Linear Algebra and its Applications 437 (2012) 783–808

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

From polynomial matrices to Markov parameters and back: Theory and numerical algorithms

Matthew S. Holzel *, Dennis S. Bernstein

Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, United States

ARTICLE INFO

Article history:
Received 10 June 2011
Accepted 2 March 2012
Available online 12 April 2012

Submitted by V. Mehrmann

AMS classification:
93A30
15A22

Keywords:
Linear systems
Polynomial matrices
Markov parameters
Realization theory

ABSTRACT

We consider polynomial matrix representations of MIMO linear systems and their connection to Markov parameters. Specifically, we consider polynomial matrix models in an arbitrary operator $\boldsymbol{\rho}$, and develop theory and numerical algorithms for transforming polynomial matrix models into Markov parameter models, and vice versa. We also provide numerical examples to illustrate the proposed algorithms.

1. Introduction

Polynomial matrix and state-space models provide alternative and complementary parametric representations for multivariable linear systems, with transfer function models providing an easy-to-work-with link between the two [1,18,22,23]. Similarly, frequency response models and Markov parameter models provide additional, albeit nonparametric, representations for the same systems [16,24].

The subject of realization theory then, is to transform one type of model into another [1,10,19]. For example, the transformation from a state-space model $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ to a polynomial matrix model (E, F) is given by $E(\boldsymbol{\rho}) = \det[\boldsymbol{\rho} I - \tilde{A}]I$ and $F(\boldsymbol{\rho}) = \tilde{C}\mathrm{adj}\,[\boldsymbol{\rho} I - \tilde{A}]\tilde{B} + \tilde{D}$, and the transformation from a Markov parameter model to a state-space model is well established by the Ho–Kalman algorithm [1,7,9]. Furthermore, several of these transformations turn out to be operator-invariant. For instance,

* Corresponding author.
E-mail addresses: mholzel@umich.edu (M.S. Holzel), dsbaero@umich.edu (D.S. Bernstein).

http://dx.doi.org/10.1016/j.laa.2012.03.006

the Markov parameters of the state-space model $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ in the differentiation operator **d/dt** are the same as the Markov parameters of the state-space model $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ in the forward shift operator **q**, even though the input–output behavior will in general depend greatly on the operator involved. As we show later, this "operator"-invariance property will hold for the transformations between polynomial matrix models and Markov parameters, and vice versa. This allows us to define the transformations with respect to an arbitrary operator $\boldsymbol{\rho}$, which may aid in the analysis of MIMO models in nonclassical operators (those which are not polynomic in either **d/dt** or **q**) such as δ-domain [14] or fractional-order models [24].

However, although many of these transformations are theoretically understood, some, such as transformation from a state-space model to a polynomial matrix model, may not be easy to compute, and numerical (rather than symbolic) algorithms are needed. To this end, [9] provides a robust numerical link between the Markov parameter and state-space models in the form of the eigensystem realization algorithm, which utilizes the singular value decomposition and Ho–Kalman algorithm to construct a minimal state-space model from a sufficient number of Markov parameters. Similarly, other authors have developed numerical approaches to realization theory, such as [21], although most of the available literature tends to fall into the broad class of system identification, that is, numerical algorithms for transforming input/output data into a given model type [3,12,16,20].

The goal of the present paper is to develop the numerical and theoretical link between polynomial matrix and Markov parameter models, so as to provide a complete picture of the interrelationships between different linear system representations. Furthermore, this work is important in several modern control areas, such as adaptive control [6,8] and model predictive control [4,13,15], where the use of polynomial matrix models is still preferred over state-space models and where system identification may only yield Markov parameters of the system and not the polynomial matrix system directly [3,17].

The development of the numerical and theoretical link between polynomial matrix and Markov parameter models is carried out entirely within the context of polynomial matrices without the use of rational functions; consequently, rational transfer functions do not appear. This approach removes the need to explicitly discuss poles and zeros, singularities, and cancellations, thus allowing us to focus on the essential algebraic structure of the problem in terms of polynomial matrices. Furthermore, the algorithms that we develop do not depend on symbolic computations, but rather are entirely numerical. This approach circumvents possible ill-conditioning that can arise in symbolic computations that depend on exact cancellation of the coefficients of operator powers.

The contents of the paper are as follows. First, we present the necessary preliminaries concerning polynomial matrices. Next, after introducing the problem statement, we discuss the theoretical relationship between polynomial matrices and Markov parameters. Finally, we present several numerical algorithms for transforming Markov parameter models into polynomial matrix models and vice versa, followed by numerical examples, and our conclusions.

2. Polynomial matrices

In this section, we introduce polynomial matrices in an arbitrary operator $\boldsymbol{\rho}$, employing the standard notation

$$\boldsymbol{\rho}^2 y(t) = \boldsymbol{\rho}\left[\boldsymbol{\rho}\left[y(t)\right]\right],$$

and so on, where $\boldsymbol{\rho}\left[y(t)\right]$ represents the signal that results from $\boldsymbol{\rho}$ operating on the signal y. For a complete treatment of matrices, polynomial matrices, and realization theory, refer to any of the excellent books [2,5,11,23].

Remark 2.1. Alternatively, throughout the paper, one could view $\boldsymbol{\rho}$ as an indeterminate. However, in this case, definitions such as that of a causal system (Definition 2.23) have no physical meaning.

We begin by introducing infinite polynomial matrices, or polynomial matrix expansions, since polynomial matrices can be viewed as a special case of infinite polynomial matrices.

Definition 2.2. Let $G_0, G_1, G_2, \ldots \in \mathbb{R}^{p\times m}$ and

$$G(\boldsymbol{\rho}) \triangleq \sum_{i=0}^{\infty} G_i \boldsymbol{\rho}^i. \tag{1}$$

Then we denote $G \in \mathbb{R}_{\infty}^{p\times m}[\boldsymbol{\rho}]$. Furthermore, by convention, $G(\boldsymbol{\rho}) = 0_{p\times m}$ if and only if $G_i = 0_{p\times m}$ for all $i \geqslant 0$.

Definition 2.3. Let $C_0, C_1, \ldots, C_s \in \mathbb{R}^{p\times m}$ and

$$C(\boldsymbol{\rho}) \triangleq C_0 + C_1\boldsymbol{\rho} + \cdots + C_s\boldsymbol{\rho}^s. \tag{2}$$

Then we denote $C \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$. Furthermore,

(i) We say that $C(\boldsymbol{\rho})$ is *diagonal* if $m = p$ and $C_0, \ldots, C_s$ are diagonal. If, in addition, there exists $\eta \in \mathbb{R}[\boldsymbol{\rho}]$ such that $C(\boldsymbol{\rho}) = \eta(\boldsymbol{\rho})I_p$, then $C(\boldsymbol{\rho})$ is *quasi-scalar*.

(ii) We say that C_j is the *trailing coefficient of* $C(\boldsymbol{\rho})$ if C_j is nonzero and $C_0 = \cdots = C_{j-1} = 0_{p\times m}$. If, in addition, $m = p$ and C_j is nonsingular, then we say that $C(\boldsymbol{\rho})$ is *regular*. If, in addition, $C_j = I_p$, then we say that $C(\boldsymbol{\rho})$ is *comonic*.

Remark 2.4. Given $F \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ or $F \in \mathbb{R}_{\infty}^{p\times m}[\boldsymbol{\rho}]$, we sometimes refer to F_i without explicitly defining a form for $F(\boldsymbol{\rho})$ such as (1) or (2). It should be clear that F_i refers to the ith coefficient matrix of $F(\boldsymbol{\rho})$, that is, the coefficient matrix which multiplies $\boldsymbol{\rho}^i$.

Next, note that for all $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$, the determinant and adjugate of $C(\boldsymbol{\rho})$ can be computed with addition, subtraction, and multiplication operations. Hence $\det[C(\boldsymbol{\rho})] \in \mathbb{R}[\boldsymbol{\rho}]$ and $\operatorname{adj}[C(\boldsymbol{\rho})] \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$.

Definition 2.5. Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$. Then $C(\boldsymbol{\rho})$ has *full normal rank* if $\det[C(\boldsymbol{\rho})] \neq 0$.

Fact 2.6. *Let $C, E \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $F(\boldsymbol{\rho}) \triangleq C(\boldsymbol{\rho})E(\boldsymbol{\rho})$. Then $F(\boldsymbol{\rho})$ has full normal rank if and only if $C(\boldsymbol{\rho})$ and $E(\boldsymbol{\rho})$ have full normal rank.*

Proof. $\det[F(\boldsymbol{\rho})] = \det[C(\boldsymbol{\rho})]\det[E(\boldsymbol{\rho})]$. □

Fact 2.7. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ have full normal rank and let $G, H \in \mathbb{R}_{\infty}^{p\times m}[\boldsymbol{\rho}]$. Then $C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = 0_{p\times m}$ if and only if $G(\boldsymbol{\rho}) = 0_{p\times m}$. Furthermore, $C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = C(\boldsymbol{\rho})H(\boldsymbol{\rho})$ if and only if $G(\boldsymbol{\rho}) = H(\boldsymbol{\rho})$.*

Proof. First, note that $\det\big[C(\boldsymbol{\rho})\big]$ is nonzero since $C(\boldsymbol{\rho})$ has full normal rank. Also, let α_i be the trailing coefficient of $\det\big[C(\boldsymbol{\rho})\big]$.

Next, let $C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = 0_{p\times m}$. Then

$$F(\boldsymbol{\rho}) \triangleq \operatorname{adj}[C(\boldsymbol{\rho})]C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = \det\big[C(\boldsymbol{\rho})\big]G(\boldsymbol{\rho}) = 0_{p\times m},$$

and hence $F_i = \alpha_i G_0 = 0_{p\times m}$. However, since α_i is nonzero, $G_0 = 0_{p\times m}$. Furthermore, since $G_0 = 0_{p\times m}$, it follows that $F_{i+1} = \alpha_i G_1 = 0_{p\times m}$, and therefore $G_1 = 0_{p\times m}$. Hence, by induction we have that $G_i = 0_{p\times m}$ for all $i \geqslant 0$, that is, $G(\boldsymbol{\rho}) = 0_{p\times m}$.

Third, let $G(\boldsymbol{\rho}) = 0_{p\times m}$. Then $C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = 0_{p\times m}$ follows immediately. Similarly, if $G(\boldsymbol{\rho}) = H(\boldsymbol{\rho})$, then $C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = C(\boldsymbol{\rho})H(\boldsymbol{\rho})$.

Finally, let $C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = C(\boldsymbol{\rho})H(\boldsymbol{\rho})$. Then

$$C(\boldsymbol{\rho})\,[G(\boldsymbol{\rho}) - H(\boldsymbol{\rho})] = 0_{p\times m},$$

and hence, as we already showed, $G(\boldsymbol{\rho}) - H(\boldsymbol{\rho}) = 0_{p\times m}$, that is, $G(\boldsymbol{\rho}) = H(\boldsymbol{\rho})$. □

Fact 2.8. *If $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ is regular, then $C(\boldsymbol{\rho})$ has full normal rank.*

Proof. Let C_j be the trailing coefficient of $C(\boldsymbol{\rho})$ and let $C(\boldsymbol{\rho}) \triangleq \boldsymbol{\rho}^j C'(\boldsymbol{\rho})$ where

$$C'(\boldsymbol{\rho}) \triangleq C_j + C_{j+1}\boldsymbol{\rho} + \cdots + C_{j+k}\boldsymbol{\rho}^k.$$

Next, since $C(\boldsymbol{\rho})$ is regular, C_j is nonsingular. Hence

$$\det\left[C'(0)\right] = \det\left[C_j\right] \neq 0,$$

and thus $C'(\boldsymbol{\rho})$ has full normal rank.

Finally, since $\boldsymbol{\rho}^j$ has full normal rank, from Fact 2.6, we have that $C(\boldsymbol{\rho})$ has full normal rank. □

Fact 2.9. *If $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ is nonzero and quasi-scalar, then $C(\boldsymbol{\rho})$ has full normal rank.*

Proof. Since $C(\boldsymbol{\rho})$ is nonzero and quasi-scalar, $C(\boldsymbol{\rho})$ is regular. Hence, from Fact 2.8, $C(\boldsymbol{\rho})$ has full normal rank. □

Definition 2.10. Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$. Then $C(\boldsymbol{\rho})$ is *unimodular* if there exists $E \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ such that $E(\boldsymbol{\rho})C(\boldsymbol{\rho}) = I_p$.

Remark 2.11. Equivalently, from Definition 2.10, we have that $C(\boldsymbol{\rho})$ is unimodular if and only if $\det[C(\boldsymbol{\rho})]$ is a nonzero constant.

Definition 2.12. Let $C, L \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$. Then

- (i) $L(\boldsymbol{\rho})$ is a *left factor of* (C, D) if there exist $E \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $F \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ such that $C(\boldsymbol{\rho}) = L(\boldsymbol{\rho})E(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho}) = L(\boldsymbol{\rho})F(\boldsymbol{\rho})$.
- (ii) $L(\boldsymbol{\rho})$ is a *greatest left factor of* (C, D) if $L(\boldsymbol{\rho})$ is a left factor of (C, D) and, for every left factor $L'(\boldsymbol{\rho})$ of (C, D), there exists $U \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ such that $L(\boldsymbol{\rho}) = L'(\boldsymbol{\rho})U(\boldsymbol{\rho})$.
- (iii) $L(\boldsymbol{\rho})$ is a *greatest quasi-scalar factor of* (C, D) if $L(\boldsymbol{\rho})$ is a quasi-scalar left factor of (C, D) and, for every quasi-scalar left factor $L'(\boldsymbol{\rho})$ of (C, D), there exists $\eta \in \mathbb{R}[\boldsymbol{\rho}]$ such that $L(\boldsymbol{\rho}) = L'(\boldsymbol{\rho})\eta(\boldsymbol{\rho})$.
- (iv) (C, D) is *left coprime* if every left factor of (C, D) is unimodular.

