 3.

Step 4. Set $k = 1$.

Step 5. Evaluate the rank of the matrix

$$M_k = \begin{bmatrix} t_{1,v_1^{(l)}} \\ \vdots \\ t_{k-1,v_{k-1}^{(l)}} \\ t_{k,v_k^{(l)}} \end{bmatrix}$$

(*a*) If $k = p$ and M_k has full row rank, the algorithm stops;

(*b*) If $k \leqslant p$ and M_k has full row rank, set $k = k + 1$ and return to Step 5.

(*c*) If rank $(M_k) < k$ then $t_{k,v_k^{(l)}}$ is linearly dependent on the previous rows in M_k and there exist real numbers β_i, $i = 1, \ldots, k - 1$, such that

$$t_{k,v_k^{(l)}} = \sum_{i=1}^{k-1} \beta_i t_{i,v_i^{(l)}} \tag{15}$$

Replace T_{k*} by

$$\tilde{T}_{k*} = T_{k*} - \sum_{i=1}^{k-1} \beta_i z^{v_k^{(l)} - v_i^{(l)}} T_{i*} \tag{16}$$

When, after replacement of the kth row, $v_k(T) < v_{k-1}(T)$, then reorder the rows of T again such that $v_1 \leqslant v_2 \leqslant \ldots \leqslant v_p$.
After reordering, the kth row of T becomes the jth row of T, ($j \leqslant k$); set $k = j$ and return to Step 5. □

The correctness of the algorithm can be verified by considering the following remarks.

(i) All operations performed on the polynomial matrix T are unimodular premultiplications, and consequently the dynamical system that is represented by T is not changed;

(ii) It can be verified that the essential operations on T in Step 3 and Step 5, formulated in (14), (16) force the sum of row degrees of T to be decreased by at least 1 in every step. Consequently the algorithm has to stop in a finite number of steps.

(iii) In Step 3 the leading row coefficient matrix of T is constructed to be of full row rank, while in Step 5 this operation is performed directed towards the trailing coefficient matrix. It can be verified that the operations of Step 5 do not affect the full row rank of $\Gamma_{hr}(T)$.

The verification of items (ii) and (iii), essentially relies on the fact that at every step in the algorithm, the matrix has an ordered set of row degrees, as formulated in Step 1 of the algorithm.

6. Structure indices

In the same line of thought as exhibited in § 4, results can be formulated for the structural indices of dynamical systems, represented in the generalized polynomial forms as discussed. Structural indices constitute a further specification of the order of dynamical systems and are often used for the construction of sets of dynamical systems that allow an identifiable and differentiable parametrization. Let us first consider the different structure indices that will be considered.

Definition 6.1: Let (A, B, C, D) be a state space realization with dimension n of a dynamical system $S \in \Sigma_{p,m}$, with $C = [c_1^T \; c_2^T \; \dots \; c_p^T]^T$. Consider the observability matrix $\mathcal{O} = [C^T \; A^T C^T \; \dots \; (A^T)^{n-1} C^T]^T$ and evaluate the linearly independent rows in $\mathcal{O}$ from top to bottom. Order these independent rows as

$$\{c_1, c_1 A, \dots, c_1 A^{\gamma_1 - 1}, \dots, c_p, \dots, c_p A^{\gamma_p - 1}\}$$

then the integers $(\gamma_i)_{i=1,\dots,p}$ are defined to be the *observability indices of the realization* (A, B, C, D). □

The observability indices of a realization are often called the (left) Kronecker indices of the realization, and actually are equal to the left Kronecker indices of the matrix pencil $[(zI - A)^T \; C^T]^T$ (Kailath, 1980). The notation $(\gamma_i)_{i=1,\dots,p}$ refers to a mapping $\mathbb{Z} \cap [1, p] \to \mathbb{Z}_+$, which means that the ordering of the observability indices is of importance, (e.g. $(2, 1) \neq (1, 2)$). When writing $\{\gamma_i\}_{i=1,\dots,p}$ we will refer to a set with p elements, where consequently any ordering of the elements is not relevant. In the case of a minimal realization (A, B, C, D) it follows directly that $\sum_{i=1}^{p} \gamma_i = n(S)$. The observability indices are presented as being related to state space realizations and consequently their interpretation is restricted to causal systems.

