KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Authors: Lukas Vanpoucke, Bart De Moor Title:Constructing multidimensional difference equations from a state-space representation using the generalized Cayley-Hamilton theorem Journal: IEEE Control Systems Letters Vol: 8 Pages: 7 (6 in Journal typesetting)
Archived version	☐ Submitted version ☐ Accepted (Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher) ☐ Published ■ Published (Open Access)
Acknowledgements	In this manuscript, the following funding is acknowledged: ☐ EU H22020 ■ EU Horizon Europe ■ FWO ■ KULeuven Internal Funds ■ VLAIO ☐ AI ■ ERC ■ Other
Published version	https://ieeexplore.ieee.org/document/10552341
Copyright statement	978-1-5386-5541-2/18/$31.00 ©2018 IEEE
Author contact	lukas.vanpoucke@esat.kuleuven.be



Constructing multidimensional difference equations from a state-space representation using the generalized Cayley-Hamilton Theorem

Lukas Vanpoucke, and Bart De Moor, *Fellow, IEEE, IFAC & SIAM*

***Abstract*— We show that applying a generalization of the Cayley-Hamilton Theorem to a state-space representation of a single-output, multidimensional (mD), linear, shift-invariant, causal, autonomous system, with distinct eigentuples results in an equivalent difference equation representation. The proposed method is also applicable to parameterize a set of mD shift-invariant difference equations in terms of a given set of eigenvalues. Lastly, a closed-form expression in terms of the eigenvalues of the state-transition matrices is derived.**

I. Introduction

Discrete-time, linear, time-invariant (LTI) models provide a great modeling tool for many engineering applications. Nowadays, multidimensional data, where many other variables besides time are considered, is abundantly available. To this end, many of the tools provided by LTI systems have been generalized to the multidimensional (mD) setting. In this work, we focus on the class of single-output, mD, linear, shift-invariant, causal, autonomous models, as described by the state-space representation in Equation (1), where $\boldsymbol{e}_k \in \mathbb{Z}^m$ denotes the kth column of the identity matrix, the multi-index of an m-dimensional gridpoint is denoted by[1] $\boldsymbol{\kappa} \in \mathbb{N}_0^m$, $\boldsymbol{A}_k \in \mathbb{R}^{n\times n}$, $k = 1, \ldots, m$ and $\boldsymbol{C} \in \mathbb{R}^{1\times n}$:

$$\begin{aligned} &\boldsymbol{x}_{\boldsymbol{\kappa}+\boldsymbol{e}_1} = \boldsymbol{A}_1\boldsymbol{x}_{\boldsymbol{\kappa}};\ \ldots;\ \boldsymbol{x}_{\boldsymbol{\kappa}+\boldsymbol{e}_m} = \boldsymbol{A}_m\boldsymbol{x}_{\boldsymbol{\kappa}};\ y_{\boldsymbol{\kappa}} = \boldsymbol{C}\boldsymbol{x}_{\boldsymbol{\kappa}} \\ &\text{where } \forall i,j \in \{1,\ldots,m\} : \boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{A}_j\boldsymbol{A}_i. \end{aligned} \tag{1}$$

This model class, known by various different names, has been studied extensively in the literature, [1], [2], [3], [4]. Through Equation (1), any initial state vector $\boldsymbol{x_0} \in \mathbb{R}^n$, where n is the model order, defines an output signal $y_{\boldsymbol{\kappa}} : \mathbb{N}_0^m \to \mathbb{R}; \boldsymbol{\kappa} \mapsto y_{\boldsymbol{\kappa}}$ over the whole positive orthant $\mathbb{N}_0^m$. Note that the commutativity of the state-transition matrices $\boldsymbol{A}_k \in \mathbb{R}^{n\times n}$ is required for the well-definedness of the state and output signals. For example, in 2D, when stepping from $\boldsymbol{x}_{(0,0)}$ to $\boldsymbol{x}_{(1,1)}$, the outcome should be path-independent, such that $\boldsymbol{A}_2\boldsymbol{A}_1\boldsymbol{x}_{(0,0)} = \boldsymbol{A}_2\boldsymbol{x}_{(1,0)} = \boldsymbol{x}_{(1,1)} = \boldsymbol{A}_1\boldsymbol{x}_{(0,1)} = \boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{x}_{(0,0)}$. Consequently, by grouping the state-transition matrices in the tuple $\boldsymbol{A} = (\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m)$ and introducing a shorthand notation[2], the model output can be compactly written as: $y_{\boldsymbol{\kappa}} = \boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}}\boldsymbol{x_0} = \boldsymbol{C}\prod_{k=1}^{m}\boldsymbol{A}_k^{\kappa_k}\boldsymbol{x_0}$. This shows that the model dynamics are invariant under a common similarity transform $\boldsymbol{T}$ on the state-transition matrices. Indeed, defining $\boldsymbol{A}_{\text{new}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}, \boldsymbol{x}_{\boldsymbol{0}}^{\text{new}} = \boldsymbol{T}^{-1}\boldsymbol{x_0}$ and $\boldsymbol{C}_{\text{new}} = \boldsymbol{C}\boldsymbol{T}$, the output remains: $y_{\boldsymbol{\kappa}}^{\text{new}} = \boldsymbol{C}_{\text{new}}\boldsymbol{A}_{\text{new}}^{\boldsymbol{\kappa}}\boldsymbol{x}_{\boldsymbol{0}}^{\text{new}} = \boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}}\boldsymbol{x_0}$.

In the current work, we consider the problem of constructing an equivalent difference equation representation for a given, possibly parametric, single-output, autonomous state-space model (1), with distinct eigentuples, that is eigenvalues of the state-transition matrices grouped into m-tuples[3]. Indeed, akin to the 1D case of LTI systems, it has been shown in [3] that the considered class of mD models can be equivalently described using difference equations with constant coefficients, i.e., shift-invariant, also called partial difference equations. For a minimal, single-output, autonomous, LTI state-space model, the Cayley-Hamilton Theorem implies that the characteristic polynomial of the state-transition matrix $\boldsymbol{A}$, readily expressed in terms of the eigenvalues, forms an equivalent difference equation representation for the model [5]. This contribution aims to generalize this procedure to the mD setting.

A similar problem has been tackled in [5] using the Buchberger-Möller algorithm [6], briefly illustrated in Example I.1. From a system theoretic point of view, the algorithm in [5] constructs a multidimensional *observability matrix* $\boldsymbol{\Gamma}$[4] [4], whose rows are $\boldsymbol{\Gamma}_{l,:} = \boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}_l}$, for a given set of N gridpoints $\{\boldsymbol{\kappa}_l |\ l = 1, ..., N\}$ and an ordering on them. The columns of an observability matrix then form a basis for the model-compliant signals $y_{\boldsymbol{\kappa}}$ restricted to the given gridpoints. The algorithm then sequentially searches for a set of linearly independent gridpoints $\boldsymbol{\kappa}$ to serve as *initial states*, from which linear recurrence relations, a so-called *Gröbner basis*, can be derived that define the signal values in subsequent gridpoints. A drawback of this approach is that the resulting set of initial states may vary depending on the precise eigenvalues of the provided system matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ [7]. In cases where the precise values of these eigenvalues are unknown, e.g., when optimizing over them as model parameters to find those that optimally fit the given data [8], this may lead to invalid parameterizations of the considered models, as illustrated in the 2D Example I.1.