Analogous definitions apply for right factors, greatest right factors, and right coprime.

Note that, when referring to a pair (C, D), we drop the argument $\boldsymbol{\rho}$, for conciseness. Also, note that for every (C, D), there exist greatest left and right factors of (C, D) [22].

Fact 2.13. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$. The zero polynomial is a left factor of (C, D) if and only if $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ are both zero.*

Proof. First, let $L(\boldsymbol{\rho}) = 0_{p\times p}$ be a left factor of (C, D). Then there exist $C' \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D' \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ such that $C(\boldsymbol{\rho}) = L(\boldsymbol{\rho})C'(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho}) = L(\boldsymbol{\rho})D'(\boldsymbol{\rho})$. However, since $L(\boldsymbol{\rho})C'(\boldsymbol{\rho}) = 0_{p\times p}$ and $L(\boldsymbol{\rho})D'(\boldsymbol{\rho}) = 0_{p\times m}$ for every $C' \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D' \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, it follows that $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ are both zero.

Second, let $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ both be zero. Then for every $C' \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D' \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, $C(\boldsymbol{\rho}) = [0_{p\times p}]C'(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho}) = [0_{p\times p}]D'(\boldsymbol{\rho})$. Hence $L(\boldsymbol{\rho}) = 0_{p\times p}$ is a left factor of (C, D). □

Fact 2.14. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$, where $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ are not both zero. Then the greatest comonic quasi-scalar factor of (C, D) is unique.*

Proof. First, since $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ are not both zero, then from Fact 2.13, the zero polynomial is not a left factor of (C, D). Hence greatest quasi-scalar factors of (C, D) are nonzero.

Next, let $L, L' \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ be greatest comonic quasi-scalar factors of (C, D). Then $L(\boldsymbol{\rho})$ and $L'(\boldsymbol{\rho})$ are nonzero, and from Definition 2.12, there exist $\eta, \mu \in \mathbb{R}[\boldsymbol{\rho}]$ such that $L(\boldsymbol{\rho}) = L'(\boldsymbol{\rho})\eta(\boldsymbol{\rho})$ and $L'(\boldsymbol{\rho}) = L(\boldsymbol{\rho})\mu(\boldsymbol{\rho})$. Furthermore, $\eta(\boldsymbol{\rho})$ and $\mu(\boldsymbol{\rho})$ are nonzero since $L(\boldsymbol{\rho})$ and $L'(\boldsymbol{\rho})$ are nonzero.

Third, note that $L(\boldsymbol{\rho}) = \eta(\boldsymbol{\rho})\mu(\boldsymbol{\rho})L(\boldsymbol{\rho})$. Furthermore, since $L(\boldsymbol{\rho})$ is nonzero and quasi-scalar, then from Fact 2.9, $L(\boldsymbol{\rho})$ has full normal rank, and from Fact 2.7, it follows that $\eta(\boldsymbol{\rho})\mu(\boldsymbol{\rho}) = 1$. Hence $\eta, \mu \in \mathbb{R}$.

Finally, since $L(\boldsymbol{\rho})$ and $L'(\boldsymbol{\rho})$ are both comonic, it follows that $\eta(\boldsymbol{\rho}) = \mu(\boldsymbol{\rho}) = 1$, that is, $L(\boldsymbol{\rho}) = L'(\boldsymbol{\rho})$. Thus the greatest comonic quasi-scalar factor of (C, D) is unique. $\square$

Definition 2.15. Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$. Then the *principal factor of* (C, D) is I_p if $C(\boldsymbol{\rho}) = 0_{p\times p}$ and $D(\boldsymbol{\rho}) = 0_{p\times m}$. Otherwise, the *principal factor of* (C, D) is the greatest comonic quasi-scalar factor of (C, D).

Definition 2.16. Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and let $D(\boldsymbol{\rho}) \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ be the principal factor of $(\text{adj}\,[C], \det[C]I_p)$. Then $E(\boldsymbol{\rho})$ is the *minimal adjugate of* $C(\boldsymbol{\rho})$ if $\text{adj}\,[C(\boldsymbol{\rho})] = D(\boldsymbol{\rho})E(\boldsymbol{\rho})$, and $\beta(\boldsymbol{\rho})$ is the *minimal determinant of* $C(\boldsymbol{\rho})$ if $\det\,[C(\boldsymbol{\rho})]I_p = D(\boldsymbol{\rho})\beta(\boldsymbol{\rho})$. Specifically, we write

$$\begin{aligned}\text{adj}\,[C(\boldsymbol{\rho})] &= D(\boldsymbol{\rho})\text{madj}\,[C(\boldsymbol{\rho})],\\ \det\,[C(\boldsymbol{\rho})]I_p &= D(\boldsymbol{\rho})\text{mdet}\,[C(\boldsymbol{\rho})].\end{aligned}$$

Fact 2.17. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$. Then the minimal adjugate and minimal determinant of $C(\boldsymbol{\rho})$ are unique.*

Proof. Let $D(\boldsymbol{\rho})$ be the principal factor of $(\text{adj}\,[C], \det[C]I_p)$ and suppose that $E, F \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ are both minimal adjugates of $C(\boldsymbol{\rho})$. Then $\text{adj}\,[C(\boldsymbol{\rho})] = D(\boldsymbol{\rho})E(\boldsymbol{\rho}) = D(\boldsymbol{\rho})F(\boldsymbol{\rho})$. Furthermore, since the principal factor $D(\boldsymbol{\rho})$ is defined to be comonic, it is nonzero. Therefore, from Fact 2.9, $D(\boldsymbol{\rho})$ has full normal rank, and it follows from Fact 2.7 that $E(\boldsymbol{\rho}) = F(\boldsymbol{\rho})$, that is, the minimal adjugate is unique.

Similarly, if $\beta, \gamma \in \mathbb{R}[\boldsymbol{\rho}]$ are both minimal determinants of $C(\boldsymbol{\rho})$, then $\det\,[C(\boldsymbol{\rho})]I_p = D(\boldsymbol{\rho})\beta(\boldsymbol{\rho}) = D(\boldsymbol{\rho})\gamma(\boldsymbol{\rho})$, and from Fact 2.7, $\beta(\boldsymbol{\rho}) = \gamma(\boldsymbol{\rho})$, that is, the minimal determinant is unique. $\square$

Remark 2.18. $\det[C(\boldsymbol{\rho})] = \text{mdet}\,[C(\boldsymbol{\rho})] = 0$ if and only if $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ does not have full normal rank.

Example 1. Let $\alpha, \beta, \gamma \in \mathbb{R}[\boldsymbol{\rho}]$ be nonzero, let α_i be the trailing coefficient of $\alpha(\boldsymbol{\rho})$, and let (β, γ) be left coprime, that is, $\beta(\boldsymbol{\rho})$ and $\gamma(\boldsymbol{\rho})$ have no common zeros. Finally, let

$$E(\boldsymbol{\rho}) \triangleq \alpha(\boldsymbol{\rho})\begin{bmatrix}\beta(\boldsymbol{\rho}) & 0\\ 0 & \gamma(\boldsymbol{\rho})\end{bmatrix}.$$

Then

(1) The determinant of $E(\boldsymbol{\rho})$ is given by

$$\det\,[E(\boldsymbol{\rho})] = \alpha^2(\boldsymbol{\rho})\beta(\boldsymbol{\rho})\gamma(\boldsymbol{\rho}).$$

(2) The adjugate of $E(\boldsymbol{\rho})$ is given by

$$\text{adj}\,[E(\boldsymbol{\rho})] = \alpha(\boldsymbol{\rho})\begin{bmatrix}\gamma(\boldsymbol{\rho}) & 0\\ 0 & \beta(\boldsymbol{\rho})\end{bmatrix}.$$

(3) The principal factor of (adj $[E]$, $\det[E]I_p$) is given by

$$\left[\frac{1}{\alpha_i}\right]\alpha(\boldsymbol{\rho})I_p.$$

(4) The minimal determinant of $E(\boldsymbol{\rho})$ is given by

$$\operatorname{mdet}\,[E(\boldsymbol{\rho})] = \alpha_i\alpha(\boldsymbol{\rho})\beta(\boldsymbol{\rho})\gamma(\boldsymbol{\rho}).$$

(5) The minimal adjugate of $E(\boldsymbol{\rho})$ is given by

$$\operatorname{madj}\left[E(\boldsymbol{\rho})\right] = \alpha_i\begin{bmatrix}\gamma(\boldsymbol{\rho}) & 0\\ 0 & \beta(\boldsymbol{\rho})\end{bmatrix}.$$

Definition 2.19. Let $C, E \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D, F \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$. Then (E, F) is a *multiple of* (C, D) if there exists $L \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ with full normal rank such that $E(\boldsymbol{\rho}) = L(\boldsymbol{\rho})C(\boldsymbol{\rho})$ and $F(\boldsymbol{\rho}) = L(\boldsymbol{\rho})D(\boldsymbol{\rho})$. Furthermore,

(i) (E, F) is a *comonic multiple of* (C, D) if $E(\boldsymbol{\rho})$ is comonic.
(ii) (E, F) is a *quasi-scalar multiple of* (C, D) if $E(\boldsymbol{\rho})$ is quasi-scalar.

Remark 2.20. A quasi-scalar multiple is analogous to a transfer function representation of a MIMO system since the system can be written as a rational polynomial matrix. For instance, if $(\alpha I_p, F)$ is a quasi-scalar multiple of (C, D), where $\alpha \in \mathbb{R}[\boldsymbol{\rho}]$ and

$$F(\boldsymbol{\rho}) \triangleq \begin{bmatrix} f_{1,1}(\boldsymbol{\rho}) & \cdots & f_{1,m}(\boldsymbol{\rho})\\ \vdots & & \vdots\\ f_{p,1}(\boldsymbol{\rho}) & \cdots & f_{p,m}(\boldsymbol{\rho})\end{bmatrix} \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}],$$

then a transfer function representation of the system (C, D) would be

$$\begin{bmatrix} \dfrac{f_{1,1}(\boldsymbol{\rho})}{\alpha(\boldsymbol{\rho})} & \cdots & \dfrac{f_{1,m}(\boldsymbol{\rho})}{\alpha(\boldsymbol{\rho})}\\ \vdots & & \vdots\\ \dfrac{f_{p,1}(\boldsymbol{\rho})}{\alpha(\boldsymbol{\rho})} & \cdots & \dfrac{f_{p,m}(\boldsymbol{\rho})}{\alpha(\boldsymbol{\rho})}\end{bmatrix}.$$

Note however, that we have made no definition of the meaning $1/\boldsymbol{\rho}$, and one must be particularly careful in defining rational functions of operators since in general, an operator is not a one-to-one mapping.

Definition 2.21. Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$. Also, let s be the smallest nonnegative integer such that $C(\boldsymbol{\rho})$ is of the form (2). Then the *degree of* $C(\boldsymbol{\rho})$ is s if $C(\boldsymbol{\rho})$ is nonzero, and $-\infty$ if $C(\boldsymbol{\rho})$ is zero. Finally, let s be the degree of $C(\boldsymbol{\rho})$ and let t be the degree of $D(\boldsymbol{\rho})$. Then the *degree of* (C, D) is $\max(s, t)$.

Next we show that the minimal adjugate provides us with a quasi-scalar multiple $(C^\star, D^\star)$ of (C, D) with the lowest possible degree, where $C^\star(\boldsymbol{\rho})$ is the minimal determinant of $C(\boldsymbol{\rho})$.

Proposition 2.22. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ have full normal rank, $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, $E(\boldsymbol{\rho}) \triangleq \operatorname{madj}[C(\boldsymbol{\rho})]$, $\beta(\boldsymbol{\rho}) \triangleq \operatorname{mdet}[C(\boldsymbol{\rho})]$, and β_i be the trailing coefficient of $\beta(\boldsymbol{\rho})$. Then*

$$(C^\star, D^\star) \triangleq (EC/\beta_i, ED/\beta_i) = (\beta/\beta_i I_p, ED/\beta_i)$$

is the unique comonic quasi-scalar multiple of (C, D) of the lowest degree.

Proof. First, since $C(\boldsymbol{\rho})$ has full normal rank, $\beta(\boldsymbol{\rho}) \neq 0$. Hence, from Fact 2.9, $\beta(\boldsymbol{\rho})I_p$ has full normal rank. Furthermore, since $\beta(\boldsymbol{\rho})I_p$ has full normal rank and $E(\boldsymbol{\rho})C(\boldsymbol{\rho}) = \beta(\boldsymbol{\rho})I_p$, then from Fact 2.6, $E(\boldsymbol{\rho})$ also has full normal rank. Hence $(EC, ED) = (\beta I_p, ED)$ is a quasi-scalar multiple of (C, D).