Related structure indices can also be defined on the basis of transfer functions. The definition of these left Kronecker indices will be adopted from Forney (1975), Kailath (1980) and Janssen (1988 *a*), and will not be further elaborated here. For more details the reader is referred to these references.

Definition 6.2: Let $H(z)$ be a rational matrix, $H(z) \in \mathbb{R}^{p \times m}(z)$, satisfying $\text{rank}_{\mathbb{R}(z)} H = p$. The *set of left Kronecker*† *indices* $\{\kappa_i\}_{i=1,\dots,p}$ of $H(z)$, is defined as the set of row degrees of any minimal polynomial basis in $\mathbb{R}[z]$ for the rational vector space, generated by the rows of the matrix $[I_p \mid H(z)]$.

Minimality of this basis is considered with respect to the sum of the row degrees. □

In Janssen (1988 *a*) it is proven that these left Kronecker indices of $H(z)$ sum up to its McMillan degree:

$$\sum_{i=1}^{p} \kappa_i = \delta_M(H)$$

Note that these left Kronecker indices are defined on rational matrices, i.e. on the level of transfer functions. This means that these structure indices do not contain information on any non-controllable part of a dynamical system. Again, as in the

† The use of the term *Kronecker indices* may be confusing in this definition, since Kronecker indices formally are defined on matrix pencils and not on general rational matrices. For this reason Janssen (1988 b) has introduced the term *left dynamical indices* for the structure indices defined. In this paper we will follow the terms used by Kailath (1980).

case of the system order $n(S)$, both sets of structure indices as defined in definitions 6.1, 6.2 can be related to each other in terms of a polynomial matrix representation, in order to deal with both non-controllable and non-causal dynamical systems. To this end the following definition is given.

Definition 6.3: Let S be a dynamical system $S \in \Sigma_{p,m}$, and let T be a polynomial matrix $T \in \mathbb{R}^{p \times (p+m)}[z, z^{-1}]$. The *set of observability indices of* S, denoted by $\{\rho_i\}_{i=1,\dots,p}$, is defined as the set of row degrees $\{v_i(T)\}_{i=1,\dots,p}$ of any bilaterally row proper matrix T that induces S. □

It can be verified that this set of observability indices is an invariant of the system S, i.e. it is not dependent on the specific bilaterally row proper matrix T that it is induced by. Its relation with the structure indices previously defined, is given in the following proposition.

Proposition 6.4: *Let S be a dynamical system $S \in \Sigma_{p,m}$ being induced by a polynomial matrix $T \in \mathbb{R}^{p \times (p+m)}[z, z^{-1}]$ then the following statements hold true.*

(*a*) *The set of observability indices of S is equal to the set of left Kronecker indices $\{\kappa_i\}_{i=1,\dots,p}$ of its transfer function if and only if S is controllable.*

(*b*) *The set of observability indices of S is equal to the set of observability indices $\{\gamma_i\}_{i=1,\dots,p}$ of any minimal space realization (A, B, C, D) that satisfies $M_{ss}(A, B, C, D) = M_p(TX)$ and $X \in \mathbb{R}^{(p+m)\times(p+m)}$ a non-singular matrix.*

Proof: Part (*a*) of this theorem is an extension of results of Forney (1975) and Kailath (1980); for the situation that $T \in \mathbb{R}^{p \times (p+m)}[z]$, it has been shown that the set of row degrees of T is equal to the set of left Kronecker indices of $P^{-1}Q$ if and only if P and Q are left coprime and T is row proper over $\mathbb{R}[z]$. Extension of this result to polynomials in $\mathbb{R}[z, z^{-1}]$ follows from the consideration that {left coprimeness and row properness over $\mathbb{R}[z]$} $\Leftrightarrow$ {left coprimeness and bilaterally row properness over $\mathbb{R}[z, z^{-1}]$}. Part (*b*) follows from results of Willems (1986), and the consideration that the set of row degrees of T is invariant for postmultiplication of T with a constant non-singular matrix. □