Example I.1. Consider the following state-space model with eigentuples, (a_1, b_1) and $(a_2, b_2) \in \mathbb{C}^2$ and system matrices:

$$\boldsymbol{A}_1 = \begin{bmatrix} a_1 & \\ & a_2 \end{bmatrix}, \boldsymbol{A}_2 = \begin{bmatrix} b_1 & \\ & b_2 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

This work was supported in part by the KU Leuven: Research Fund (Projects iBOF/23/064, C3/20/117 and C3I-21-00316), Industrial Research Fund and several LRD bilateral industrial projects; in part by Flemish Government agencies: FWO (SBO project S005319N and TBM Project T001919N); EWI (Flanders AI Research Program); and VLAIO (CSBO project HBC.2021.0076); and in part by the European Commission (ERC Adv. Grant) under Grant 885682.

The authors are with the Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS), Department of Electrical Engineering (ESAT), KU Leuven, 3001 Leuven, Belgium (e-mail: lukas.vanpoucke@esat.kuleuven.be; bart.demoor@esat.kuleuven.be).

[1] **Notation:** $\mathbb{N}_0$ denotes the set of natural numbers including zero and $\mathbb{N}$ is assumed to not include zero.

[2] **Notation:** the shorthand notation $\boldsymbol{x}^{\boldsymbol{\alpha}} = \prod_i x_i^{\alpha_i}$ will be used for general multivariate monomials described by the multi-exponent (multi-index) $\boldsymbol{\alpha}$.

[3] This grouping is defined precisely in Section III, after Lemma III.1

[4] **Notation:** The index ':' indicates all rows or columns are selected.

To find a set of difference equations that the outputs $y_{\boldsymbol{\kappa}}$ of this model must satisfy, we first construct a multidimensional observability matrix $\boldsymbol{\Gamma}$. To this end, we define an appropriate ordering on the gridpoints, for instance, consider the ordering used in [4], which, in (2D), is : $(0,0) \prec (1,0) \prec (0,1) \prec (2,0) \prec (1,1), \ldots$. Then, we respectively arrange the output signals $y_{\boldsymbol{\kappa}} = \boldsymbol{x_0}(\boldsymbol{A}^{\boldsymbol{\kappa}})\boldsymbol{C}$ corresponding to the initial states $\boldsymbol{x_0}^{\mathsf{T}} = [1 \quad 0]$ and $\boldsymbol{x_0}^{\mathsf{T}} = [0 \quad 1]$ in the rows of $\boldsymbol{\Gamma}^{\mathsf{T}}$ below. Note that, due to the linearity of the model, the rows of $\boldsymbol{\Gamma}^{\mathsf{T}}$ form a basis for the model-compliant output signals, since any initial state can be written as a linear combination of the two considered states. Considering the points $\boldsymbol{\kappa} \in \mathbb{N}_0^2$, where $\|\boldsymbol{\kappa}\|_1 \leq 2$, this results in the following matrix:

$$\boldsymbol{\Gamma}^{\mathsf{T}} = \begin{matrix} \boldsymbol{y}_{(0,0)} & \boldsymbol{y}_{(1,0)} & \boldsymbol{y}_{(0,1)} & \boldsymbol{y}_{(2,0)} & \boldsymbol{y}_{(1,1)} & \boldsymbol{y}_{(0,2)} \end{matrix} \\ \begin{bmatrix} 1 & a_1 & b_1 & a_1^2 & a_1 b_1 & b_1^2 \\ 1 & a_2 & b_2 & a_2^2 & a_2 b_2 & b_2^2 \end{bmatrix}.$$

When $a_1 \neq a_2$, the first and second column of $\boldsymbol{\Gamma}^{\mathsf{T}}$ are linearly independent. Clearly, all subsequent columns can then be written as linear combinations of these independent columns, e.g., $(a_1 - a_2)\boldsymbol{y}_{(0,1)} = (a_1 b_2 - a_2 b_1)\boldsymbol{y}_{(0,0)} + (b_1 - b_2)\boldsymbol{y}_{(1,0)}$ and $\boldsymbol{y}_{(2,0)} = (a_1 + a_2)\boldsymbol{y}_{(1,0)} - a_1 a_2 \boldsymbol{y}_{(0,0)}$, where $\boldsymbol{y}_{\boldsymbol{\kappa}} \in \mathbb{R}^{2\times 1}$. Note that due to the multi-shift invariance of the model outputs, these relations can be 'shifted' throughout the grid $\mathbb{N}_0^m$ by left multiplying by ${\boldsymbol{A}_k}^{\mathsf{T}}$, leading to linear relations that define $y_{\boldsymbol{\kappa}}$ for $\boldsymbol{\kappa} \in \mathbb{N}_0^2$ with $\boldsymbol{y}_{(0,0)}, y_{(1,0)}$ as initial values. However, in case $a_1 = a_2$, the relation $\boldsymbol{y}_{(1,0)} = a_1 \boldsymbol{y}_{(0,0)}$ holds, so that $\boldsymbol{y}_{(1,0)}$ is linearly dependent on $\boldsymbol{y}_{(0,0)}$ and thus cannot be considered an initial value. Instead, $\boldsymbol{y}_{(0,0)}$ and $\boldsymbol{y}_{(0,1)}$ then act as initial values if $b_1 \neq b_2$.

When considering the eigentuples as variables, it is not know which case applies, though. As it is the generic case, one might use the recurrence equations provided by the case $a_1 \neq a_2$ as a parameterization for the model in terms of these variables. Then, substituting $a_1 = a_2$ into these recurrence relations yields $(b_1 - b_2)y_{(1,0)} = a_1(b_1 - b_2)y_{(0,0)}$ and $y_{(2,0)} = 2a_1 y_{(1,0)} - a_1^2 y_{(0,0)}$, which acts as a set of 1D difference equations. Consequently, all signals of the form $y_{\boldsymbol{\kappa}} = a_1^{\kappa_1} y_{\kappa_2}$, where y_{κ_2} can be any 1D signal, satisfy this set of difference equations, as will be explained in Example I.3. As such, these equations then describe an infinite-order systems, instead of the intended second-order system. In essence, the parameterization 'breaks down' in this non-generic case, describing an infinite order model, which does not allow a finite dimensional state-space representation (1), instead of the intended second-order model [3], [7].

Aside from the theoretical value in offering an alternative transition from a state-space to a difference equation representation, the approach presented in this work provides the advantage of a more general parameterization of models w.r.t. the Gröbner-based methods. The presented construction relies on combining the properties of commuting matrices, with the Generalized Cayley-Hamilton Theorem, introduced in Section II. To this end, we first explain the link between multivariate polynomials and mD difference equations.

Formally, difference equations can be generalized to the multidimensional setting by introducing the linear shift operators: $\sigma_i y_{\boldsymbol{\kappa}} = y_{\boldsymbol{\kappa}+\boldsymbol{e}_i}$, that a set of difference equations can be characterized by a multivariate polynomial matrix in these operators $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_m]$, as illustrated by the following example of 2D homogeneous difference equations.