Next, from Fact 2.17, the minimal adjugate is unique. Hence (FC, FD) is a quasi-scalar multiple of (C, D) if and only if there exists a nonzero $\mu \in \mathbb{R}[\boldsymbol{\rho}]$ such that $F(\boldsymbol{\rho}) = \mu(\boldsymbol{\rho})E(\boldsymbol{\rho})$. Furthermore, if the degree of $\mu(\boldsymbol{\rho})$ is greater than zero, the degree of (FC, FD) is greater than the degree of $(C^\star, D^\star)$. Hence $(C^\star, D^\star)$ is the unique comonic quasi-scalar multiple of (C, D) of the lowest degree. □

Definition 2.23. Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$. Also, assume there exists $G \in \mathbb{R}^{p\times m}_\infty[\boldsymbol{\rho}]$ such that

$$C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = D(\boldsymbol{\rho}).$$

Then (C, D) is *causal*, $G(\boldsymbol{\rho})$ is a *Markov parameter polynomial of* (C, D), and G_i is an ith *Markov parameter of* (C, D).

Remark 2.24. In Section 8, we show that this definition of Markov parameters is consistent with the usual state-space definition of Markov parameters.

Remark 2.25. Whether or not a system

$$C(\boldsymbol{\rho})y(t) = D(\boldsymbol{\rho})u(t) \tag{3}$$

is causal in the sense that $y(t)$ is a function only of $u(\tau)$ for $\tau \leqslant t$, is dependent on the operator $\boldsymbol{\rho}$. For $\boldsymbol{\rho} \triangleq \mathbf{q}^{-1}$, this sense of causality and Definition 2.23 are equivalent. However, for $\boldsymbol{\rho} \triangleq \mathbf{d/dt}$ or $\boldsymbol{\rho} \triangleq \mathbf{q}$, the two definitions are not equivalent. For these operators, one would say that a system (3) is causal in the classical sense if there exists $G \in \mathbb{R}^{p\times m}_\infty[\boldsymbol{\rho}]$ such that

$$C(1/\boldsymbol{\rho})G(\boldsymbol{\rho}) = D(1/\boldsymbol{\rho}).$$

Fact 2.26. *Let $C, E \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $D, F \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$. Also, let (E, F) be a multiple of (C, D). Then (E, F) is causal if and only if (C, D) is causal.*

Proof. First, since (E, F) is a multiple of (C, D), there exists $L \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ with full normal rank such that $(E, F) = (LC, LD)$.

Next, let (E, F) be causal. Then there exists $G \in \mathbb{R}^{p\times m}_\infty[\boldsymbol{\rho}]$ such that

$$E(\boldsymbol{\rho})G(\boldsymbol{\rho}) = L(\boldsymbol{\rho})C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = L(\boldsymbol{\rho})D(\boldsymbol{\rho}) = F(\boldsymbol{\rho}).$$

Hence from Fact 2.7, $C(\boldsymbol{\rho})G(\boldsymbol{\rho}) = D(\boldsymbol{\rho})$, that is, (C, D) is causal.

Finally, let (C, D) be causal. Then there exists $G' \in \mathbb{R}^{p\times m}_\infty[\boldsymbol{\rho}]$ such that $C(\boldsymbol{\rho})G'(\boldsymbol{\rho}) = D(\boldsymbol{\rho})$. Hence

$$L(\boldsymbol{\rho})C(\boldsymbol{\rho})G'(\boldsymbol{\rho}) = E(\boldsymbol{\rho})G'(\boldsymbol{\rho}) = F(\boldsymbol{\rho}) = L(\boldsymbol{\rho})D(\boldsymbol{\rho}),$$

and therefore, (E, F) is causal. □

Fact 2.27. *Let $E \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ be comonic, let $F \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, and let*

$$E(\boldsymbol{\rho}) = I_p\boldsymbol{\rho}^{\ell} + E_{\ell+1}\boldsymbol{\rho}^{\ell+1} + \cdots + E_s\boldsymbol{\rho}^{s},$$
$$F(\boldsymbol{\rho}) = F_0 + F_1\boldsymbol{\rho} + \cdots + F_s\boldsymbol{\rho}^{s},$$

where $0 \leqslant \ell \leqslant s$. If (E, F) is causal, then $F_0 = \cdots = F_{\ell-1} = 0_{p\times m}$ and the Markov parameter polynomial of (E, F) is given later in Theorem 4.1.

Proof. Since (E, F) is causal, there exists $G \in \mathbb{R}_{\infty}^{p\times m}[\boldsymbol{\rho}]$ such that $E(\boldsymbol{\rho})G(\boldsymbol{\rho}) = F(\boldsymbol{\rho})$. Hence, computing the product $E(\boldsymbol{\rho})G(\boldsymbol{\rho})$, it follows that $F_0 = \cdots = F_{\ell-1} = 0_{p\times m}$. □

Fact 2.28. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ have full normal rank, let $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, and let (C, D) be left coprime and causal. Then C_0 is nonsingular.*

Proof. First, from Proposition 2.22, there exists a comonic multiple $(E, F) = (LC, LD)$ of (C, D). Furthermore, from Fact 2.26, since (C, D) is causal, (E, F) is causal. Hence, letting E_ℓ denote the trailing coefficient of $E(\boldsymbol{\rho})$, from Fact 2.27 we have that $F_0 = \cdots = F_{\ell-1} = 0_{p\times m}$. Therefore $\boldsymbol{\rho}^{\ell}I_p$ is a left factor of (E, F).

Next, since (C, D) is left coprime, $L(\boldsymbol{\rho})$ is a greatest left factor of (E, F). Hence there exists $L' \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ such that $L(\boldsymbol{\rho}) = \boldsymbol{\rho}^{\ell}L'(\boldsymbol{\rho})$.

Finally, letting $(E', F') \triangleq (L'C, L'D)$, we have that $E'_0 = I_p$. Furthermore, since $E'_0 = L'_0C_0 = I_p$, it follows that C_0 is nonsingular. □

Fact 2.29. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ have full normal rank and let $C_0 \in \mathbb{R}^{p\times p}$ be nonsingular. Also, let*

$$\beta(\boldsymbol{\rho}) \triangleq \text{mdet}\,[C(\boldsymbol{\rho})] = \beta_0 + \beta_1\boldsymbol{\rho} + \cdots + \beta_s\boldsymbol{\rho}^{s}.$$

Then β_0 is nonzero.

Proof. Let $D(\boldsymbol{\rho}) = \mu(\boldsymbol{\rho})I_p$ denote the principal factor of $(\text{adj}\,[C], \det[C]I_p)$, where $\mu \in \mathbb{R}[\boldsymbol{\rho}]$. Then $\det[C(\boldsymbol{\rho})] = \mu(\boldsymbol{\rho})\beta(\boldsymbol{\rho})$. Furthermore, since C_0 is nonsingular, we have that

$$0 \neq \det[C_0] = \det[C(0)] = \mu(0)\beta(0).$$

Hence $\beta_0 = \beta(0)$ is nonzero. □

3. Problem formulation

Consider the linear time-invariant system

$$A(\boldsymbol{\rho})y(t) = B(\boldsymbol{\rho})u(t), \tag{4}$$

where $\boldsymbol{\rho}$ is an operator, $A \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ has full normal rank, $B \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, $y \in \mathbb{R}^p$ is the output, $u \in \mathbb{R}^m$ is the input, (A, B) is left coprime and causal, and (4) holds for all $t \in T$. Also, let $(A^\star, B^\star)$ be the unique comonic quasi-scalar multiple of (A, B) given by Proposition 2.22, and let

$$A(\boldsymbol{\rho}) \triangleq A_0 + A_1\boldsymbol{\rho} + \cdots + A_n\boldsymbol{\rho}^{n},$$
$$B(\boldsymbol{\rho}) \triangleq B_0 + B_1\boldsymbol{\rho} + \cdots + B_n\boldsymbol{\rho}^{n},$$
$$A^\star(\boldsymbol{\rho}) \triangleq I_p + a_1^\star I_p\boldsymbol{\rho} + \cdots + a_{n^\star}^\star I_p\boldsymbol{\rho}^{n^\star},$$
$$B^\star(\boldsymbol{\rho}) \triangleq B_0^\star + B_1^\star\boldsymbol{\rho} + \cdots + B_{n^\star}^\star\boldsymbol{\rho}^{n^\star},$$

where $a_1^\star, \ldots, a_{n^\star}^\star \in \mathbb{R}$. This notation is assumed for the rest of the paper.

Throughout the paper, we have two objectives in mind, namely

(1) Given a not necessarily coprime multiple of (A, B), compute the Markov parameters of (A, B).
(2) Given a sufficient number of the Markov parameters of (A, B), compute a multiple of (A, B).

We show how to obtain both of these objectives numerically.

Remark 3.1. The trailing coefficient of $A^{\star}(\boldsymbol{\rho})$ is the identity as a result of Proposition 2.22, Fact 2.28, and Fact 2.29.

Remark 3.2. Let (C, D) be a multiple of (A, B). Since $A(\boldsymbol{\rho})$ has full normal rank, then from Fact 2.6, $C(\boldsymbol{\rho})$ also has full normal rank.

4. Markov parameters

In this section, we develop Markov parameters algebraically from polynomial matrices. Furthermore, we show that the Markov parameters of (A, B) and the Markov parameters of every multiple of (A, B) are equal and unique.

Theorem 4.1. *Let (E, F) be a comonic multiple of (A, B) given by*

$$E(\boldsymbol{\rho}) \triangleq I_p\boldsymbol{\rho}^{\ell} + E_{\ell+1}\boldsymbol{\rho}^{\ell+1} + \cdots + E_s\boldsymbol{\rho}^{s},$$

$$F(\boldsymbol{\rho}) \triangleq F_0 + F_1\boldsymbol{\rho} + \cdots + F_s\boldsymbol{\rho}^{s},$$

where $0 \leqslant \ell \leqslant s$. Also, let $H(\boldsymbol{\rho}) \triangleq \sum_{i=0}^{\infty} H_i\boldsymbol{\rho}^i$, where, for all $i \geqslant 0$,

$$H_i \triangleq F_{\ell+i} - \sum_{j=1}^{\min(s-\ell,i)} E_{\ell+j}H_{i-j}, \tag{5}$$

and $F_j \triangleq 0_{p\times m}$ for $j > s$. Then $H(\boldsymbol{\rho})$ is the Markov parameter polynomial of (A, B) and every multiple of (A, B).

Proof. First, from Proposition 2.22, there exists a comonic multiple of (A, B). Furthermore, recalling that (A, B) is causal, Fact 2.26 implies that (E, F) is also causal. Hence, from Fact 2.27, $F_0 = \cdots = F_{\ell-1} = 0_{p\times m}$.

Next, let $G(\boldsymbol{\rho}) \triangleq E(\boldsymbol{\rho})H(\boldsymbol{\rho})$. Then for $i \geqslant 0$, we have that

$$G_{\ell+i} = H_i + \sum_{j=1}^{\min(s-\ell,i)} E_{\ell+j}H_{i-j}. \tag{6}$$

Hence from (5) and (6), it follows that

$$G_i = \begin{cases} F_i, & \ell \leqslant i \leqslant s, \\ 0_{p\times m}, & \text{otherwise,} \end{cases}$$

that is, $G(\boldsymbol{\rho}) = F(\boldsymbol{\rho})$. Therefore $E(\boldsymbol{\rho})H(\boldsymbol{\rho}) = F(\boldsymbol{\rho})$, and thus $H(\boldsymbol{\rho})$ is a Markov parameter polynomial of (E, F).

Next, since (E, F) is a multiple of (A, B), there exists $L \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ with full normal rank such that $E(\boldsymbol{\rho}) = L(\boldsymbol{\rho})A(\boldsymbol{\rho})$ and $F(\boldsymbol{\rho}) = L(\boldsymbol{\rho})B(\boldsymbol{\rho})$. Hence

$$E(\boldsymbol{\rho})H(\boldsymbol{\rho}) = L(\boldsymbol{\rho})A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = L(\boldsymbol{\rho})B(\boldsymbol{\rho}) = F(\boldsymbol{\rho}),$$

and from Fact 2.7, $A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B(\boldsymbol{\rho})$. Thus $H(\boldsymbol{\rho})$ is a Markov parameter polynomial of (A, B). Furthermore, if $H'' \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ is also a Markov parameter polynomial of (A, B), then $A(\boldsymbol{\rho})H''(\boldsymbol{\rho}) = B(\boldsymbol{\rho})$, and hence $A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = A(\boldsymbol{\rho})H''(\boldsymbol{\rho})$. Thus from Fact 2.7, $H''(\boldsymbol{\rho}) = H(\boldsymbol{\rho})$, and it follows that $H(\boldsymbol{\rho})$ is the unique Markov parameter polynomial of (A, B).

Finally, for every multiple (A', B') of (A, B), there exists $M \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ with full normal rank such that $A'(\boldsymbol{\rho}) = M(\boldsymbol{\rho})A(\boldsymbol{\rho})$ and $B'(\boldsymbol{\rho}) = M(\boldsymbol{\rho})B(\boldsymbol{\rho})$. Hence $A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B(\boldsymbol{\rho})$ implies $M(\boldsymbol{\rho})A(\boldsymbol{\rho}) = A'(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B'(\boldsymbol{\rho}) = M(\boldsymbol{\rho})B(\boldsymbol{\rho})$, and it follows that $H(\boldsymbol{\rho})$ is a Markov parameter polynomial of (A', B'). Furthermore, if $H''' \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ is also a Markov parameter polynomial of (A', B'), then $A'(\boldsymbol{\rho})H'''(\boldsymbol{\rho}) = B'(\boldsymbol{\rho})$, and hence $A'(\boldsymbol{\rho})H(\boldsymbol{\rho}) = A'(\boldsymbol{\rho})H'''(\boldsymbol{\rho})$. Thus from Fact 2.7, $H'''(\boldsymbol{\rho}) = H(\boldsymbol{\rho})$, and it follows that $H(\boldsymbol{\rho})$ is the unique Markov parameter polynomial of (A', B'). □

Theorem 4.2. *Let $C \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ have full normal rank, let $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, and let $H \in \mathbb{R}_\infty^{p\times m}[\boldsymbol{\rho}]$ be the Markov parameter polynomial of (A, B) and (C, D), that is,*

$$A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B(\boldsymbol{\rho}),$$

$$C(\boldsymbol{\rho})H(\boldsymbol{\rho}) = D(\boldsymbol{\rho}).$$

Then (C, D) is a multiple of (A, B).