The set of observability indices of S coincides with the set of left Kronecker indices in case of controllability, and it coincides with the set of observability indices of a state space realization in case of causality. Note that equality, as meant in the proposition, refers to the sets of indices, i.e. equality of the indices up to ordering. The observability indices as defined, act as a further structural specification of dynamical systems with order

$$n(S) = \sum_{i=1}^{p} \rho_i(S) \tag{17}$$

They have been defined on the basis of a bilaterally row proper polynomial representation of the system. It has to be stressed that this set of indices can also be formulated purely on the basis of system behaviour as is the case with the system order $n(S)$. We will briefly state this result but for further details we refer to Willems (1986).

Proposition 6.5: *Let $S = (\mathbb{Z}, W, \mathcal{B})$ be a dynamical system $S \in \Sigma_{p,m}$. Then the set of observability indices $\{\rho_i\}_{i=1,\dots,p}$ of S can be determined as follows. Let*

$$\beta^N := \dim \mathcal{B}^N - \dim \mathcal{B}^{N-1} \quad \textit{with } \mathcal{B}^0 := 0 \tag{18}$$

and

$$\tau^N := \beta^N - \beta^{N+1} \quad \text{with } \beta^0 := p + m \tag{19}$$

If $\tau^j = k > 0$ *then* j *appears* k*-times as element of* $\{\rho_i\}_{i=1,\ldots,p}$.

In the case of non-causal systems, reflected by non-proper transfer functions, the set of observability indices can also be shown to be related to the left Kronecker indices of the matrix pencil $[(zE - A)^{\mathrm{T}} \; C^{\mathrm{T}}]^{\mathrm{T}}$ generated by a generalized state space representation; this representation is analysed by Kuijper and Schumacher (1990).

Similar results as formulated in Proposition 4.7 and Corollary 4.8 with respect to the system order, can now be given for the evaluation of the set of observability indices of a system in some specific situations. We will briefly pay attention to the situation when the set of observability indices can be determined on the basis of polynomial matrix P only, as a submatrix of the matrix T that induces S. Moreover the specific situation of MFD and ARMA forms will be taken into account. These results will be presented in the following two corollaries.

Corollary 6.6: *Let* S *be a dynamical system* $S \in \Sigma_{p,m}$ *induced by a bilaterally row proper polynomial matrix* $T \in \mathbb{R}^{p \times (p+m)}[z, z^{-1}]$, *with* $T = [P \mid -Q]$, *then*

(*a*) $(\nu_i^{(u)}(T))_{i=1,\ldots,p} = (\nu_i^{(u)}(P))_{i=1,\ldots,p}$ *if and only if* $P^{-1}Q$ *has no poles in* $z = \infty$;

(*b*) $(\nu_i^{(l)}(T))_{i=1,\ldots,p} = (\nu_i^{(l)}(P))_{i=1,\ldots,p}$ *if and only if* $P^{-1}Q$ *has no poles in* $z = 0$.

Proof: If T is bilaterally row proper it follows that $\pi_p^{(u)}(T) = \sum_{i=1}^p \nu_i^{(u)}(T)$ and $\pi_p^{(l)}(T) = \sum_{i=1}^p \nu_i^{(l)}(T)$. Applying this property to Proposition 4.7 shows that the expression $\pi_p^{(u)}(T) = \delta^{(u)}(\det(P))$ can be replaced by $\sum_{i=1}^p \nu_i^{(u)}(T) = \sum_{i=1}^p \nu_i^{(u)}(P)$. Since it is impossible that $\nu_i^{(u)}(T) < \nu_i^{(u)}(P)$ it follows that the latter expression is equivalent to $\nu_i^{(u)}(T) = \nu_i^{(u)}(P)$ for $i = 1, \ldots, p$. Applying a similar reasoning for the lower row degrees (part (*b*)) shows that the corollary follows directly from Proposition 4.7. □