Example I.2. Consider the following set of 2D difference equations with constant coefficients in the scalar signal $y_{\boldsymbol{\kappa}}$:

$$\begin{bmatrix} y_{\boldsymbol{\kappa}} - y_{\boldsymbol{\kappa}+(1,1)} + y_{\boldsymbol{\kappa}+(2,0)} \\ 6y_{\boldsymbol{\kappa}} - 5y_{\boldsymbol{\kappa}+(0,1)} + y_{\boldsymbol{\kappa}+(0,2)} \end{bmatrix} = \begin{bmatrix} 1 - \sigma_1\sigma_2 + \sigma_1^2 \\ 6 - 5\,\sigma_2 + \sigma_2^2 \end{bmatrix} y_{\boldsymbol{\kappa}} \\ = \boldsymbol{R}(\boldsymbol{\sigma})y_{\boldsymbol{\kappa}} = \boldsymbol{0},\ \forall \boldsymbol{\kappa} \in \mathbb{N}_0^2.$$

In the *behavioral framework* for systems theory, model-compliant signals are defined as signals that satisfy such a set of difference equations at each gridpoint $\boldsymbol{\kappa} \in \mathbb{N}_0^m$, i.e., any $y_{\boldsymbol{\kappa}}$ that lies in the kernel of the difference operator, i.e., $\boldsymbol{R}(\boldsymbol{\sigma})y_{\boldsymbol{\kappa}} = \boldsymbol{0}, \forall \boldsymbol{\kappa} \in \mathbb{N}_0^m$. Hence, this is type of representation is a kernel representation. Due to the linearity of the shift operators $\boldsymbol{\sigma}$, this kernel forms a vector space over $\mathbb{C}$, which can be thought of as the column space of an infinitely long mD observability matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{\infty\times n}$, i.e., constructed using the whole grid $\mathbb{N}_0^m$ [9], [10], [11]. Since the model class considered in this paper admits a finite-dimensional state-space system description as described in Equation (1), the kernel of the operator $\boldsymbol{R}(\boldsymbol{\sigma})$ that characterizes the system dynamics is also finite-dimensional [3].

As previously stated, we restrict ourselves to single-output, autonomous systems, such that this leads to homogeneous scalar difference equations. For the considered class of difference equations, a basis for the kernel of $\boldsymbol{R}(\boldsymbol{\sigma})$ is readily obtained by filling in an exponential of the form $y_{\boldsymbol{\kappa}} = \boldsymbol{\lambda}^{\boldsymbol{\kappa}}$ and subsequently solving $\boldsymbol{R}(\boldsymbol{\sigma})$ as a set of multivariate polynomials, as illustrated in Example I.3 below [12].

Example I.3. Substituting a signal of the form $y_{\boldsymbol{\kappa}} = \lambda_1^{\kappa_1}\lambda_2^{\kappa_2}$ into the difference equations from Example I.2 we obtain:

$$\begin{cases} \lambda_1^{\kappa_1}\lambda_2^{\kappa_2}\left(1 - \lambda_1\lambda_2 + \lambda_1^2\right) = 0 \\ \lambda_1^{\kappa_1}\lambda_2^{\kappa_2}\left(6 - 5\lambda_2 + \lambda_2^2\right) = 0 \end{cases} \quad \forall \boldsymbol{\kappa} \in \mathbb{N}_0^2,$$

which corresponds to the set of polynomial equations $\boldsymbol{R}(\lambda_1, \lambda_2)$ after dividing away the factor $\lambda_1^{\kappa_1}\lambda_2^{\kappa_2}$.

Example I.3 illustrates that when the polynomials $\boldsymbol{R}(\boldsymbol{\sigma})$ has a finite number of common roots $\boldsymbol{\lambda}_i, i = 1, \ldots, n$, its is n-dimensional with $\boldsymbol{\lambda}_i^{\boldsymbol{\kappa}}$ as its basis signals. These roots then correspond to the eigentuples of an equivalent state-space model [3], [4]. As such, the considered problem, that is, constructing an equivalent difference equation representation $\boldsymbol{R}(\boldsymbol{\sigma})$ for a given state-space model (1), can be thought of as constructing a set of multivariate polynomials that precisely describes a given a set of n mD points as its set of common roots. In general, these basis functions are not necessarily of the form $\boldsymbol{\lambda}^{\boldsymbol{\kappa}}$, for example, when there are eigentuples with a multiplicity greater than one [4]. Such cases are not considered in this work, however.

The rest of this paper is organized as follows: Section II introduces the generalized Cayley-Hamilton Theorem, and the multivariate polynomials associated with it. Then, Section III shows that, under an observability-like condition, for systems with distinct eigentuples, these polynomials are a

difference equation representation of a single-output, mD, linear, shift-invariant, causal, autonomous system. A closed-form expression for this set of difference equations in terms of the eigentuples is then presented in Section IV.

II. The generalized Cayley-Hamilton Theorem

Like the classical Cayley-Hamilton Theorem, the Generalized Cayley-Hamilton (GCH) Theorem, as first presented in [13], describe relations between sets of matrices and their higher-order powers. The ideas behind this construction are illustrated in Example II.1 which is quite simple, but it generalizes easily. While the given matrices are not required to commute for the GCH Theorem, since it will be applied to the state-transition matrices, we present a version adapted to commuting matrices in Theorem II.1 to avoid confusion.

Example II.1. Reconsider the state-space model from Example I.1 and define the matrix $\boldsymbol{B} = c_1\boldsymbol{A}_1 + c_2\boldsymbol{A}_2,\ c_k \in \mathbb{C}$. Now compute the characteristic polynomial of $\boldsymbol{B}$: $\chi(x) = x^2 + x(-b_2c_2 - b_1c_2 - a_2c_1 - a_1c_1) + (b_1c_2 + a_1c_1)(b_2c_2 + a_2c_1)$. Introducing the variables λ_1 and λ_2 and substituting $x = c_1\lambda_1 + c_2\lambda_2$, results in the following equation:

$$\begin{aligned}&\chi(c_1\lambda_1 + c_2\lambda_2) = \\ &\underbrace{(\lambda_1^2 - \lambda_1(a_2+a_1) + a_1a_2)}_{\mu_{(2,0)}(\lambda_1,\lambda_2)}c_1^2 + \underbrace{(\lambda_2^2 - \lambda_2(b_2+b_1) + b_1b_2)}_{\mu_{(0,2)}(\lambda_1,\lambda_2)}c_2^2 \\ &+ \underbrace{(2\lambda_1\lambda_2 - \lambda_2(a_2+a_1) - \lambda_1(b_2+b_1) + a_1b_2 + a_2b_1)}_{\mu_{(1,1)}(\lambda_1,\lambda_2)}c_1c_2,\end{aligned} \tag{2}$$

where the multi-index $\boldsymbol{\alpha}$ of the coefficient $\mu_{\boldsymbol{\alpha}}$ refers to the monomial $\boldsymbol{c}^{\boldsymbol{\alpha}}$ with which it is associated. Due to the Cayley-Hamilton Theorem, expression (2) evaluated in $(\lambda_1, \lambda_2) = (\boldsymbol{A}_1, \boldsymbol{A}_2)$, equals zero as a polynomial in c_1, c_2, implying that each of its coefficients $\mu_{\boldsymbol{\alpha}}$ are zero. Consequently, it holds that $\boldsymbol{0} = \mu_{(2,0)}(\boldsymbol{A}_1, \boldsymbol{A}_2) = \mu_{(1,1)}(\boldsymbol{A}_1, \boldsymbol{A}_2) = \mu_{(0,2)}(\boldsymbol{A}_1, \boldsymbol{A}_2)$. Note that $\mu_{(2,0)}$ and $\mu_{(0,2)}$ are the characteristic polynomials of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, respectively.