Proof. Let $C_R \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ be a greatest right factor of (A, C). Then there exist $A_L \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ and $C_L \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$, such that

$$A(\boldsymbol{\rho}) = A_L(\boldsymbol{\rho})C_R(\boldsymbol{\rho}),$$

$$C(\boldsymbol{\rho}) = C_L(\boldsymbol{\rho})C_R(\boldsymbol{\rho}).$$

Furthermore, since $A(\boldsymbol{\rho})$ and $C(\boldsymbol{\rho})$ have full normal rank, from Fact 2.6, we have that $A_L(\boldsymbol{\rho})$, $C_L(\boldsymbol{\rho})$, and $C_R(\boldsymbol{\rho})$ have full normal rank.

Next, let $E(\boldsymbol{\rho}) \triangleq \operatorname{madj}[A_L(\boldsymbol{\rho})]$ and $\beta(\boldsymbol{\rho}) \triangleq \operatorname{mdet}[A_L(\boldsymbol{\rho})]$. Since $A(\boldsymbol{\rho}) = A_L(\boldsymbol{\rho})C_R(\boldsymbol{\rho})$, then $E(\boldsymbol{\rho})A(\boldsymbol{\rho}) = \beta(\boldsymbol{\rho})C_R(\boldsymbol{\rho})$, and hence

$$C_L(\boldsymbol{\rho})E(\boldsymbol{\rho})A(\boldsymbol{\rho}) = C_L(\boldsymbol{\rho})\beta(\boldsymbol{\rho})C_R(\boldsymbol{\rho}) = \beta(\boldsymbol{\rho})C(\boldsymbol{\rho}).$$

Third, since $C(\boldsymbol{\rho})H(\boldsymbol{\rho}) = D(\boldsymbol{\rho})$, then

$$\beta(\boldsymbol{\rho})C(\boldsymbol{\rho})H(\boldsymbol{\rho}) = \beta(\boldsymbol{\rho})D(\boldsymbol{\rho}) = C_L(\boldsymbol{\rho})E(\boldsymbol{\rho})A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = C_L(\boldsymbol{\rho})E(\boldsymbol{\rho})B(\boldsymbol{\rho}),$$

and hence $(C_LEA, C_LEB) = (\beta C, \beta D)$.

Next, since (A, B) is left coprime and $C_L(\boldsymbol{\rho})E(\boldsymbol{\rho})$ has full normal rank, $C_L(\boldsymbol{\rho})E(\boldsymbol{\rho})$ is a greatest left factor of (C_LEA, C_LEB). Furthermore, since $\beta(\boldsymbol{\rho})I_p$ is also a left factor of (C_LEA, C_LEB), it follows that there exists $F \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ such that $C_L(\boldsymbol{\rho})E(\boldsymbol{\rho}) = \beta(\boldsymbol{\rho})F(\boldsymbol{\rho})$ and hence $(\beta FA, \beta FB) = (\beta C, \beta D)$.

Finally, since $A_L(\boldsymbol{\rho})$ has full normal rank, βI_p has full normal rank. Thus from Fact 2.7, $F(\boldsymbol{\rho})A(\boldsymbol{\rho}) = C(\boldsymbol{\rho})$ and $F(\boldsymbol{\rho})B(\boldsymbol{\rho}) = D(\boldsymbol{\rho})$. Furthermore, since $C_L(\boldsymbol{\rho})E(\boldsymbol{\rho})$ has full normal rank and $C_L(\boldsymbol{\rho})E(\boldsymbol{\rho}) = \beta(\boldsymbol{\rho})F(\boldsymbol{\rho})$, from Fact 2.6 it follows that $F(\boldsymbol{\rho})$ has full normal rank. □

5. Numerical manipulation of polynomial matrices

In this section, we introduce notation and definitions that we use to numerically manipulate polynomial matrices.

Definition 5.1. Let $F \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ have degree n and be given by

$$F(\boldsymbol{\rho}) \triangleq F_0 + F_1\boldsymbol{\rho} + \cdots + F_n\boldsymbol{\rho}^n.$$

Then for $s \geqslant 0$ and $t \geqslant 0$,

$$\theta\,(F) \triangleq \begin{bmatrix} F_0 & \cdots & F_n \end{bmatrix} \in \mathbb{R}^{p\times m(n+1)},$$

$$\theta_s\,(F) \triangleq \begin{bmatrix} F_0 & \cdots & F_s \end{bmatrix} \in \mathbb{R}^{p\times m(s+1)},$$

$$\mathcal{T}_s\,(F) \triangleq \begin{bmatrix} F_0 & F_1 & \cdots & F_s \\ 0_{p\times m} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & F_1 \\ 0_{p\times m} & \cdots & 0_{p\times m} & F_0 \end{bmatrix} \in \mathbb{R}^{p(s+1)\times m(s+1)},$$

$$\mathcal{T}_{s,t}\,(F) \triangleq \begin{bmatrix} F_0 & F_1 & \cdots & F_s & F_{s+1} & \cdots & F_{s+t} \\ 0_{p\times m} & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & F_1 & F_2 & & \vdots \\ 0_{p\times m} & \cdots & 0_{p\times m} & F_0 & F_1 & \cdots & F_t \end{bmatrix} \in \mathbb{R}^{p(s+1)\times m(s+t+1)},$$

$$\mathcal{K}_{s,t}(F) \triangleq \begin{bmatrix} \mathcal{T}_{s,t}(F) \\ \begin{bmatrix} I_{m(s+1)} & 0_{m(s+1)\times mt} \end{bmatrix} \end{bmatrix} \in \mathbb{R}^{(p+m)(s+1)\times m(s+t+1)},$$

$$\overline{\mathcal{K}}_{s,t}(F) \triangleq \begin{bmatrix} 0_{(m[s+1]+ps)\times p} & I_{(m[s+1]+ps)} \end{bmatrix} \mathcal{K}_{s,t}(F) \in \mathbb{R}^{(m[s+1]+ps)\times m(s+t+1)},$$

where $F_i = 0_{p\times m}$ for all $i > n$, and we drop the argument $\boldsymbol{\rho}$ for conciseness.

Remark 5.2. Note that $\theta\,(F) = \theta_n\,(F)$ and $\overline{\mathcal{K}}_{s,t}(F)$ is obtained by removing the first p rows of $\mathcal{K}_{s,t}(F)$.

Fact 5.3. *Let $E \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ be of degree n and let $C \in \mathbb{R}^{m\times l}[\boldsymbol{\rho}]$ be of degree s. Then the first $n+1$ matrix coefficients of the product $D(\boldsymbol{\rho}) \triangleq E(\boldsymbol{\rho})C(\boldsymbol{\rho})$ are given by*

$$\theta_n\,(D) = \theta\,(E)\,\mathcal{T}_n\,(C)\,,$$

and all of the $n+s+1$ matrix coefficients of $D(\boldsymbol{\rho})$ are given by

$$\theta\,(D) = \theta\,(E)\,\mathcal{T}_{n,s}\,(C)\,.$$

Proof. For all $i = 0, \ldots, n+s$, we have that

$$D_i = \sum_{j=\max(0,i-s)}^{\min(n,i)} E_j C_{i-j},$$

from which Fact 5.3 follows. □

Fact 5.4. *Let $E \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ be of degree n, $F \in \mathbb{R}_\infty^{m\times l}[\boldsymbol{\rho}]$, and $t \geqslant 0$. Then the first $n+t+1$ matrix coefficients of the product $G(\boldsymbol{\rho}) \triangleq E(\boldsymbol{\rho})F(\boldsymbol{\rho})$ are given by*

$$\theta_{n+t}\,(G) = \theta\,(E)\,\mathcal{T}_{n,t}\,(F)\,.$$

Proof. For all $i = 0, \ldots, n + t$, we have that

$$G_i = \sum_{j=0}^{\min(n,i)} E_j F_{i-j},$$

from which Fact 5.4 follows. □

Remark 5.5. Let $D \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ be given by

$$D(\boldsymbol{\rho}) \triangleq \begin{bmatrix} d_1(\boldsymbol{\rho}) & \cdots & d_m(\boldsymbol{\rho}) \end{bmatrix},$$

where $d_1, \ldots, d_m \in \mathbb{R}^{p\times 1}[\boldsymbol{\rho}]$. Then

$$\mathrm{vec}\left[D(\boldsymbol{\rho})\right] \triangleq \begin{bmatrix} d_1(\boldsymbol{\rho}) \\ \vdots \\ d_m(\boldsymbol{\rho}) \end{bmatrix} \in \mathbb{R}^{pm\times 1}[\boldsymbol{\rho}],$$

$$I_s \otimes D(\boldsymbol{\rho}) \triangleq \begin{bmatrix} D(\boldsymbol{\rho}) & & \\ & \ddots & \\ & & D(\boldsymbol{\rho}) \end{bmatrix} \in \mathbb{R}^{ps\times ms}[\boldsymbol{\rho}].$$

Fact 5.6. *Let $C \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ and $D \in \mathbb{R}^{m\times \ell}[\boldsymbol{\rho}]$. Then*

$$\mathrm{vec}\left[C(\boldsymbol{\rho})D(\boldsymbol{\rho})\right] = \left[I_\ell \otimes C(\boldsymbol{\rho})\right]\mathrm{vec}\left[D(\boldsymbol{\rho})\right].$$

Proof. See [2]. □

6. Numerical algorithms for computing the Markov parameters

Here we demonstrate how to compute the Markov parameters of (A, B) from a multiple of (A, B) numerically. Since Theorem 4.1 is constructive given a comonic multiple of (A, B), first we present two methods of computing a comonic multiple of (A, B) numerically.

Proposition 6.1. *Let (C, D) be a multiple of (A, B). Then there exists a nonnegative t such that*

$$\mathrm{rank}\begin{bmatrix} \begin{bmatrix} 0_{p\times pt} & I_p \end{bmatrix} \\ \mathcal{T}_t\,(C) \end{bmatrix} = \mathrm{rank}\left[\mathcal{T}_t\,(C)\right]. \tag{7}$$

Furthermore, let $U \in \mathbb{R}^{p\times p(t+1)}$ be a solution of

$$\begin{bmatrix} 0_{p\times pt} & I_p \end{bmatrix} = U\mathcal{T}_t\,(C)\,, \tag{8}$$

and let $L \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ be the polynomial matrix of degree t such that

$$\theta\,(L) \triangleq U.$$

Then $(E, F) = (LC, LD)$ is a comonic multiple of (A, B) and (C, D).

Proof. From Proposition 2.22, there exists a comonic multiple $(E', F') = (L'C, L'D)$ of (A, B) and (C, D). Hence, letting E_γ be the trailing coefficient of $E'(\boldsymbol{\rho})$, we have that

$$\theta_\gamma \left(E'\right) = \left[\, 0_{p\times p\gamma} \;\; I_p \,\right],$$

where, from Fact 5.3, it follows that

$$\theta_\gamma \left(E'\right) = \left[\, 0_{p\times p\gamma} \;\; I_p \,\right] = \theta_\gamma \left(L'\right) \mathcal{T}_\gamma \left(C\right).$$

Thus there exists a nonnegative t such that (7) holds.

Finally, since there exists a nonnegative t such that (7) holds, there exists a $U \in \mathbb{R}^{p\times p(t+1)}$ such that (8) holds ($\theta_t(L')$ being one such U). Hence,

$$\theta \left(L\right) \mathcal{T}_t \left(C\right) = \theta_t(E) = \left[\, 0_{p\times pt} \;\; I_p \,\right],$$

that is, $E(\boldsymbol{\rho})$ is comonic. Therefore (E, F) is a comonic multiple of (A, B) and (C, D). □

Algorithm 6.2. Let (C, D) be a given multiple of (A, B) of degree s. The following algorithm yields a comonic multiple $(E, F) = (LC, LD)$ of (A, B), as described in Proposition 6.1.

(1) $t = -1$.
(2) $t = t + 1$.
(3) $u = \operatorname{rank}\left[\mathcal{T}_t \left(C\right)\right]$.
(4) $v = \operatorname{rank}\begin{bmatrix} \left[\, 0_{p\times pt} \;\; I_p \,\right] \\ \mathcal{T}_t \left(C\right) \end{bmatrix}$.
(5) If $u < v$, go to Step 2. Otherwise, continue.
(6) $\theta \left(L\right) = \left[\, 0_{p\times pt} \;\; I_p \,\right] \mathcal{T}_t^{+} \left(C\right)$, where $(\cdot)^{+}$ denotes the Moore–Penrose generalized inverse.
(7) $\theta \left(E\right) = \theta \left(L\right) \mathcal{T}_{t,s} \left(C\right)$.
(8) $\theta \left(F\right) = \theta \left(L\right) \mathcal{T}_{t,s} \left(D\right)$.

Next, we present an alternative method for computing a comonic multiple of (A, B). Specifically, we show how to compute a comonic quasi-scalar multiple of (A, B) from an arbitrary multiple (C, D) of (A, B).