Corollary 6.7: *Let* S *be a dynamical system* $S \in \Sigma_{p,m}$, *represented in MFD or ARMA form by*

$$S = M_{\mathrm{p}}(T_{\mathrm{f}}), \quad T_{\mathrm{f}} = T_{\mathrm{f}}(z) = [P_{\mathrm{f}} \mid -Q_{\mathrm{f}}] \in \mathbb{R}^{p \times (p+m)}[z]$$

and

$$S = M_{\mathrm{p}}(T_{\mathrm{b}}), \quad T_{\mathrm{b}} = T_{\mathrm{b}}(z^{-1}) = [P_{\mathrm{b}} \mid -Q_{\mathrm{b}}] \in \mathbb{R}^{p \times (p+m)}[z^{-1}]$$

Then the following statements hold true

(*a*) *If* $\mathrm{rank}_{\mathbb{C}}\, T_{\mathrm{f}}(\lambda) = p$ *for* $\lambda = 0$ *and* T_{f} *is row proper over* $\mathbb{R}[z]$, *and* $P_{\mathrm{f}}^{-1}Q_{\mathrm{f}}$ *has no poles in* $z = \infty$ *then* $\{\rho_i(S)\}_{i=1,\ldots,p} = \{\nu_i^{(u)}(P_{\mathrm{f}})\}_{i=1,\ldots,p}$;

(*b*) *If* $\mathrm{rank}_{\mathbb{C}}\, T_{\mathrm{b}}(\lambda) = p$ *for* $\lambda = 0$ *and* T_{b} *is row proper over* $\mathbb{R}[z^{-1}]$, *and* $P_{\mathrm{b}}^{-1}Q_{\mathrm{b}}$ *has no poles in* $z = 0$ *then* $\{\rho_i(S)\}_{i=1,\ldots,p} = \{-\nu_i^{(k)}(P_{\mathrm{b}})\}_{i=1,\ldots,p}$. □

Proof: The corollary follows immediately from combining Corollaries 4.8 and 6.6. □

As mentioned before, the expressions for the determination of the set of observability indices shows a clear resemblance to similar expressions for the system order $n(S)$. In order to be able to extract these indices of S only from the submatrix P within T, additional conditions of causality (for MFD forms) and absence of delays (for ARMA forms) have to be satisfied. These two symmetrical conditions clearly show the symmetrical treatment of the two different forms. Finally, we will illustrate the results presented in this section, in the following example.

Example 6.8: Consider again the dynamical system $S \in \Sigma_{2,1}$ as introduced in example 4.6, induced by polynomial matrix $T^* = [P^* \mid -Q^*]$,

$$T^*(z, z^{-1}) = \left[\begin{array}{cc|c} z & 0 & -1 \\ z^{-1} & 1 & 0 \end{array}\right] \tag{20}$$

Since T^* is bilaterally row proper the set of observability indices $\{\rho_1, \rho_2\}$ is determined by the set of row degrees of T^*, which equals $\{1, 1\}$.

An equivalent MFD form for T^* that induces the same system, can be obtained by premultiplying the second row of T^* by z. This is a unimodular operation on T^*. The resulting matrix is

$$T_1 = [P_1 \mid -Q_1] = \left[\begin{array}{cc|c} z & 0 & -1 \\ 1 & z & 0 \end{array}\right] \tag{21}$$

With this matrix T_1 the conditions of Corollary 6.7(a) are satisfied and consequently the set of upper row degrees of P_1 is equal to the set of observability indices of S.