This construction can be easily extended to the general case of m matrices $\boldsymbol{A}_k$ and variables c_k, leading to a set of polynomials $\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda})$, giving rise to the generalized Cayley-Hamilton Theorem, Theorem II.1. Note that only monomials $\boldsymbol{c}^{\boldsymbol{\alpha}}$ of exactly degree n occur in $\chi(\sum_k c_k\boldsymbol{\lambda}_k)$, which can be readily deduced from the cofactor expansion of the determinant. To this end, we denote the set of monomials in m variables and of exactly degree n as $\mathcal{A}_n = \{\boldsymbol{\alpha} |\ \boldsymbol{\alpha} \in \mathbb{N}_0^m, \|\boldsymbol{\alpha}\|_1 = n\}$, implying that the GCH construction results in $\#\mathcal{A}_n = \binom{m+n-1}{m-1}$ equations.

Theorem II.1 (GCH [13], [14])**.** *Given m pairwise commuting matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathbb{C}^{n\times n}$. Let χ be the characteristic polynomial of the polynomial matrix $\sum_{k=1}^m c_k\boldsymbol{A}_k = \boldsymbol{B} \in \mathbb{C}^{n\times n}[\boldsymbol{c}]$. Define the variables $\boldsymbol{\lambda}$ and define the polynomials $\mu_{\boldsymbol{\alpha}}$ by writing $\chi(\sum_k c_k\lambda_k)$ such that:*

$$\chi(\sum_k c_k\lambda_k) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}_n} \mu_{\boldsymbol{\alpha}}(\lambda_1, ..., \lambda_m)\boldsymbol{c}^{\boldsymbol{\alpha}}. \tag{3}$$

It then holds that $\forall\boldsymbol{\alpha} \in \mathcal{A}_n : \mu_{\boldsymbol{\alpha}}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m) = \boldsymbol{0}$.

III. Obtaining difference equations through the Generalized Cayley-Hamilton Theorem

As illustrated in Example II.1, the application of the Generalized Cayley-Hamilton Theorem to a set of commuting matrices can be used to build a set of multivariate polynomials $\{\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda})|\ \boldsymbol{\alpha} \in \mathcal{A}_n\}$ with a finite subset of predefined roots, i.e., the eigentuples of the given matrices. Such a set of multivariate polynomials then naturally leads to a set of homogeneous, single-output mD difference equations by evaluating them in the shift operators $\boldsymbol{\sigma}$, as discussed in Section I. In this section, we show that the eigentuples of the state-transition matrices, when distinct, correspond exactly to the common roots of $\mu_{\boldsymbol{\alpha}}$, implying that both representations describe the same set of model outputs. This is first illustrated in Example III.1 below.

Example III.1. Reconsider the polynomials $\boldsymbol{R}(\boldsymbol{\sigma})^\mathsf{T} = \begin{bmatrix}\mu_{(2,0)}(\boldsymbol{\sigma}) & \mu_{(1,1)}(\boldsymbol{\sigma}) & \mu_{(0,2)}(\boldsymbol{\sigma})\end{bmatrix}$ as defined in (2). While it is obvious from Theorem II.1 that (a_1, b_1) and (a_2, b_2) are common roots of these polynomials, it is not immediately clear that they form the complete set of common roots. Indeed, using the same grid ordering as in Example I.1, writing out $\boldsymbol{R}(\boldsymbol{\sigma})y_{(0,0)} = \boldsymbol{0}$ readily leads to values for $y_{(2,0)}, y_{(1,1)}, y_{(0,2)}$, which thus seems to define a recurrence relation with $y_{(0,0)}, y_{(1,0)}, y_{(0,1)}$ as initial values. This would imply that there are three basis signals instead of the two of the original model in Example I.1. Indeed, the difference equations $\boldsymbol{R}(\boldsymbol{\sigma})$ impose restrictions on the signal at all gridpoints, i.e., $\forall\boldsymbol{\kappa} \in \mathbb{N}_0^2 : \boldsymbol{R}(\boldsymbol{\sigma})y_{\boldsymbol{\kappa}} = \boldsymbol{R}(\boldsymbol{\sigma})\boldsymbol{\sigma}^{\boldsymbol{\kappa}}y_{(0,0)} = \boldsymbol{\sigma}^{\boldsymbol{\kappa}}\boldsymbol{R}(\boldsymbol{\sigma})y_{(0,0)} = \boldsymbol{0}$, not only in $\boldsymbol{\kappa} = (0,0)$. For example, let $(a_1, b_1) = (\frac{1+i}{\sqrt{2}}, i)$ and $(a_2, b_2) = (\frac{1-i}{\sqrt{2}}, -i)$ in Equation (2), we can then write out the difference equations at the gridpoints $\|\boldsymbol{\kappa}\|_1 \leq 1$ as a linear system:

$$\boldsymbol{0} = \boldsymbol{M}\boldsymbol{y} = \begin{array}{c}\\ \boldsymbol{R}(\boldsymbol{\sigma}) \\ \\ \\ \hline \\ \sigma_1\boldsymbol{R}(\boldsymbol{\sigma}) \\ \\ \\ \hline \\ \sigma_2\boldsymbol{R}(\boldsymbol{\sigma}) \\ \\ \\ \\ \end{array} \left[\begin{array}{ccccccccccc} 1 & -\sqrt{2} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \sqrt{2} & 0 & -\sqrt{2} & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & -\sqrt{2} & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & -\sqrt{2} & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & -\sqrt{2} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & 0 & -\sqrt{2} & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} y_{(0,0)} \\ y_{(1,0)} \\ y_{(0,1)} \\ y_{(2,0)} \\ y_{(1,1)} \\ y_{(0,2)} \\ y_{(3,0)} \\ y_{(2,1)} \\ y_{(1,2)} \\ y_{(0,3)} \end{array}\right].$$

Clearly, constructing $\boldsymbol{y}$ appropriately from any signal $y_{\boldsymbol{\kappa}}$ in the kernel of $\boldsymbol{R}(\boldsymbol{\sigma})$ solves this system. Moreover, the right null space of $\boldsymbol{M}$ is of dimension two, such that the basis signals $(\frac{1+i}{\sqrt{2}})^{\kappa_1}(i)^{\kappa_2}$ and $(\frac{1-i}{\sqrt{2}})^{\kappa_1}(-i)^{\kappa_2}$ generate two basis vectors for it. Note that these basis vectors correspond to the columns of a 10×2 extended version of the observability matrix $\boldsymbol{\Gamma}$ from Example I.1, from which it was deduced that for the given ordering, $y_{(0,0)}$ and $y_{(1,0)}$ act as initial values. Now, like in Example I.1 consider the case where $a_1 = a_2$ and $b_1 \neq b_2$, e.g., let $(a_1, b_1) = (1, 1)$ and $(a_2, b_2) = (1, 2)$, a similar analysis then reveals that the basis signals $1^{\kappa_1}1^{\kappa_2}$ and $1^{\kappa_1}2^{\kappa_2}$ again lead to two basis vectors $\boldsymbol{y}$ spanning null$(\boldsymbol{M})$. Thus, as discussed in Example I.1 $y_{(0,0)}$ and $y_{(0,1)}$ act as initial values and the GCH parameterization does not break down.