Proposition 6.3. *Let (C, D) be a multiple of (A, B) of degree s and let*

$$L(\boldsymbol{\rho}) \triangleq I_m \otimes C^T(\boldsymbol{\rho}),$$

$$M(\boldsymbol{\rho}) \triangleq \operatorname{vec}\left[D(\boldsymbol{\rho})\right]^T.$$

Then there exists a nonnegative t such that

$$\operatorname{nullity}\left(W_t^T\right) \geqslant 1, \tag{9}$$

where

$$W_t \triangleq \begin{bmatrix} \mathcal{T}_{t,s} \left(M\right) \\ \mathcal{T}_{t,s} \left(L\right) \end{bmatrix}.$$

Furthermore, let $U \in \mathbb{R}^{(pm+1)(t+1)}$ be a nonzero vector in the nullspace of W_t^T, and let $\gamma' \in \mathbb{R}[\boldsymbol{\rho}]$ and $F' \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ be the polynomial matrices of degree t such that

$$\theta\left(\gamma'\right) \triangleq U^T \begin{bmatrix} I_{t+1} \\ 0_{pm(t+1)\times t+1} \end{bmatrix},$$

$$\theta\left(\mathrm{vec}\left[F'\right]^T\right) \triangleq -U^T \begin{bmatrix} 0_{t+1\times pm(t+1)} \\ I_{pm(t+1)} \end{bmatrix}.$$

Then $\gamma'(\boldsymbol{\rho})$ is nonzero. Finally, let γ_i be the trailing coefficient of $\gamma'(\boldsymbol{\rho})$, and let $\gamma(\boldsymbol{\rho}) \triangleq \gamma'(\boldsymbol{\rho})/\gamma_i$ and $F(\boldsymbol{\rho}) \triangleq F'(\boldsymbol{\rho})/\gamma_i$. Then $(\gamma I_p, F)$ is a comonic quasi-scalar multiple of (A, B).

Proof. First, letting $E(\boldsymbol{\rho}) \triangleq \mathrm{madj}\,[C(\boldsymbol{\rho})]$ and $\beta(\boldsymbol{\rho}) \triangleq \mathrm{mdet}\,[C(\boldsymbol{\rho})]$, it follows that

$$C(\boldsymbol{\rho})E(\boldsymbol{\rho})D(\boldsymbol{\rho}) = \beta(\boldsymbol{\rho})D(\boldsymbol{\rho}).$$

Hence, from Fact 5.6, we have that

$$[I_m \otimes C(\boldsymbol{\rho})]\mathrm{vec}\left[E(\boldsymbol{\rho})D(\boldsymbol{\rho})\right] = \beta(\boldsymbol{\rho})\mathrm{vec}\left[D(\boldsymbol{\rho})\right],$$

where $\beta(\boldsymbol{\rho})$ is nonzero since $C(\boldsymbol{\rho})$ has full normal rank. Thus, letting η denote the degree of $(\beta I_p, ED)$, from Fact 5.3 we have that

$$\theta\left(\mathrm{vec}\left[ED\right]^T\right)\mathcal{T}_{\eta,s}(L) = \theta(\beta)\mathcal{T}_{\eta,s}(M).$$

Thus there exists a nonnegative t such that (9) holds.

Next, since there exists a nonnegative t such that (9) holds, there exists a nonzero $U \in \mathbb{R}^{(pm+1)(t+1)}$ in the nullspace of W_t^T. Furthermore, from the definition of $\gamma'(\boldsymbol{\rho})$ and $F'(\boldsymbol{\rho})$, it follows that

$$\theta\left(\mathrm{vec}\left[F'\right]^T\right)\mathcal{T}_{t,s}(L) = \theta(\gamma')\mathcal{T}_{t,s}(M),$$

and hence from Fact 5.3, we have that $C(\boldsymbol{\rho})F'(\boldsymbol{\rho}) = \gamma'(\boldsymbol{\rho})D(\boldsymbol{\rho})$.

Next, suppose that $\gamma'(\boldsymbol{\rho})$ is zero. Then $C(\boldsymbol{\rho})F'(\boldsymbol{\rho}) = 0_{p\times m}$ and therefore, since $C(\boldsymbol{\rho})$ has full normal rank, from Fact 2.7, $F'(\boldsymbol{\rho}) = 0_{p\times m}$. However this contradicts the fact that U is nonzero. Hence $\gamma'(\boldsymbol{\rho})$ is nonzero.

Finally, letting $H \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$ denote the Markov parameter polynomial of (A, B) and (C, D), it follows that

$$C(\boldsymbol{\rho})H(\boldsymbol{\rho}) = D(\boldsymbol{\rho}),$$

$$C(\boldsymbol{\rho})\gamma(\boldsymbol{\rho})H(\boldsymbol{\rho}) = \gamma(\boldsymbol{\rho})D(\boldsymbol{\rho}) = C(\boldsymbol{\rho})F(\boldsymbol{\rho}).$$

Therefore from Fact 2.7, $\gamma(\boldsymbol{\rho})H(\boldsymbol{\rho}) = F(\boldsymbol{\rho})$. Furthermore, since $\gamma(\boldsymbol{\rho})I_p$ is comonic and quasi-scalar, from Fact 2.9, $\gamma(\boldsymbol{\rho})$ has full normal rank. Therefore, from Theorem 4.2, $(\gamma I_p, F)$ is a comonic quasi-scalar multiple of (A, B). □

Algorithm 6.4. Let (C, D) be a given multiple of (A, B) of degree s. The following algorithm yields a comonic quasi-scalar multiple $(\gamma I_p, F)$ of (A, B), as described in Proposition 6.3.

(1) $t = -1$.
(2) $L(\boldsymbol{\rho}) = I_m \otimes C^T(\boldsymbol{\rho})$.
(3) $M(\boldsymbol{\rho}) = \mathrm{vec}\left[D(\boldsymbol{\rho})\right]^T$.

(4) $t = t + 1$.
(5) Compute the singular value decomposition of

$$W_t = \begin{bmatrix} \mathcal{T}_{t,s}(M) \\ \mathcal{T}_{t,s}(L) \end{bmatrix}.$$

(6) If nullity $\left(W_t^T\right) = 0$, go to Step 4. Otherwise, continue.
(7) Choose a nonzero vector $U \in \mathbb{R}^{(pm+1)(t+1)}$ in the nullspace of W_t^T, and scale U such that the first nonzero component is 1.
(8) $\theta(\gamma) = U^T \begin{bmatrix} I_{t+1} \\ 0_{pm(t+1)\times t+1} \end{bmatrix}$.
(9) $\theta\left(\mathrm{vec}\ [F]^T\right) = -U^T \begin{bmatrix} 0_{t+1\times pm(t+1)} \\ I_{pm(t+1)} \end{bmatrix}$.

Remark 6.5. Proposition 6.1 and 6.3 provide two alternative ways of obtaining a comonic multiple of (A, B) numerically, with the main difference being that Proposition 6.1 provides a comonic multiple of both (A, B) and (C, D), while Proposition 6.3 provides a quasi-scalar comonic multiple that is only guaranteed to be a multiple of (A, B). Typically, Proposition 6.1 will provide a comonic multiple of lower degree than Proposition 6.3, due to the quasi-scalar requirement in Proposition 6.3, however this is not always the case. One of the benefits of Proposition 6.3 is that quasi-scalar multiples exhibit a direct link to transfer function, and thus state-space, models, as shown in Section 8, albeit at the expense of increased computational complexity.

Now that we have shown how to compute a comonic multiple of (A, B) numerically, Theorem 4.1 can be used to compute the Markov parameters of (A, B) algebraically. Specifically, we have the following Proposition:

Proposition 6.6. *Let (C, D) be a multiple of (A, B) and let (E, F) be a comonic multiple of (A, B) computed using either Proposition 6.1 (Algorithm 6.2) or Proposition 6.3 (Algorithm 6.4). Then the Markov parameters of (A, B) are given by (5).*

7. Numerical algorithms for computing a multiple of (A, B)

Here we present two methods of computing a multiple of (A, B) numerically from the Markov parameters of (A, B).

Proposition 7.1. *Let $H \in \mathbb{R}_\infty^{p\times m}[\boldsymbol{\rho}]$ be the Markov parameter polynomial of (A, B) and let $\bar{n} \geqslant n^\star$. Then for all nonnegative t,*

$$\mathrm{rank}\left[\mathcal{K}_{t,n^\star}(H)\right] = \mathrm{rank}\left[\mathcal{K}_{t,\bar{n}}(H)\right], \tag{10}$$
$$\mathrm{rank}\left[\overline{\mathcal{K}}_{t,n^\star}(H)\right] = \mathrm{rank}\left[\overline{\mathcal{K}}_{t,\bar{n}}(H)\right]. \tag{11}$$

Furthermore, there exists a nonnegative $s \leqslant n^\star$ such that

$$\mathrm{rank}\left[\overline{\mathcal{K}}_{s,n^\star}(H)\right] = \mathrm{rank}\left[\mathcal{K}_{s,n^\star}(H)\right]. \tag{12}$$

Finally, letting (12) hold, letting $E \in \mathbb{R}^{p\times p}[\boldsymbol{\rho}]$ have full normal rank, letting $F \in \mathbb{R}^{p\times m}[\boldsymbol{\rho}]$, and

$$\begin{bmatrix} \theta(E) & -\theta(F) \end{bmatrix} \mathcal{K}_{s,\bar{n}}(H) = 0_{p\times m(s+\bar{n}+1)}, \tag{13}$$

then (E, F) is a multiple of (A, B).

Proof. First, since $(A^\bigstar, B^\bigstar)$ is a quasi-scalar multiple of (A, B), then

$$A^\bigstar(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B^\bigstar(\boldsymbol{\rho}).$$

Furthermore, from (5), for all $j \geqslant 1$ we have that

$$a^\bigstar_{n^\bigstar}H_j + \cdots + a^\bigstar_1 H_{n^\bigstar+j-1} + H_{n^\bigstar+j} = 0_{p\times m},$$

where, since $a^\bigstar_1, \ldots, a^\bigstar_{n^\bigstar} \in \mathbb{R}$,

$$H_j a^\bigstar_{n^\bigstar} + \cdots + H_{n^\bigstar+j-1}a^\bigstar_1 + H_{n^\bigstar+j} = 0_{p\times m}. \tag{14}$$

Next, suppose that $\bar{n} = n^\bigstar + 1$. Then from (14), the columns of $\mathcal{K}_{t,\bar{n}}(H)$ beginning with $H_{t+n^\bigstar+1}$ are in the column space of the previous $mn^\bigstar$ columns, specifically,

$$\begin{bmatrix} H_{t+1} \\ \vdots \\ H_1 \\ 0_{m(t+1)\times m} \end{bmatrix} a^\bigstar_n + \cdots + \begin{bmatrix} H_{t+n^\bigstar} \\ \vdots \\ H_{n^\bigstar} \\ 0_{m(t+1)\times m} \end{bmatrix} a^\bigstar_1 + \begin{bmatrix} H_{t+n^\bigstar+1} \\ \vdots \\ H_{n^\bigstar+1} \\ 0_{m(t+1)\times m} \end{bmatrix} = 0_{(p+m)(t+1)\times m}.$$

Similarly, for all $j \geqslant 1$, the columns of $\mathcal{K}_{t,n^\bigstar+j}(H)$ beginning with $H_{t+n^\bigstar+j}$ are in the column space of the previous $mn^\bigstar$ columns. Hence, by induction, we have (10). Furthermore, (11) follows directly from (10) since $\overline{\mathcal{K}}_{t,\bar{n}}(H)$ is obtained by removing the first p rows of $\mathcal{K}_{t,\bar{n}}(H)$.

Next, since $(A^\bigstar, B^\bigstar)$ is a comonic quasi-scalar multiple of (A, B), from Fact 5.4, we have that

$$\theta(A^\bigstar)\mathcal{T}_{n^\bigstar,n^\bigstar}(H) = \theta_{2n^\bigstar}(B^\bigstar) = \begin{bmatrix} \theta(B^\bigstar) & 0_{p\times mn^\bigstar} \end{bmatrix},$$

and hence

$$\begin{bmatrix} \theta(A^\bigstar) & -\theta(B^\bigstar) \end{bmatrix} \mathcal{K}_{n^\bigstar,n^\bigstar}(H) = 0_{p\times m(2n^\bigstar+1)},$$

where, since $A^\bigstar(\boldsymbol{\rho})$ is comonic with $A^\bigstar_0 = I_p$, we have (12).

Finally, let $E(\boldsymbol{\rho})$ have full normal rank and let (13) hold. Then from (10), for all $j \geqslant 1$, we have that

$$\begin{bmatrix} \theta\,(E) & -\theta\,(F) \end{bmatrix} \mathcal{K}_{s,\bar{n}+j}(H) = 0_{p\times m(s+\bar{n}+j+1)},$$

and hence

$$\theta(E)\mathcal{T}_{s,\bar{n}+j}(H) = \begin{bmatrix} \theta(F) & 0_{p\times m(\bar{n}+j)} \end{bmatrix}.$$

Therefore, from Fact 5.4, $E(\boldsymbol{\rho})H(\boldsymbol{\rho}) = F(\boldsymbol{\rho})$, and from Theorem 4.2, (E, F) is a multiple of (A, B). □

Algorithm 7.2. Let $\bar{n}$ be a known upper bound for $n^\bigstar$, that is, $\bar{n} \geqslant n^\bigstar$. Also, let $H(\boldsymbol{\rho})$ be the Markov parameter polynomial of (A, B), and let $H_0, \ldots, H_{2\bar{n}+1}$ be given. Then following algorithm yields a comonic multiple (E, F) of (A, B), as described in Proposition 7.1.