An equivalent ARMA form for T^* that induces the same system, can be obtained by premultiplying the first row of T^* by z^{-1}. This is a unimodular operation on T^*. The resulting matrix is

$$T_2 = [P_2 \mid -Q_2] = \left[\begin{array}{cc|c} 1 & 0 & -z^{-1} \\ z^{-1} & 1 & 0 \end{array}\right] \tag{22}$$

This matrix T_2 satisfies the first two conditions of Corollary 6.7(b); however since the transfer function of S, $H(z) = [(z^{-1}) \quad -(z^{-2})]^{\mathrm{T}}$, contains poles in $z = 0$, the third condition is not satisfied. It follows that the set of observability indices of S cannot be determined from the row degrees of P_2. □

7. Conclusions

In this paper system order and structure indices have been presented as generalizations of, respectively, the McMillan degree and the (Kronecker) observability indices of linear systems, being applicable to systems that are not necessarily causal and/or controllable. To this end the concept of dynamical system in terms of its behaviour is adopted. General expressions for the evaluation of system order and observability indices are derived in terms of system representations in polynomial matrix form. The polynomial forms discussed exhibit two shift operators, being a generalization of both MFD and ARMA forms. Consequently a unifying theory results for the evaluation of system order and observability indices of linear systems in polynomial forms, in which MFD and ARMA forms are special cases that are completely symmetrical.

Appendix

Proof of Theorem 4.5: In order to prove parts (a) and (c), first it will be shown that

$$'\delta_{\mathrm{M}}(P^{-1}Q) \leqslant \pi_p^{(u)}(T) - \pi_p^{(l)}(T) \tag{A 1}$$

with equality if and only if $\mathrm{rank}_{\mathbb{C}}\, T(\lambda, \lambda^{-1}) = p$ for all $0 \neq \lambda \in \mathbb{C}$. Secondly, part ($a$) will be proven.

The following additional notation will be required; let $G(z) \in \mathbb{R}^{p\times m}(z)$, then denote

$$\delta_a^p(G) := \text{the degree of a pole in } z = a; \tag{A 2}$$

$$\delta_a^z(G) := \text{the degree of a zero in } z = a; \tag{A 3}$$

$$\text{exc}_a\,(G) := \delta_a^p(G) - \delta_a^z(G); \tag{A 4}$$

$$\text{exc}_{\text{tot}}\,(G) := \sum_{a\in\mathbb{C}^*} \delta_a^p(G) - \delta_a^z(G), \text{ with } \mathbb{C}^* = \mathbb{C} \cup \{\infty\}; \tag{A 5}$$

$$c_a^{(i)}(G) := \text{the maximum of exc}_a\,(g^{(i)}(z)) \text{ over all } i\times i\text{-minors } g^{(i)}(z) \text{ of } G. \tag{A 6}$$

By definition

$$\delta_{\text{M}}(G) = \sum_{a\in\mathbb{C}^*} \delta_a^p(G) = \sum_{a\in\mathbb{C}^*} \max_{i\geqslant 0}\,[c_a^{(i)}(G)] \quad \text{for all } a \in \mathbb{C}^* \tag{A 7}$$

with $c_a^{(0)}(G) := 0$ (Verghese and Kailath 1981).

For $T = [P \mid -Q]$ being polynomial in z only, and $G = P^{-1}Q$, it has been shown by Janssen (1988 *a*) that $\delta_{\text{M}}(G) = \text{exc}_{\text{tot}}\,(T)$. It can easily be verified that this result remains valid for $T \in \mathbb{R}^{p\times(p+m)}[z, z^{-1}]$. Since T has only poles in $z = 0$ and $z = \infty$, it follows that

$$\delta_{\text{M}}(G) = \text{exc}_{\text{tot}}\,(T) \leqslant \text{exc}_0\,(T) + \text{exc}_\infty\,(T) \tag{A 8}$$

with equality if and only if T has no finite zeros $\neq 0$, or equivalently $\text{rank}_{\mathbb{C}}\, T(\lambda, \lambda^{-1}) = p$ for all $\lambda \in \mathbb{C}\backslash\{0\}$.

Using a result of Verghese and Kailath (1981), showing that $\text{exc}_a\,(H) = c_a^{(r)}(H)$ for any rational matrix H with $\text{rank}_{\mathbb{R}(z)} H = r$, it follows from (A 8) that

$$\delta_{\text{M}}(G) \leqslant c_0^{(p)}(T) + c_\infty^{(p)}(T) \tag{A 9}$$

with equality if and only if $\text{rank}_{\mathbb{C}}\, T(\lambda, \lambda^{-1}) = p$ for all $\lambda \in \mathbb{C}\backslash\{0\}$.