In general, as long as enough gridpoints are considered, i.e., $\|\boldsymbol{\kappa}\|_1 \leq d$ for d 'sufficiently large', it has been shown that for any set of polynomials $\boldsymbol{R}(\boldsymbol{\sigma})$, the kernel of the associated *Macaulay matrix* $\boldsymbol{M}$ is spanned by vectors similar to $\boldsymbol{y}$, generated by the basis signals[5] $\boldsymbol{\lambda}_i^{\boldsymbol{\kappa}}$, where $\boldsymbol{\lambda}_i$ are the common roots of the multivariate polynomials $\boldsymbol{R}(\boldsymbol{\sigma})$ [15].

We now extend the observation of Example III.1 to general n and m. First, note that the state-transition matrices in (1) can be brought into upper triangular form through a common Schur transform, as stated in Lemma III.1. Notice that this reveals the eigenvalues of the state-transition matrices on the diagonals, allowing us to group them into eigentuples.

Lemma III.1 ([16])**.** ***Schur triangularization for commuting matrices:*** *Let $\boldsymbol{A}_k \in \mathbb{C}^{n\times n}, k = 1,\dots,m$ be a set of commuting matrices, then there exists a unitary $\boldsymbol{Q} \in \mathbb{C}^{n\times n}$ such that $\boldsymbol{T}_k = \boldsymbol{Q}^{-1}\boldsymbol{A}_k\boldsymbol{Q}$ are upper triangular.*

For the practical computation of $\boldsymbol{Q}$, we refer the reader to [17]. Now, by Lemma III.1 and since $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q} = \boldsymbol{Q}^{-1}(c_1\boldsymbol{A}_1 + \cdots + c_m\boldsymbol{A}_m)\boldsymbol{Q} = c_1\boldsymbol{T}_1 + \cdots + c_m\boldsymbol{T}_m$, the characteristic polynomial of $\boldsymbol{B}$ and, by extension, the GCH equations defined in Theorem II.1 are invariant under a common similarity transform on the matrices $\boldsymbol{A}_k$, so that we can assume these to be upper triangular without loss of generality. Denoting the eigentuples, i.e., the m-tuples of ith diagonal entries, as $\boldsymbol{a}_i = [(\boldsymbol{T}_1)_{i,i} \quad (\boldsymbol{T}_2)_{i,i} \quad \cdots \quad (\boldsymbol{T}_m)_{i,i}]^\mathsf{T}$, $i = 1,\dots,n$, the characteristic polynomial of $\boldsymbol{B}$ is then $\chi(x) = \prod_{i=1}^n (x - \boldsymbol{c}\cdot\boldsymbol{a}_i)$, where $\cdot$ denotes the dot product, since the determinant of an upper triangular matrix is equal to the product of its diagonal entries. Finally, to construct the GCH equations, the variables $\boldsymbol{\lambda} \in \mathbb{C}^m$ are substituted into the polynomial:

$$\chi(\boldsymbol{c}\cdot\boldsymbol{\lambda}) = \prod_{i=1}^{n}(\boldsymbol{c}\cdot\boldsymbol{\lambda} - \boldsymbol{c}\cdot\boldsymbol{a}_i) = \prod_{i=1}^{n}\boldsymbol{c}\cdot(\boldsymbol{\lambda} - \boldsymbol{a}_i). \tag{4}$$

This can then be rewritten to obtain the set of m-variate polynomials $\{\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}) \mid \boldsymbol{\alpha} \in \mathcal{A}_n\}$, as constructed in Theorem II.1, whose common roots of are described in Theorem III.1.

Theorem III.1. *Given a set of m pairwise commuting matrices $\mathcal{M} = \{\boldsymbol{A}_k \in \mathbb{C}^{n\times n} \mid k = 1,\dots,m\}$. As in Lemma III.1, jointly triangularize these matrices resulting in $\boldsymbol{T}_k$ and define the eigentuples as $\boldsymbol{a}_i = [(\boldsymbol{T}_1)_{i,i} \quad \cdots \quad (\boldsymbol{T}_m)_{i,i}]^\mathsf{T}$, where $i = 1,\dots,n$. Let the set of $\binom{m+n-1}{m-1}$ equations $\{\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}) \mid \boldsymbol{\alpha} \in \mathcal{A}_n\}$ be the GCH polynomials constructed from $\mathcal{M}$, as described in Theorem II.1. The set of affine common solutions[6] of these GCH polynomials is then precisely $V = \{\boldsymbol{a}_i \mid i = 1,\dots,n\}$.*

Moreover, when $\boldsymbol{a}_1,\dots,\boldsymbol{a}_n$ comprise n unique tuples, i.e., $\forall i,j \in \{1,\dots,n\},\ i \neq j:\ \boldsymbol{a}_i \neq \boldsymbol{a}_j$, each solution in V has a multiplicity of precisely one.

[5]The common roots of the polynomials are revealed by the eigentuples of the state-transition matrices, which can be obtained through the shift equations $\boldsymbol{\Gamma}_{\boldsymbol{\kappa}_l,:}\boldsymbol{A}_k = \boldsymbol{\Gamma}_{\boldsymbol{\kappa}_l+\boldsymbol{e}_k,:}$ built from the observability matrix $\boldsymbol{\Gamma}$ in the null space of a Macaulay matrix.

[6]Additional solutions in projective space, i.e., *solutions at infinity*, are excluded since homogenizing with a variable λ_0 and setting $\lambda_0 = 0$ results in $\mu_{\boldsymbol{\alpha}}(0,\boldsymbol{\lambda}) = \boldsymbol{\lambda}^{\boldsymbol{\alpha}}$ (see Section IV), which only vanishes in $(\lambda_0,\boldsymbol{\lambda}) = \boldsymbol{0}$.

Proof. As previously noted, the GCH polynomials are invariant under common similarity transforms applied to $\boldsymbol{A}_k$. We thus assume the upper triangular format without loss of generality. Let $\boldsymbol{\lambda}^*$ be a solution for the set of polynomials $\{\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}) \mid \boldsymbol{\alpha} \in \mathcal{A}_n\}$. By Equations (4) and (3), this implies that $\forall\ \boldsymbol{c} \in \mathbb{C}^m:\ \chi(\boldsymbol{c}\cdot\boldsymbol{\lambda}^*) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}_n}\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}^*)\cdot\boldsymbol{c}^{\boldsymbol{\alpha}} = \prod_{i=1}^n \boldsymbol{c}\cdot(\boldsymbol{\lambda}^* - \boldsymbol{a}_i) = 0$. This can only equal zero, regardless of the value of $\boldsymbol{c}$, when $\exists i \in \{1,\dots,n\}:\ \boldsymbol{\lambda}^* - \boldsymbol{a}_i = \boldsymbol{0}$.