(1) $s = 0$.
(2) $s = s + 1$.
(3) $u = \text{rank}\,[\overline{\mathcal{K}}_{s,\bar{n}}(H)]$.
(4) $v = \text{rank}\,[\mathcal{K}_{s,\bar{n}}(H)]$.
(5) If $u < v$, go to Step 2. Otherwise, continue.
(6) $W = \theta_{s+\bar{n}}(H)\overline{\mathcal{K}}^{+}_{s,\bar{n}}(H)$, where $(\cdot)^+$ denotes the Moore–Penrose generalized inverse.

(7) $\theta\,(E) = \begin{bmatrix} I_p & -W \begin{bmatrix} I_{ps} \\ 0_{m(s+1)\times ps} \end{bmatrix} \end{bmatrix}.$

(8) $\theta\,(F) = W \begin{bmatrix} 0_{ps\times m(s+1)} \\ I_{m(s+1)} \end{bmatrix}.$

Next, we present an alternative method for computing a comonic multiple of (A, B). Specifically, we show how to compute a comonic quasi-scalar multiple of (A, B) from the Markov parameters of (A, B).

Proposition 7.3. *Let $H \in \mathbb{R}_\infty^{p\times m}[\boldsymbol{\rho}]$ be the Markov parameter polynomial of (A, B), $\bar{n} \geqslant n^\star$, and $H^\star(\boldsymbol{\rho}) \triangleq \operatorname{vec}\left[H(\boldsymbol{\rho})\right]^T$. Then for all nonnegative t,*

$$\operatorname{rank}\left[\mathcal{K}_{t,n^\star}(H^\star)\right] = \operatorname{rank}\left[\mathcal{K}_{t,\bar{n}}(H^\star)\right]. \tag{15}$$

Furthermore, there exists a nonnegative $s \leqslant n^\star$ such that

$$\operatorname{nullity}\left[\mathcal{K}^T_{s,n^\star}(H^\star)\right] \geqslant 1. \tag{16}$$

Finally, letting $\gamma \in \mathbb{R}[\boldsymbol{\rho}]$ be nonzero, $D \in \mathbb{R}^{1\times pm}[\boldsymbol{\rho}]$, and

$$\begin{bmatrix} \theta(\gamma) & -\theta(D) \end{bmatrix} \mathcal{K}_{s,\bar{n}}(H^\star) = 0_{1\times pm(s+\bar{n}+1)}, \tag{17}$$

$$F(\boldsymbol{\rho}) \triangleq \operatorname{unvec}\left[D^T(\boldsymbol{\rho})\right],$$

then $(\gamma I_p, F)$ is a quasi-scalar multiple of (A, B).

Proof. First, since $(A^\star, B^\star)$ is a quasi-scalar multiple of (A, B), then

$$A^\star(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B^\star(\boldsymbol{\rho}).$$

Hence letting $\overline{B}^\star(\boldsymbol{\rho}) \triangleq \operatorname{vec}\,[B(\boldsymbol{\rho})]^T$, from Fact 5.6, we have that

$$a^\star(\boldsymbol{\rho})H^\star(\boldsymbol{\rho}) = \overline{B}^\star(\boldsymbol{\rho}).$$

Therefore, from (5), for all $j \geqslant 1$ we have that

$$a^\star_{n^\star}H^\star_j + \cdots + a^\star_1 H^\star_{n^\star+j-1} + H^\star_{n^\star+j} = 0_{1\times pm},$$

where, since $a^\star_1, \ldots, a^\star_{n^\star} \in \mathbb{R}$,

$$H^\star_j a^\star_{n^\star} + \cdots + H^\star_{n^\star+j-1}a^\star_1 + H^\star_{n^\star+j} = 0_{1\times pm}.$$

Thus (15) follows directly from the proof of Proposition 7.1.

Next, since $(A^\star, B^\star)$ is a comonic quasi-scalar multiple of (A, B), from Fact 5.4, we have that

$$\theta(a^\star)\mathcal{T}_{n^\star,n^\star}(H^\star) = \theta_{2n^\star}\left(\overline{B}^\star\right) = \begin{bmatrix} \theta\left(\overline{B}^\star\right) & 0_{1\times pmn^\star} \end{bmatrix},$$

and hence

$$\begin{bmatrix} \theta(a^\star) & -\theta\left(\overline{B}^\star\right) \end{bmatrix} \mathcal{K}_{n^\star,n^\star}(H^\star) = 0_{1\times pm(2n^\star+1)},$$

where, since $a^\star(\boldsymbol{\rho})$ is comonic with $a^\star_0 = 1$, we have (16).

Finally, let $\gamma(\boldsymbol{\rho})$ be nonzero and let (17) hold. Then from (15), for all $j \geqslant 1$, we have that

$$\left[\, \theta(\gamma) \;\; -\theta(D) \,\right] \mathcal{K}_{s,\bar{n}+j}(H^{\star}) = 0_{1\times pm(s+\bar{n}+j+1)},$$

and hence

$$\theta(\gamma)\mathcal{T}_{s,\bar{n}+j}(H^{\star}) = \left[\, \theta(D) \;\; 0_{1\times pm(\bar{n}+j)} \,\right].$$

Therefore, from Fact 5.4, $\gamma(\boldsymbol{\rho})H^{\star}(\boldsymbol{\rho}) = D(\boldsymbol{\rho})$, and from Fact 5.6,

$$\gamma(\boldsymbol{\rho})H(\boldsymbol{\rho}) = F(\boldsymbol{\rho}).$$

Furthermore, since $\gamma(\boldsymbol{\rho})$ is nonzero and quasi-scalar, from Fact 2.9, $\gamma(\boldsymbol{\rho})I_p$ has full normal rank. Hence, from Theorem 4.2, $(\gamma I_p, F)$ is a multiple of (A, B). □

Algorithm 7.4. Let $\bar{n}$ be a known upper bound for $n^{\star}$, that is, $\bar{n} \geqslant n^{\star}$. Also, let $H(\boldsymbol{\rho})$ be the Markov parameter polynomial of (A, B), and let $H_0, \ldots, H_{2\bar{n}+1}$ be given. Finally, for all $i = 0, \ldots, 2\bar{n}+1$, let $H_i^{\star} \triangleq \operatorname{vec}\left[H_i\right]^T$. Then the following algorithm yields a quasi-scalar comonic multiple $(\gamma I_p, F)$ of (A, B), as described in Proposition 7.3.

(1) $s = 0$.
(2) $s = s + 1$.
(3) Compute the singular value decomposition of $\mathcal{K}_{s,\bar{n}}(H^{\star})$.
(4) If nullity $\left(\mathcal{K}_{s,\bar{n}}^T(H^{\star})\right) = 0$, go to Step 2. Otherwise, continue.
(5) Choose a nonzero vector $U \in \mathbb{R}^{1\times(pm+1)(s+1)}$ in the left nullspace of $\mathcal{K}_{s,\bar{n}}(H^{\star})$.
(6) $\theta(\gamma) = U \begin{bmatrix} I_{s+1} \\ 0_{pm(s+1)\times(s+1)} \end{bmatrix}$.
(7) $\theta(D) = -U \begin{bmatrix} 0_{(s+1)\times pm(s+1)} \\ I_{pm(s+1)} \end{bmatrix}$.
(8) $F(\boldsymbol{\rho}) = \operatorname{unvec}\left[D^T(\boldsymbol{\rho})\right]$.

Remark 7.5. As in the previous section, Proposition 7.1 and 7.3 provide two alternative ways of obtaining a comonic multiple of (A, B) numerically from the Markov parameters, with the main difference being that Proposition 7.1 provides a comonic multiple, while Proposition 7.3 provides a quasi-scalar comonic multiple. Proposition 7.1 will always provide a comonic multiple of degree less than or equal to Proposition 7.3, due to the quasi-scalar requirement in Proposition 7.3. However, one of the benefits of Proposition 7.3 is that quasi-scalar multiples exhibit a direct link to transfer function, and thus state-space, models, as we demonstrate in the following section, albeit at the expense of increased computational complexity.

Remark 7.6. In both Algorithm 7.2 and Algorithm 7.4, it is required that an upper bound $\bar{n}$ for $n^{\star}$ is known. However, in practice, this may be difficult or impossible to ascertain. In this case, we would advise the reader to take an initial guess of for the upper bound, say n_1, and run the algorithms as proposed. If in Algorithms 7.2 and 7.4, the rank conditions are not satisfied for $s \leqslant n_1$, then increase n_1, provide more Markov parameters, and run the algorithms again.

8. Connection with state-space models

Here we consider the connection between polynomial matrix models, state-space models, and Markov parameters. Specifically, we review the well-known method of obtaining a polynomial matrix model from a state-space model, and then show that, using the Markov parameters of the state-space model, we can obtain the same polynomial matrix model using the algorithms in the present paper, particularly Proposition 7.1. Furthermore, we show that all of the same rank properties presented in Proposition 7.1 still hold when the Markov parameters are generated from a state-space model, where $n^\star$ is replaced by the order of the state-space model which generates the Markov parameters.

Proposition 8.1. *Consider the state-space system*

$$x(t) = \boldsymbol{\rho}\tilde{A}x(t) + \boldsymbol{\rho}\tilde{B}u(t),$$

$$y(t) = \tilde{C}x(t) + \tilde{D}u(t),$$

where $\tilde{A} \in \mathbb{R}^{n\times n}$, $\tilde{B} \in \mathbb{R}^{n\times m}$, $\tilde{C} \in \mathbb{R}^{p\times n}$, *and* $\tilde{D} \in \mathbb{R}^{p\times m}$, $x \in \mathbb{R}^n$ *is the state,* $u \in \mathbb{R}^m$ *is the input, and* $y \in \mathbb{R}^p$ *is the output. Also, let*

$$A(\boldsymbol{\rho}) \triangleq \det\left[I_n - \boldsymbol{\rho}\tilde{A}\right],$$

$$E(\boldsymbol{\rho}) \triangleq \operatorname{adj}\left[I_n - \boldsymbol{\rho}\tilde{A}\right],$$

$$B(\boldsymbol{\rho}) \triangleq \boldsymbol{\rho}\tilde{C}E(\boldsymbol{\rho})\tilde{B} + A(\boldsymbol{\rho})\tilde{D}.$$

Then $A(\boldsymbol{\rho})y(t) = B(\boldsymbol{\rho})u(t)$.

Proof

$$A(\boldsymbol{\rho})y(t) = \tilde{C}A(\boldsymbol{\rho})x(t) + A(\boldsymbol{\rho})\tilde{D}u(t) = \tilde{C}\left[E(\boldsymbol{\rho})\boldsymbol{\rho}\tilde{B}u(t)\right] + A(\boldsymbol{\rho})\tilde{D}u(t) = B(\boldsymbol{\rho})u(t). \quad \square$$

Definition 8.2. Let $\tilde{A} \in \mathbb{R}^{n\times n}$, $\tilde{B} \in \mathbb{R}^{n\times m}$, $\tilde{C} \in \mathbb{R}^{p\times n}$, and $\tilde{D} \in \mathbb{R}^{p\times m}$. Also, for $i \geqslant 1$, let

$$H_0 \triangleq \tilde{D}, \quad H_1 \triangleq \tilde{C}\tilde{B}, \quad H_2 = \tilde{C}\tilde{A}\tilde{B}, \quad \cdots, \quad H_i \triangleq \tilde{C}\tilde{A}^{i-1}\tilde{B}.$$

Then H_j is the jth *Markov parameter* of $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$, and

$$H(\boldsymbol{\rho}) \triangleq \sum_{j=0}^{\infty} H_j \boldsymbol{\rho}^j$$

is the *Markov parameter polynomial* of $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$.