Let any $p\times p$-minor of T be equal to $t(z) = a_1 z^{n_1} + \cdots + a_{n_1-n_2+1} z^{n_2}$ with $n_1, n_2 \in \mathbb{Z}$ and $n_1 \geqslant n_2$, then $\text{exc}_\infty\,(t(z)) = n_1$ and $\text{exc}_0\,(t(z)) = -n_2$; consequently

$$\delta_{\text{M}}(G) \leqslant \pi_p^{(u)}(T) - \pi_p^{(l)}(T) \tag{A 10}$$

with equality if and only if $\text{rank}_{\mathbb{C}}\, T(\lambda, \lambda^{-1}) = p$ for all $0 \neq \lambda \in \mathbb{C}$. This proves the equivalence between the statements (*a*) and (*c*) of the theorem.

Secondly, part (*a*) of the theorem will be proven.

If the polynomial matrix T is bilaterally row proper, it can be written as:

$$T(z, z^{-1}) = \text{diag}\,(z^{u_1}, \ldots, z^{u_p})\Gamma_{hr} + T_1(z, z^{-1}) \tag{A 11}$$

or

$$T(z, z^{-1}) = T_2(z, z^{-1}) + \text{diag}\,(z^{l_1}, \ldots, z^{l_p})\Gamma_{lr} \tag{A 12}$$

with $u_i = v_i^{(u)}(T)$, $l_i = v_i^{(l)}(T)$; $T_1(z, z^{-1})$ satisfying $v_i^{(u)}(T_1) < u_i$ and $T_2(z, z^{-1})$ satisfying $v_2^{(l)}(T_2) > l_i$, $i = 1, \ldots, p$.

Since Γ_{hr} and Γ_{lr} have full row rank, $\pi_p^{(l)}(T) = \sum_i u_i$, and $\pi_p^{(l)}(T) = \sum_i l_i$, leading to $\pi_p^{(u)}(T) - \pi_p^{(l)}(T) = \sum_i (u_i - l_i) = \sum_{i=1}^p v_i(T) = n(S)$.

If T is not bilaterally row proper, it can always be brought to a bilaterally row proper form by unimodular premultiplication. Let $T'(z, z^{-1}) = U(z, z^{-1})T(z, z^{-1})$ with U unimodular over $\mathbb{R}[z, z^{-1}]$, then any $p\times p$-minor of T' equals $\det\,(U)\cdot$ (the corresponding $p \times p$-minor of T). Since $\det(U) = cz^d$, $0 \neq c \in \mathbb{R}$, $d \in \mathbb{Z}$, it follows that $\pi_p^{(u)}(T') - \pi_p^{(l)}(T') = \pi_p^{(u)}(T) - \pi_p^{(l)}(T)$, showing that $\pi_p^{(u)}(\cdot) - \pi_p^{(l)}(\cdot)$ is representation independent, which proves statement *a* of the theorem. □

Proof of Proposition 4.7: Let $G(z) = P(z, z^{-1})^{-1}Q(z, z^{-1})$ and $T = [P \mid -Q]$. In this proof we use lemma P4A-4 from Janssen (1988 *a*), stating that, using the notation of (A 5), $\delta_\alpha^p(G) = \text{exc}_\alpha(T) - \text{exc}_\alpha(P)$ for $\alpha \in \mathbb{C}^*$.

(*a*) G proper $\Leftrightarrow \delta_\infty^p(G) = 0$. Using the lemma mentioned above it follows that this is equivalent with $\text{exc}_\infty(T) = \text{exc}_\infty(P)$. Since polynomial matrices T and P satisfy $\text{rank}_{\mathbb{R}(z)} T = \text{rank}_{\mathbb{R}(z)}(P) = p$ the previously mentioned result of Verghese and Kailath (1981) ensures that $\text{exc}_\infty(T) = \text{exc}_\infty(P) \Leftrightarrow c_\infty^{(p)}(T) = c_\infty^{(p)}(P) \Leftrightarrow \pi_p^{(u)}(T) = \pi_p^{(u)}(P) = \delta^{(u)}\{\det(P)\}$.