To prove the absence of higher-order multiplicities when there are n unique solutions, let $\boldsymbol{v}\cdot\nabla_{\boldsymbol{\lambda}}$ be any linear combination of first-order partial differentiation operators w.r.t. the variables $\boldsymbol{\lambda}$, that is, $\nabla_{\boldsymbol{\lambda}} = \left[\frac{\partial}{\partial\lambda_1} \quad \cdots \quad \frac{\partial}{\partial\lambda_m}\right]^\mathsf{T}$ and $\boldsymbol{v} \in \mathbb{C}^m$. Applying this differentiation operator to $\chi(\boldsymbol{c}\cdot\boldsymbol{\lambda})$ leads to the following expression, where we denote $\frac{\mathrm{d}\chi(x)}{\mathrm{d}x}$ by $\chi'(x)$:

$$\begin{aligned}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{\lambda}})\chi(\boldsymbol{c}\cdot\boldsymbol{\lambda}) &= \sum_{\boldsymbol{\alpha}\in\mathcal{A}_n}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{\lambda}})\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda})\boldsymbol{c}^{\boldsymbol{\alpha}} \\ &= \chi'(\boldsymbol{c}\cdot\boldsymbol{\lambda})(\boldsymbol{v}\cdot\nabla_{\boldsymbol{\lambda}})(\boldsymbol{c}\cdot\boldsymbol{\lambda}) = \chi'(\boldsymbol{c}\cdot\boldsymbol{\lambda})(\boldsymbol{v}\cdot\boldsymbol{c}).\end{aligned} \tag{5}$$

Now assume that the set of equations has a multiplicity of at least two for some solution $\boldsymbol{\lambda}^*$, that is, $\exists\boldsymbol{v} \in \mathbb{C}^m\backslash\{0\}$ such that $\forall\boldsymbol{\alpha}:\ \mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}^*) = (\boldsymbol{v}\cdot\nabla_{\boldsymbol{\lambda}})\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}^*) = 0$ [18]. It then follows that $\forall\boldsymbol{c} \in \mathbb{C}^m:\ (\boldsymbol{v}\cdot\nabla_{\boldsymbol{\lambda}})\chi(\boldsymbol{c}\cdot\boldsymbol{\lambda}^*) = 0$, due to the first equality in (5). Then, due to the third equality in (5), it is clear that: $\forall\boldsymbol{c} \in \mathbb{C}^m$ such that $\boldsymbol{c}\cdot\boldsymbol{v} \neq 0$ it holds that $\chi'(\boldsymbol{c}\cdot\boldsymbol{\lambda}^*) = 0$.

Applying the product rule for differentiation to $\chi'(\boldsymbol{c}\cdot\boldsymbol{\lambda})$ results in:

$$\chi'(\boldsymbol{c}\cdot\boldsymbol{\lambda}) = \sum_{j=1}^{n}\prod_{i=1,i\neq j}^{n}\boldsymbol{c}\cdot(\boldsymbol{\lambda} - \boldsymbol{a}_i). \tag{6}$$

Since any solution $\boldsymbol{\lambda}^*$ is an element of V, i.e., $\boldsymbol{\lambda}^* = \boldsymbol{a}_k$ for some $k \in \{1,\dots,n\}$. However, when $\boldsymbol{a}_1,\dots,\boldsymbol{a}_n$ comprise n unique tuples, all the factors $\boldsymbol{c}\cdot(\boldsymbol{\lambda} - \boldsymbol{a}_i)$ are unique, implying that when we evaluate (6) in $\boldsymbol{\lambda} = \boldsymbol{a}_k$, the kth term is the only one without the factor $\boldsymbol{c}\cdot(\boldsymbol{\lambda} - \boldsymbol{a}_k)$ and thus the only one to not vanish. Additionally, when $\boldsymbol{\lambda} = \boldsymbol{a}_k$ it holds that $\boldsymbol{\lambda} - \boldsymbol{a}_i \neq 0$, where $i \neq k$, such that the kth term cannot vanish for all values of $\boldsymbol{c}$ in the open set $\{\boldsymbol{c} \in \mathbb{C}^m : \boldsymbol{c}\cdot\boldsymbol{v} \neq 0\}$, thus contradicting the initial assumption that a solution to the equations with a multiplicity of at least two exists. □

Remark: Note that Equation (6) implies that in case there is a multiple in the set of diagonal entries, i.e., $\exists i,j \in \{1,\dots,n\}, i \neq j:\ \boldsymbol{a}_i = \boldsymbol{a}_j$, $\chi'(\boldsymbol{c}\cdot\boldsymbol{\lambda})$ will vanish in said solution point $\boldsymbol{a}_i$, regardless of $\boldsymbol{c}$. By (5), this implies that all the first-order partial derivatives $(\boldsymbol{v}\cdot\nabla_{\boldsymbol{\lambda}})\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda})$ vanish in $\boldsymbol{a}_i$, implying a multiplicity of at least $m+1$, while generally, multiplicities in the range 2 to m can be achieved [18]. As such, parameterizations based on the GCH construction may describe models of an order higher than n in these cases, though not of infinite order. Moreover, the parameterization is more general than those obtained by the Buchberger-Möller algorithm, since for any choice of initial values, the so-called *normal set*, an infinite, measure zero collection of sets of distinct eigentuples exists for which this normal set is invalid, as illustrated in Example I.1. [7]

Lastly, we show that, under a condition akin to observability, the state-space model (1) and the difference equations

built from the GCH equations describe exactly the same set of output signals. Theorem III.2 shows this equivalence.

Theorem III.2. *Let $\boldsymbol{A}_k \in \mathbb{C}^{n\times n}, k = 1,\dots,m$ be a set of upper triangular commuting matrices, define a state-space model* (1) *with these matrices and $\boldsymbol{C} \in \mathbb{R}^{1\times n}$, and denote $\boldsymbol{a}_i = [(\boldsymbol{A}_1)_{i,i} \quad \cdots \quad (\boldsymbol{A}_m)_{i,i}]^\mathsf{T}$. If $\forall i,j \in \{1,\dots,n\},\ i \neq j:\ \boldsymbol{a}_i \neq \boldsymbol{a}_j$, and there exists a finite set of gridpoints $\boldsymbol{\kappa}_l$, $l = 1,\dots,N$ such that the matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{N\times n} : \boldsymbol{\Gamma}_{l,:} = \boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}_l}$, is of rank n, then, the set of signals described by the GCH difference equations constructed from $\boldsymbol{A}_k$ in Theorem III.1 equals the set of signals described by the state-space model.*