Proposition 8.3. *Consider the controllable state-space model*

$$x(t) = \boldsymbol{\rho}\tilde{A}x(t) + \boldsymbol{\rho}\tilde{B}u(t),$$

$$y(t) = \tilde{C}x(t) + \tilde{D}u(t),$$

where $\tilde{A} \in \mathbb{R}^{n\times n}$, $\tilde{B} \in \mathbb{R}^{n\times m}$, $\tilde{C} \in \mathbb{R}^{p\times n}$, *and* $\tilde{D} \in \mathbb{R}^{p\times m}$, $x \in \mathbb{R}^n$ *is the state,* $u \in \mathbb{R}^m$ *is the input, and* $y \in \mathbb{R}^p$ *is the output. Furthermore, let* $\bar{n} \geqslant n$ *and let* $H \in \mathbb{R}_\infty^{p\times m}[\boldsymbol{\rho}]$ *be the Markov parameter polynomial of* $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$. *Then for all nonnegative* t,

$$\operatorname{rank}\left[\mathcal{K}_{t,n}(H)\right] = \operatorname{rank}\left[\mathcal{K}_{t,\bar{n}}(H)\right], \tag{18}$$

$$\operatorname{rank}\left[\overline{\mathcal{K}}_{t,n}(H)\right] = \operatorname{rank}\left[\overline{\mathcal{K}}_{t,\bar{n}}(H)\right]. \tag{19}$$

Furthermore, letting

$$A(\boldsymbol{\rho}) \triangleq \det\left[I_n - \boldsymbol{\rho}\tilde{A}\right],$$
$$E(\boldsymbol{\rho}) \triangleq \operatorname{adj}\left[I_n - \boldsymbol{\rho}\tilde{A}\right],$$
$$B(\boldsymbol{\rho}) \triangleq \boldsymbol{\rho}\tilde{C}E(\boldsymbol{\rho})\tilde{B} + A(\boldsymbol{\rho})\tilde{D},$$

then

$$\left[\ \theta_{\bar{n}}\left(AI_p\right)\ -\theta_{\bar{n}}\left(B\right)\ \right]\mathcal{K}_{n,\bar{n}}\left(H\right) = 0_{p\times m(n+\bar{n}+1)}, \tag{20}$$

and there exists a nonnegative $s \leqslant n$ such that

$$\operatorname{rank}\left[\overline{\mathcal{K}}_{s,n}(H)\right] = \operatorname{rank}\left[\mathcal{K}_{s,n}(H)\right]. \tag{21}$$

Proof. First, note that from Definition 5.1 and Definition 8.2, for all $\bar{n} \geqslant n$ and $t \geqslant 0$, we have that

$$\mathcal{K}_{t,\bar{n}}(H) = \begin{bmatrix} \mathcal{T}_t(H) & \mathcal{O}_t\left(\tilde{A},\tilde{C}\right)\mathcal{C}_{\bar{n}}\left(\tilde{A},\tilde{B}\right) \\ I_{m(t+1)} & 0_{m(t+1)\times m\bar{n}} \end{bmatrix},$$
$$\mathcal{O}_t\left(\tilde{A},\tilde{C}\right) \triangleq \left[\ \left(\tilde{C}\tilde{A}^t\right)^T\ \cdots\ \left(\tilde{C}\tilde{A}\right)^T\ \tilde{C}^T\ \right]^T,$$
$$\mathcal{C}_{\bar{n}}\left(\tilde{A},\tilde{B}\right) \triangleq \left[\ \tilde{B}\ \tilde{A}\tilde{B}\ \cdots\ \tilde{A}^{\bar{n}-1}\tilde{B}\ \right],$$

where $\mathcal{O}_n\left(\tilde{A},\tilde{C}\right)$ is the reordered observability matrix of $(\tilde{A},\tilde{C})$, and $\mathcal{C}_n\left(\tilde{A},\tilde{B}\right)$ is the controllability matrix of $(\tilde{A},\tilde{B})$. Furthermore, since $(\tilde{A},\tilde{B},\tilde{C},\tilde{D})$ is controllable, then for all $\bar{n} \geqslant n$, $\mathcal{C}_{\bar{n}}\left(\tilde{A},\tilde{B}\right)$ has full row rank. Hence for all $\bar{n} \geqslant n$, it follows that

$$\operatorname{rank}\left[\mathcal{O}_t\left(\tilde{A},\tilde{C}\right)\mathcal{C}_n\left(\tilde{A},\tilde{B}\right)\right] = \operatorname{rank}\left[\mathcal{O}_t\left(\tilde{A},\tilde{C}\right)\mathcal{C}_{\bar{n}}\left(\tilde{A},\tilde{B}\right)\right] = \operatorname{rank}\left[\mathcal{O}_t\left(\tilde{A},\tilde{C}\right)\right] \leqslant n,$$

that is, the final $m(\bar{n}-n)$ columns of $\mathcal{K}_{t,\bar{n}}(H)$ are in the column space of the previous mn columns and therefore (18). Similarly, we have (19).

Next, note that

$$B(\boldsymbol{\rho}) - A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = \boldsymbol{\rho}\tilde{C}E(\boldsymbol{\rho})\tilde{B} + A(\boldsymbol{\rho})\tilde{D} - A(\boldsymbol{\rho})\left(\tilde{D} + \sum_{i=1}^{\infty}\tilde{C}\tilde{A}^{i-1}\tilde{B}\boldsymbol{\rho}^i\right) = \boldsymbol{\rho}\tilde{C}\left(E(\boldsymbol{\rho}) - A(\boldsymbol{\rho})\sum_{i=0}^{\infty}\tilde{A}^i\boldsymbol{\rho}^i\right)\tilde{B}.$$

Furthermore, since

$$\left[I_n - \boldsymbol{\rho}\tilde{A}\right]\sum_{i=0}^{\infty}\tilde{A}^i\boldsymbol{\rho}^i = I_n,$$

it follows that

$$\left[I_n - \boldsymbol{\rho}\tilde{A}\right]\left(E(\boldsymbol{\rho}) - A(\boldsymbol{\rho})\sum_{i=0}^{\infty}\tilde{A}^i\boldsymbol{\rho}^i\right) = \det\left[I_n - \boldsymbol{\rho}\tilde{A}\right] - A(\boldsymbol{\rho}) = 0_{n\times n},$$

where, since $\left[I_n - \boldsymbol{\rho}\tilde{A}\right]$ is regular, from Fact 2.8, $\left[I_n - \boldsymbol{\rho}\tilde{A}\right]$ has full row rank. Hence, from Fact 2.7,

$$E(\boldsymbol{\rho}) - A(\boldsymbol{\rho})\sum_{i=0}^{\infty}\tilde{A}^i\boldsymbol{\rho}^i = 0_{n\times n},$$

and therefore

$$B(\boldsymbol{\rho}) - A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = \boldsymbol{\rho}\tilde{C}\left(0_{n\times n}\right)\tilde{B} = 0_{p\times m},$$

that is, $A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B(\boldsymbol{\rho})$.

Finally, note that $A(\boldsymbol{\rho})$ has degree less than or equal n from the definition of the determinant, and from the definition of the adjugate in terms of the cofactor matrix, it follows that $E(\boldsymbol{\rho})$ has degree less than or equal $n-1$. Hence $B(\boldsymbol{\rho})$ has degree less than or equal to n. Therefore, since (A, B) has degree less than or equal n, and $A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B(\boldsymbol{\rho})$, we have (20). Furthermore, since

$$A(0) = \det\left[I_n - 0 \times \tilde{A}\right] = 1 = A_0,$$

we have (21). □

9. Numerical examples

In the following, we illustrate Algorithm 6.2, Algorithm 6.4, Proposition 6.6, Algorithm 7.2, and Algorithm 7.4 with a low-degree example for conciseness. Let

$$A(\boldsymbol{\rho}) \triangleq \begin{bmatrix} (2+\boldsymbol{\rho}) & (3+\boldsymbol{\rho}) \\ (5+\boldsymbol{\rho}) & (7+\boldsymbol{\rho}) \end{bmatrix}, \tag{22}$$

$$B(\boldsymbol{\rho}) \triangleq \begin{bmatrix} (1+\boldsymbol{\rho}) & (2+\boldsymbol{\rho}) & (3+\boldsymbol{\rho}) \\ (4+\boldsymbol{\rho}) & (5+\boldsymbol{\rho}) & (6+\boldsymbol{\rho}) \end{bmatrix}, \tag{23}$$

$$N(\boldsymbol{\rho}) \triangleq \begin{bmatrix} (1+\boldsymbol{\rho}) & (2+\boldsymbol{\rho}) \\ (3+\boldsymbol{\rho}) & (6+\boldsymbol{\rho}) \end{bmatrix}, \tag{24}$$

and $(C, D) \triangleq (NA, NB)$. Then (C, D) is a multiple of (A, B), and

$$C(\boldsymbol{\rho}) = \begin{bmatrix} 12+10\boldsymbol{\rho}+2\boldsymbol{\rho}^2 & 17+13\boldsymbol{\rho}+2\boldsymbol{\rho}^2 \\ 36+16\boldsymbol{\rho}+2\boldsymbol{\rho}^2 & 51+19\boldsymbol{\rho}+2\boldsymbol{\rho}^2 \end{bmatrix}, \tag{25}$$

$$D(\boldsymbol{\rho}) = \begin{bmatrix} 9+8\boldsymbol{\rho}+2\boldsymbol{\rho}^2 & 12+10\boldsymbol{\rho}+2\boldsymbol{\rho}^2 & 15+12\boldsymbol{\rho}+2\boldsymbol{\rho}^2 \\ 27+14\boldsymbol{\rho}+2\boldsymbol{\rho}^2 & 36+16\boldsymbol{\rho}+2\boldsymbol{\rho}^2 & 45+18\boldsymbol{\rho}+2\boldsymbol{\rho}^2 \end{bmatrix}. \tag{26}$$

Furthermore,

$$\theta(C) = \left[\begin{array}{cc|cc|cc} 12 & 17 & 10 & 13 & 2 & 2 \\ 36 & 51 & 16 & 19 & 2 & 2 \end{array}\right],$$

$$\theta(D) = \left[\begin{array}{ccc|ccc|ccc} 9 & 12 & 15 & 8 & 10 & 12 & 2 & 2 & 2 \\ 27 & 36 & 45 & 14 & 16 & 18 & 2 & 2 & 2 \end{array}\right],$$

where we insert vertical lines in $\theta(\cdot)$ to separate coefficients.

Example 2. Let $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ be given by (25) and (26), respectively. The following example illustrates Algorithm 6.2.

First, the following table displays the normalized singular values ($\bar{\sigma}_i \triangleq \sigma_i/\sigma_{\max}$) of $\mathcal{T}_t(C)$ and

$$\mathcal{I}_t(C) \triangleq \begin{bmatrix} \begin{bmatrix} 0_{p\times pt} & I_p \end{bmatrix} \\ \mathcal{T}_t(C) \end{bmatrix}$$

for $t = 0$ and $t = 1$. Since the ranks both equal 3 for $t = 1$, we move to Step 6.

	$\bar{\sigma}\,[\mathcal{T}_t(C)]$			$\bar{\sigma}\,[\mathcal{I}_t(C)]$		
t	$\bar{\sigma}_2$	$\bar{\sigma}_3$	$\bar{\sigma}_4$	$\bar{\sigma}_2$	$\bar{\sigma}_3$	$\bar{\sigma}_4$
0	6.2×10^{-17}			0.015		
1	0.65	3.7×10^{-4}	2.0×10^{-21}	0.65	1.2×10^{-2}	1.6×10^{-17}

Next, from Step 6, we have that

$$\theta(L) = \left[\begin{array}{cc|cc} -25.5 & 8.5 & 1 & 3 \\ 18 & -6 & -0.7 & -2.1 \end{array}\right],$$

and from Step 7 and 8, we have that

$$\theta(E) = \left[\begin{array}{cc|cc|cc|cc} \epsilon_1 & \epsilon_1 & 1 & \epsilon_1 & 24 & 36 & 8 & 8 \\ \epsilon_2 & \epsilon_2 & \epsilon_3 & 1 & -16.6 & -25 & -5.6 & -5.6 \end{array}\right],$$

$$\theta(F) = \left[\begin{array}{ccc|ccc|ccc|ccc} \epsilon_2 & \epsilon_1 & \epsilon_1 & 5 & 1 & -3 & 16 & 24 & 32 & 8 & 8 & 8 \\ \epsilon_4 & \epsilon_2 & \epsilon_2 & -3 & \epsilon_3 & 3 & -11 & -16.6 & -22.2 & -5.6 & -5.6 & -5.6 \end{array}\right],$$

where

$$\epsilon_1 \triangleq 1.1369 \times 10^{-13},$$

$$\epsilon_2 \triangleq -5.6843 \times 10^{-14},$$

$$\epsilon_3 \triangleq 9.9476 \times 10^{-14},$$

$$\epsilon_4 \triangleq -2.8422 \times 10^{-14}.$$

Therefore, we can see that the multiple $(E, F) = (LC, LD) = (LNA, LNB)$ is comonic.

Example 3. Let $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ be given by (25) and (26), respectively. The following example illustrates Algorithm 6.4.

First, we begin by constructing $L(\boldsymbol{\rho}) = I_3 \otimes C^T(\boldsymbol{\rho})$ and $M(\boldsymbol{\rho}) = \text{vec}\,[D(\boldsymbol{\rho})]^T$. Then, examining the following table, which displays the inverse condition number of W_t for $t = 0$ and $t = 1$, we see that nullity $\left(W_t^T\right) > 0$ for $t = 1$.

t	$\sigma_{\min}\left[W_t^T\right]/\sigma_{\max}\left[W_t^T\right]$
0	8.8×10^{-4}
1	8.9×10^{-17}

Hence, proceeding to Steps 7–10 with $t = 1$, we find that

$$\theta(\gamma) = \left[\begin{array}{c|c} 1 & -1 \end{array}\right],$$

$$\theta(F) = \left[\begin{array}{ccc|ccc} 5 & 1 & -3 & -1 & -1 & -1 \\ -3 & 5 \times 10^{-14} & 3 & -9 \times 10^{-14} & -5 \times 10^{-15} & -4 \times 10^{-14} \end{array}\right].$$

Next, we would like to verify that $(\gamma I_2, F)$ is indeed a multiple of (A, B). To accomplish this, note that if $H \in \mathbb{R}_{\infty}^{2\times 3}[\boldsymbol{\rho}]$ is the Markov parameter polynomial of (A, B), we should have that

$$A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B(\boldsymbol{\rho}),$$
$$\gamma(\boldsymbol{\rho})H(\boldsymbol{\rho}) = F(\boldsymbol{\rho}),$$

and therefore

$$\gamma(\boldsymbol{\rho})A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = A(\boldsymbol{\rho})\gamma(\boldsymbol{\rho})H(\boldsymbol{\rho}) = A(\boldsymbol{\rho})F(\boldsymbol{\rho}) = \gamma(\boldsymbol{\rho})B(\boldsymbol{\rho}).$$

Thus, to compare the accuracy of our computed quasi-scalar comonic multiple, let $\varepsilon_1(\boldsymbol{\rho}) \triangleq A(\boldsymbol{\rho})F(\boldsymbol{\rho}) - \gamma(\boldsymbol{\rho})B(\boldsymbol{\rho})$ and $\varepsilon_2(\boldsymbol{\rho}) \triangleq A(\boldsymbol{\rho})F(\boldsymbol{\rho})$. Then one type of percent error metric is

$$\frac{\|\theta(\varepsilon_1)\|_F}{\|\theta(\varepsilon_2)\|_F} = 1.018 \times 10^{-14},$$

where $\|\cdot\|_F$ denotes the Frobenius norm of $(\cdot)$, and this type of percent error is meant to give us some indication of how far the product $\gamma(\boldsymbol{\rho})B(\boldsymbol{\rho})$ is from $A(\boldsymbol{\rho})F(\boldsymbol{\rho})$. Since this number is small, numerically we have that $A(\boldsymbol{\rho})F(\boldsymbol{\rho}) = \gamma(\boldsymbol{\rho})B(\boldsymbol{\rho})$.