(*b*) Using reasoning similar to that in part (*a*), it follows that $\delta_0^p(G) = 0 \Leftrightarrow \text{exc}_0(T) = \text{exc}_0(P) \Leftrightarrow c_0^{(p)}(T) = c_0^{(p)}(P) \Leftrightarrow \pi_p^{(l)}(T) = \pi_p^{(l)}(P) = \delta^{(l)}\{\det(P)\}$. □

References

Bokor, J., and Keviczky, L., 1987, ARMA canonical forms obtained from constructibility invariants. *International Journal of Control*, **45**, 861–873.

Deistler, M., 1983, The properties of the parametrization of ARMAX systems and their relevance for structural estimation. *Econometrica*, **51**, 1187–1207.

Deistler, M., Dunsmuir, W., and Hanan, E. J., 1978, Vector linear time series models: corrections and extensions. *Advances in Applied Probability*, **10**, 360–372.

Forney, G. D., 1975, Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM Journal on Control and Optimization*, **13**, 493–520.

Gevers, M., 1986, ARMA models, their Kronecker indices and their McMillan degree. *International Journal of Control*, **43**, 1745–1761.

Guidorzi, R. P., 1975, Canonical structures in the identification of multivariable systems. *Automatica*, **11**, 361–374; 1981, Invariants and canonical forms for systems structural and parametric identification. *Ibid* **17**, 117–133.

Hanan, E. J., and Kavalieris, L., 1984, Multivariable linear time series models. *Advances in Applied Probability*, **16**, 492–561.

Janssen, P. H. M., 1988 *a*, On model parametrization and model structure selection for identification of MIMO systems. Doctoral dissertation, Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands; 1988 *b*, General results on the McMillan degree and the Kronecker indices of ARMA and MFD models. *International Journal of Control*, **48**, 591–608.

Kailath, T., 1980, *Linear Systems* (Englewood Cliffs, NJ: Prentice Hall).

Kuijper, M., and Schumacher, J. M., 1990, Realization of autoregressive equations in pencil and descriptor form. *SIAM Journal on Control and Optimization*, **28**, 1162–1189.

Ljung, L., 1987, *System Identification: Theory for the User* (Englewood Cliffs, NJ: Prentice Hall).

Rosenbrock, H. H., 1970, *State Space and Multivariable Theory* (London: Nelson).

Van den Hof, P. M. J., 1989 *a*, On residual-based parametrization and identification of multivariable systems. Doctoral dissertation, Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands; 1989 *b*, A deterministic approach to approximate modelling of input-output data, *Proceedings of the 28th I.E.E.E. Conference on Decision and Control*, Tampa, Florida, pp. 659–664; 1991, On the order and structural indices of linear systems represented in polynomial form. *Automatic Control in the Service of Mankind: Proceedings of the 11th IFAC World Congress*, Tallinn, Estonia (1990), edited by H. Jaaksoo and V. Utkin, IFAC Proceedings Series 1991, No. 1, pp. 293–298.

VERGHESE, G. C., and KAILATH, T., 1981, Rational matrix structure. *I.E.E.E. Transactions on Automatic Control*, **26**, 434–439.

WILLEMS, J. C., 1986, From time series to linear system—Part I: Finite dimensional linear time-invariant systems. *Automatica*, **22**, 561–580; 1988, Models for dynamics. *Dynamics Reported*, Vol. 2, edited by U. Kirchgraber, and H. O. Walther (Wiley and Teubner), pp. 171–269.

WOLOVICH, W. A., 1974, *Linear Multivariable Systems*. Applied Mathematical Sciences, Vol. 11 (New York: Springer Verlag).

WOLOVICH, W. A., and ELLIOTT, H., 1983, Discrete models for linear multivariable systems. *International Journal of Control*, **38**, 337–357.