Proof. Let $\boldsymbol{R}(\boldsymbol{\sigma}) \in \mathbb{C}^{\#\mathcal{A}_n\times 1}[\boldsymbol{\sigma}]$ denote the GCH difference equations. As established in Section I and due to Theorem III.1 any signal $y_{\boldsymbol{\kappa}} : \mathbb{N}_0^m \to \mathbb{R}; \boldsymbol{\kappa} \mapsto y_{\boldsymbol{\kappa}}$ that is model-compliant, i.e., $\boldsymbol{R}(\boldsymbol{\sigma})y_{\boldsymbol{\kappa}} = \boldsymbol{0}$, can be written as $y_{\boldsymbol{\kappa}} = \sum_{i=1}^n \gamma_i \boldsymbol{a}_i^{\boldsymbol{\kappa}}$, where $\gamma_i \in \mathbb{C}$, such that the associated kernel is the span of signals $\mathcal{Y}_{GCH} = \text{span}(\boldsymbol{a}_1^{\boldsymbol{\kappa}},\dots,\boldsymbol{a}_n^{\boldsymbol{\kappa}})$. The outputs of the state-space model lie in the span of signals $\text{span}(\boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}}\boldsymbol{e}_1,\dots,\boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}}\boldsymbol{e}_n) = \mathcal{Y}_{ss}$, where $\boldsymbol{e}_i$ is the ith column of the identity matrix. All entries of the vector signal $\boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}}$ then lie in $\mathcal{Y}_{GCH}$, since by Theorem II.1, $\forall \boldsymbol{\kappa} \in \mathbb{N}_0^m : \boldsymbol{R}(\boldsymbol{\sigma})\boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}} = \boldsymbol{C}\boldsymbol{R}(\boldsymbol{A}_1,\dots,\boldsymbol{A}_m)\boldsymbol{A}^{\boldsymbol{\kappa}} = \boldsymbol{0}$. Consequently, $\mathcal{Y}_{ss}$ is a linear subspace of $\mathcal{Y}_{GCH}$, more precisely: $\exists \boldsymbol{V} \in \mathbb{C}^{n\times n} : \forall \boldsymbol{\kappa} \in \mathbb{N}_0^m : [\boldsymbol{a}_1^{\boldsymbol{\kappa}} \quad \cdots \quad \boldsymbol{a}_n^{\boldsymbol{\kappa}}]\boldsymbol{V} = \boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}}$. Now, $\boldsymbol{V}$ is of rank n and thus invertible, because $\boldsymbol{\Gamma}$ is of rank n and $\forall l : \boldsymbol{\Gamma}_{l,:} = [\boldsymbol{a}_1^{\boldsymbol{\kappa}_l} \quad \cdots \quad \boldsymbol{a}_n^{\boldsymbol{\kappa}_l}]\boldsymbol{V}$. This implies that $\forall \boldsymbol{\kappa} \in \mathbb{N}_0^m : [\boldsymbol{a}_1^{\boldsymbol{\kappa}} \quad \cdots \quad \boldsymbol{a}_n^{\boldsymbol{\kappa}}] = \boldsymbol{C}\boldsymbol{A}^{\boldsymbol{\kappa}}\boldsymbol{V}^{-1}$, such that $\mathcal{Y}_{GCH}$ is a linear subspace[7] of $\mathcal{Y}_{ss}$, and thus $\mathcal{Y}_{GCH} = \mathcal{Y}_{ss}$. □

IV. Closed-form expression

In this section, we derive an explicit expression for the GCH polynomials $\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda})$ in terms of their given set of roots. The construction of a specific equation $\mu_{\boldsymbol{\alpha}}$ can then be viewed as a combinatorial problem, which will first be outlined using the Example IV.1, Afterwards, we derive a general formulation in similar fashion. For an easier exposition, we first introduce the function $\text{count}(\boldsymbol{k}, j) : \mathbb{N}_0^n \times \mathbb{N} \to \mathbb{N}_0; (\boldsymbol{k}, j) \mapsto \#\{i \in \{1,\dots,n\} |\ k_i = j\}$, which counts the number of occurrences of j in the n-tuple $\boldsymbol{k}$.

Now, let $a_{i,k}$ denote the kth entry of $\boldsymbol{a}_i$, then by (3) and (4), it holds that:

$$\sum_{\boldsymbol{\alpha}\in\mathcal{A}_n} \mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}) \cdot \boldsymbol{c}^{\boldsymbol{\alpha}} = \prod_{i=1}^n \left(\sum_{k=1}^m c_k(\lambda_k - a_{i,k}) \right). \tag{7}$$

Example IV.1. Consider the GCH polynomials constructed from the matrices $\boldsymbol{A}_1 = \text{diag}(1,2,3)$ and $\boldsymbol{A}_2 = \text{diag}(3,2,1)$, such that $\boldsymbol{a}_1 = [1 \quad 3]^\mathsf{T}, \boldsymbol{a}_2 = [2 \quad 2]^\mathsf{T}$ and $\boldsymbol{a}_3 = [3 \quad 1]^\mathsf{T}$ and $\boldsymbol{c} = [c_1 \quad c_2]^\mathsf{T}$. To keep the exposition brief, we focus on building the single polynomial $\mu_{(2,1)}$, i.e., for $\boldsymbol{\alpha} = (2,1)$. For notational conciseness, denote $d_{i,k} = (\lambda_k - a_{i,k})$ and expand the right-hand side of (7):

$$(c_1 d_{1,1} + c_2 d_{1,2})(c_1 d_{2,1} + c_2 d_{2,2})(c_1 d_{3,1} + c_2 d_{3,2}). \tag{8}$$

[7]Note that $\boldsymbol{V}^{-1}$ essentially transforms the state-space outputs into the outputs of a model defined by the diagonal matrices $\text{diag}(\boldsymbol{A}_k)$, implying that a similarity transform exists that jointly diagonalizes the matrices $\boldsymbol{A}_k$.

This expression can be expanded into eight individual terms, each corresponding to one of the four third-degree monomials in $\boldsymbol{c}$. To construct $\mu_{(2,1)}$, we filter out the terms associated with $\boldsymbol{c}^{(2,1)} = c_2^2 c_1$, as stated in the construction procedure in Theorem II.1, resulting in:

$$\mu_{(2,1)} = \underbrace{d_{1,1}d_{2,1}d_{3,2}}_{(1,1,2)} + \underbrace{d_{1,1}d_{2,2}d_{3,1}}_{(1,2,1)} + \underbrace{d_{1,2}d_{2,1}d_{3,1}}_{(2,1,1)}. \tag{9}$$

The $(1,1,2)$-term in (9) then corresponds to selecting the first term in the first factor of (8), the first term in the second factor and the second term in the third factor, as encoded by the tuple $(1,1,2)$. The selection pattern is similarly encoded for the other terms. Notice that for each tuple $\boldsymbol{k}$ associated with a term in (9), it holds that $\text{count}(\boldsymbol{k},1) = 2$ and $\text{count}(\boldsymbol{k},2) = 1$, ensuring the product involves a factor $\boldsymbol{c}^{(2,1)}$.

Now, we further expand each of these terms. Let us focus on the $(1,1,2)$-term. Filling in $d_{i,k} = (\lambda_k - a_{i,k})$ results in:

$$d_{1,1}d_{2,1}d_{3,2} = (\lambda_1 - 1)(\lambda_1 - 2)(\lambda_2 - 1). \tag{10}$$

To further expand this in terms of monomials in (λ_1, λ_2), one has to choose between the constant or the λ_k-term in each of these three factors. We focus on the λ_k terms first. Notice that λ_1 can achieve any exponent $\beta_1 \in \{0,1,2\}$ and, similarly, λ_2 can achieve any exponent $\beta_2 \in \{0,1\}$ in the terms of the expansion of Equation (10). In short, $\boldsymbol{\beta} = (\beta_1, \beta_2) \leq \boldsymbol{\alpha} = (2,1)$, where $\leq$ denotes elementwise comparison. A Similar observation holds for the other selections patterns.