Finally, since $A(\boldsymbol{\rho})F(\boldsymbol{\rho}) = \gamma(\boldsymbol{\rho})B(\boldsymbol{\rho})$, we have that

$$\gamma(\boldsymbol{\rho})B(\boldsymbol{\rho}) = A(\boldsymbol{\rho})F(\boldsymbol{\rho}) = A(\boldsymbol{\rho})\gamma(\boldsymbol{\rho})H(\boldsymbol{\rho}),$$

and hence from Fact 2.7, it follows that $\gamma(\boldsymbol{\rho})H(\boldsymbol{\rho}) = F(\boldsymbol{\rho})$. Furthermore, since $\gamma(\boldsymbol{\rho})$ is nonzero and quasi-scalar, from Fact 2.9, $\gamma(\boldsymbol{\rho})I_2$ has full normal rank. Hence from Theorem 4.2, $(\gamma I_2, F)$ is a comonic quasi-scalar multiple of (A, B).

Remark 9.1. The comonic multiple of (A, B) generated in Example 2 has a higher degree, 3, than the quasi-scalar comonic multiple of (A, B) generated in Example 3, which has a degree of 1. While this may seem counterintuitive since the constraint of generating a quasi-scalar comonic multiple appears to be more restrictive, the reason lies in how the multiple is generated. Specifically, in Algorithm 6.2 (Proposition 6.1 and Example 2), we search for a comonic multiple of (C, D). Hence the degree of the multiple generated by Algorithm 6.2 will always be greater than or equal to the degree of (C, D). However, in Algorithm 6.4 (Proposition 6.3 and Example 3), we search for a quasi-scalar multiple of (A, B) directly, that is, the quasi-scalar comonic multiple (γ, F) of (A, B) is in general not a multiple of (C, D).

Example 4. Let $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ be given by (25) and (26), respectively. The following example illustrates Proposition 6.6.

First, we compute the Markov parameters of (A, B) using the multiples of (A, B) generated in Examples 2 and 3. For both multiples we find that

$$H_0 = \begin{bmatrix} 5 & 1 & -3 \\ -3 & 0 & 3 \end{bmatrix},$$

$$H_1 = \begin{bmatrix} 4 & 0 & -4 \\ -3 & 0 & 3 \end{bmatrix},$$

and $H_i = H_1$ for every $i \geqslant 1$.

Next, computing the error $\varepsilon(\boldsymbol{\rho}) \triangleq A(\boldsymbol{\rho})H(\boldsymbol{\rho}) - B(\boldsymbol{\rho})$, we find that

$$\frac{\|\theta_9(\varepsilon)\|_F}{\|\theta(B)\|_F} = 1.191 \times 10^{-13}.$$

Hence numerically, we find that $A(\boldsymbol{\rho})H(\boldsymbol{\rho}) = B(\boldsymbol{\rho})$, that is, the Markov parameters are indeed the Markov parameters of (A, B).

Example 5. Let $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ be given by (25) and (26), respectively. The following example illustrates Algorithm 7.2.

First, assume that $\bar{n} = 4$ is an upper bound for $n^\star$. Then, since $\bar{n} = 4$, we use the first 9 Markov parameters from Example 4.

Next, the following table displays the third through eighth normalized singular values ($\bar{\sigma}_i \triangleq \sigma_i/\sigma_{\max}$) of $\overline{\mathcal{K}}_{s,\bar{n}}(H)$ and $\mathcal{K}_{s,\bar{n}}(H)$ for $s = 0$ and $s = 1$. Since the ranks are equal for $s = 1$, we proceed to Step 6.

	$\bar{\sigma}\left[\overline{\mathcal{K}}_{s+\bar{n}}(H)\right]$					$\bar{\sigma}\left[\mathcal{K}_{s+\bar{n}}(H)\right]$					
s	$\bar{\sigma}_3$	$\bar{\sigma}_4$	$\cdots$	$\bar{\sigma}_7$	$\bar{\sigma}_8$	$\bar{\sigma}_3$	$\bar{\sigma}_4$	$\bar{\sigma}_5$	$\bar{\sigma}_6$	$\bar{\sigma}_7$	$\bar{\sigma}_8$
0	1					0.04	0.04	8×10^{-16}			
1	0.05	0.05	$\cdots$	0.04	8×10^{-16}	0.05	0.05	0.03	0.03	0.03	3×10^{-15}

Next, from Steps 6–8, we have that

$$\theta(E) = \left[\begin{array}{cc|cc} 1 & 0 & -4/13 & 12/13 \\ 0 & 1 & +3/13 & -9/13 \end{array}\right],$$
$$\theta(F) = \left[\begin{array}{ccc|ccc} +5 & +1 & -3 & -4/13 & -4/13 & -4/13 \\ -3 & -5 \times 10^{-16} & +3 & +3/13 & +3/13 & +3/13 \end{array}\right].$$

Furthermore, letting

$$L(\boldsymbol{\rho}) \triangleq I_2 + \begin{bmatrix} 12/13 & 16/13 \\ 12/13 & 16/13 \end{bmatrix} \boldsymbol{\rho},$$

it follows that $(E, F) = (LA, LB)$. Hence (E, F) is a comonic multiple of (A, B).

Example 6. Let $C(\boldsymbol{\rho})$ and $D(\boldsymbol{\rho})$ be given by (25) and (26), respectively. The following example illustrates Algorithm 7.4.

First, assume that $\bar{n} = 4$ is an upper bound for $n^\star$. Then, since $\bar{n} = 4$, we use the first 9 Markov parameters from Example 4. Furthermore, for every $i \in [0, 8]$, we construct $H_i^\star = \mathrm{vec}\ [H_i]^T$, that is,

$$H_0^\star = \begin{bmatrix} 5 & -3 & 1 & 0 & -3 & 3 \end{bmatrix},$$
$$H_1^\star = \begin{bmatrix} 4 & -3 & 0 & 0 & -4 & 3 \end{bmatrix},$$

and so on.

Next, examining the following table, which displays the inverse condition number of $\mathcal{K}_{s,\bar{n}}(H^\star)$ for $s = 0$ and $s = 1$, we see that nullity $\left[\mathcal{K}_{s,\bar{n}}^T(H^\star)\right] > 0$ for $s = 1$.

s	$\sigma_{\min}\left[\mathcal{K}_{s,\bar{n}}(H^\star)\right]/\sigma_{\max}\left[\mathcal{K}_{s,\bar{n}}(H^\star)\right]$
0	0.042
1	2.9×10^{-15}

Hence, proceeding to Steps 5–8 with $s = 1$, we find that

$$\theta(\gamma) = \left[\begin{array}{c|c} 1 & -1 \end{array}\right],$$

$$\theta(F) = \left[\begin{array}{ccc|ccc} 5 & 1 & -3 & -1 & -1 & -1 \\ -3 & 5\times 10^{-16} & 3 & 3\times 10^{-13} & 8\times 10^{-17} & -3\times 10^{-13} \end{array}\right],$$

which is similar to the quasi-scalar comonic multiple $(\gamma I_2, F)$ generated in Example 3 up to rounding errors.

Finally, as in Example 3, we should find that $A(\boldsymbol{\rho})F(\boldsymbol{\rho}) = \gamma(\boldsymbol{\rho})B(\boldsymbol{\rho})$. Thus, letting $\varepsilon_1(\boldsymbol{\rho}) \triangleq A(\boldsymbol{\rho})F(\boldsymbol{\rho}) - \gamma(\boldsymbol{\rho})B(\boldsymbol{\rho})$ and $\varepsilon_2(\boldsymbol{\rho}) \triangleq A(\boldsymbol{\rho})F(\boldsymbol{\rho})$, we find that

$$\frac{\|\theta\,(\varepsilon_1)\|_F}{\|\theta\,(\varepsilon_2)\|_F} = 2.041\times 10^{-14},$$

and hence, numerically we have that $A(\boldsymbol{\rho})F(\boldsymbol{\rho}) = \gamma(\boldsymbol{\rho})B(\boldsymbol{\rho})$. Furthermore, as in Example 3, we find that this implies that $(\gamma I_2, F)$ is a comonic quasi-scalar multiple of (A, B).

Remark 9.2. As evidenced by the previous examples, all of the proposed algorithms require, at some point, one to determine the rank of a matrix, which is always a very delicate task, even for these small examples. Furthermore, we do not suggest rigid guidelines for choosing tolerances for rank conditions, since presumably these choices would be motivated by the problem at hand, specifically the conditioning of the problem. For instance, suppose that a row or column of the Markov parameter polynomial was significantly smaller than the others. Then the results would be influenced by the practitioner's determination whether the row or column in question is due to round-off errors or not.

Remark 9.3. In the examples presented here, access to the original system allows us to ascertain the accuracy of the computed object. However, this is not possible for the practitioner, who may need to develop reliability tests. These should be motivated by how the end object is to be used. For instance, if the practitioner has access to the Markov parameters of a system, and computes a multiple of (A, B) from the Markov parameters, one could save the final x Markov parameters, that is, not include them in the algorithms, then check how small $A(\boldsymbol{\rho})H(\boldsymbol{\rho}) - B(\boldsymbol{\rho})$ is using the saved Markov parameters. However, if one is interested in the accuracy of the spectral content of the system (A, B), then some other test may be required.

10. Conclusions

We have considered polynomial matrix representations of MIMO linear systems and their connection to Markov parameters. Specifically, we have developed theory and numerical algorithms for transforming polynomial matrix models into Markov parameter models, and vice verse. We have also provided numerical examples to illustrate the given algorithms.

Acknowledgements

The Ph.D of the first author was funded by the NASA Aeronautics Scholarship Program, through which this effort was made possible. Furthermore, we would like to thank the reviewers for suggesting shorter alternative proofs to some of the facts in the paper.

References

[1] P.J. Antsaklis, A.N. Michel, Linear Systems, Birkhauser, Boston, 2006.
[2] D.S. Bernstein, Matrix Mathematics, second ed., Princeton University Press, Princeton, NJ, 2009.
[3] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, Time Series Analysis: Forecasting and Control, fourth ed., Wiley, 2008.
[4] E.F. Camacho, C. Bordons, Model Predictive Control, Springer-Verlag, London, 2004.
[5] I. Gohberg, P. Lancaster, L. Rodman, Matrix Polynomials, SIAM, 2009.
[6] G.C. Goodwin, K.S. Sin, Adaptive Filtering Prediction and Control, Dover, Mineola, New York, 2009.
[7] B.L. Ho, R.E. Kalman, Effective construction of linear state-variable models from input/output functions, Regelungstechnik 14 (12) (1966) 545–592.
[8] P. Ioannou, B. Fidan, Adaptive Control Tutorial, SIAM, Philadelphia, 2006.
[9] J.N. Juang, Applied System Identification, Prentice-Hall, Englewood Cliffs, NJ, 1994.
[10] T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ, 1980.
[11] P. Lancaster, M. Tismenetsky, The Theory of Matrices, second ed., Academic Press, 1985.
[12] L. Ljung, System Identification: Theory for the User, Prentice-Hall, Upper Saddle River, NJ, 1999.
[13] J.M. Maciejowski, Predictive Control with Constraints, Prentice Hall, Englewood Cliffs, NJ, 2002.
[14] R.H. Middleton, G.C. Goodwin, Digital Control and Estimation: A Unified Approach, Prentice-Hall, Englewood Cliffs, NJ, 1990.
[15] J.E. Normay-Rico, E.F. Camacho, Control of Dead-Time Processes, Springer-Verlag, London, 2007.
[16] R. Pintelon, J. Schoukens, System Identification: A Frequency Domain Approach, Wiley-IEEE Press, 2005.
[17] E. Reynders, R. Pintelon, G. De Roeck, Consistent impulse–response estimation and system realization from noisy data, IEEE Trans. Signal Process. 56 (2008) 2696–2705.
[18] W.J. Rugh, Linear System Theory, Prentice-Hall, Englewood Cliffs, NJ, 1996.
[19] W.J. Rugh, Nonlinear System Theory: The Volterra/Wiener Approach, The Johns Hopkins University Press, 1981.
[20] T. Soderstrom, P. Stoica, System Identification, Prentice-Hall, Upper Saddle River, NJ, 1988.
[21] Z. Szabo, P.S.C. Heuberger, J. Bokora, P.M.J. Van den Hof, Extended Ho–Kalman algorithm for systems represented in generalized orthonormal bases, Automatica 36 (2000) 1809–1818.
[22] A.I.G. Vardulakis, Linear Multivariable Control: Algebraic Analysis and Synthesis Methods, Wiley, 1991.
[23] W.A. Wolovich, Linear Multivariable Systems, Springer-Verlag, New York, NY, 1974.
[24] D. Xue, Y. Chen, D.P. Atherton, Linear Feedback Control: Analysis and Design with Matlab, SIAM, 2007.