The latter observation shows which (λ_1, λ_2)-monomials occur in $\mu_{(2,1)}$. As such, we now consider a specific monomial, for instance, $\boldsymbol{\lambda}^{(1,1)}$, and attempt to construct its coefficient. Note that, by the full expansion of Equations (8), (7), this monomial can only be obtained through a factor of the form $\boldsymbol{c}^{(1,1)}\boldsymbol{\lambda}^{(1,1)}$, which still needs to be multiplied by a factor involving c_1 to arrive at the required factor $\boldsymbol{c}^{(2,1)}$. Since we fixed the monomial $\boldsymbol{\lambda}^{(1,1)}$, we can only choose factors of the form $-c_k a_{i,k}$ to achieve this. As such, disregarding the λ_k-part of $d_{i,k}$, we can inspect eligible parts of the individual factors of (8):

$$(-c_1 - 3c_2), (-2c_1 - 2c_2), (-3c_1 - c_2). \tag{11}$$

Clearly, to achieve the goal of multiplying by c_1, we can select the first term in either the first, second or third factor, selection patterns that can be encoded as $(1,0,0), (0,1,0)$ and $(0,0,1)$, respectively. Notice that for these selections $\boldsymbol{k}$: $\text{count}(\boldsymbol{k},1) = 1$ and $\text{count}(\boldsymbol{k},2) = 0$. Summing over all these contributions to the coefficient of $\boldsymbol{\lambda}^{(1,1)}$ then yields: $-1-2-3 = -6$. However, fully expanding (9) results in:

$$\mu_{(2,1)} = -26 + 26\lambda_1 + 11\lambda_2 - 6\lambda_1^2 - 12\lambda_1\lambda_2 + 3\lambda_1^2\lambda_2.$$

The missing factor two results from the different permutations by which the monomial $\boldsymbol{\lambda}^{(1,1)}$ can be obtained. For example: selecting the term $-c_1$ corresponds to the selection pattern $(1,0,0)$ in (11). For this pattern, one can either select $c_1\lambda_1$ as the second factor and $c_2\lambda_2$ as the third, or vice versa, as both choices result in a monomial $\boldsymbol{c}^{(2,1)}$ when multiplied by $-c_1$.

We now derive an expression for the general case. Using distributivity, Equation (7) can be expanded into m^n terms of the form $\prod_i^n c_{k_i}(\lambda_{k_i} - a_{i,k_i})$ for some selection of k-indices: $(k_1, \ldots, k_n) \in \mathbb{N}^n$ and $k_i \leq m$. To arrive at a specific monomial $\boldsymbol{c}^{\boldsymbol{\alpha}}$ in this expansion, one has to make a selection such that: for every $j = 1, \ldots, m$, the jth term is chosen precisely α_j times over all the n factors. This set of $\boldsymbol{\alpha}$-compatible can thus be formally described as $\mathcal{S}^n_{\boldsymbol{\alpha}} = \{\boldsymbol{k} \in \mathbb{N}_0^n \mid k_i \leq m, \forall j \in \{1, \ldots, m\} : \text{count}(\boldsymbol{k}, j) = \alpha_j\}$.

A term corresponding to some $\boldsymbol{\alpha}$-compatible selection can then be expanded further by choosing between the $c_k\lambda_k$ and $-c_k a_{i,k}$ terms in each of the n factors. First, we focus on the former type of terms. Observe that the products of subselections of the $c_k\lambda_k$-type terms lead to contributions of the form $f(\boldsymbol{a}_1, ..., \boldsymbol{a}_n)\boldsymbol{c}^{\boldsymbol{\beta}}\boldsymbol{\lambda}^{\boldsymbol{\beta}}$ for any $\boldsymbol{\beta} \in \mathbb{N}_0^m$ and $\boldsymbol{\beta} \leq \boldsymbol{\alpha}$. Now, given that a factor $\boldsymbol{c}^{\boldsymbol{\beta}}\boldsymbol{\lambda}^{\boldsymbol{\beta}}$ has been obtained through some subselection of terms, it is clear that a factor $\boldsymbol{c}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}$ is still required to arrive at the monomial $\boldsymbol{c}^{\boldsymbol{\alpha}}$. Any $(\boldsymbol{\alpha} - \boldsymbol{\beta})$-compatible subselection of $n - \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|_1$ factors $-c_k a_{i,k}$ provides the required factor $\boldsymbol{c}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}$. Such a selection can be encoded by a tuple $\boldsymbol{k} \in \mathcal{S}^n_{\boldsymbol{\alpha}-\boldsymbol{\beta}}$, in which a zero entry k_i indicates selection of a $c_k\lambda_k$-type factor at position i.

Finally, note that there are multiple subselections by which $\boldsymbol{\beta}$ can be formed. To account for this, one has to count the number of $\boldsymbol{\beta}$-compatible selections out of $\|\boldsymbol{\beta}\|_1$ positions, i.e., $\#S^{\|\boldsymbol{\beta}\|_1}_{\boldsymbol{\beta}}$. To do so, first consider the number of selections of β_1 positions out of $\|\boldsymbol{\beta}\|_1$, i.e., $\binom{\sum_{s\geq 1}\beta_s}{\boldsymbol{\beta}_1}$; next, consider the number of selections of β_2 positions out of $\|\boldsymbol{\beta}\|_1 - \beta_1$, i.e., $\binom{\sum_{s\geq 2}^m \beta_s}{\boldsymbol{\beta}_2}$, and so on. Taking the product then results in a total of $\prod_{j=1}^m \binom{\sum_{s\geq j}^m \beta_s}{\beta_j}$ possibilities.

Combining all the observations above then leads to the following formula in terms of the parameters for the equation associated with multi-index $\boldsymbol{\alpha}$:

$$\mu_{\boldsymbol{\alpha}}(\boldsymbol{\lambda}) = \sum_{\boldsymbol{0}\leq\boldsymbol{\beta}\leq\boldsymbol{\alpha}} \prod_{j=1}^{m} \binom{\sum_{s\geq j}^m \beta_s}{\beta_j} \boldsymbol{\lambda}^{\boldsymbol{\beta}} \sum_{\boldsymbol{k}\in\mathcal{S}_{\boldsymbol{\alpha}-\boldsymbol{\beta}}} (-1)^{\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|_1} \prod_{i:k_i\neq 0}^{n} a_{i,k_i}.$$

To construct the set $\mathcal{S}^n_{\boldsymbol{\alpha}-\boldsymbol{\beta}}$, first construct all selections of $(\boldsymbol{\alpha} - \boldsymbol{\beta})_1$ out of n positions and assign '1' to the selected positions; then, combine this with all possible selections of $(\boldsymbol{\alpha} - \boldsymbol{\beta})_2$ out of the remaining $n - (\boldsymbol{\alpha} - \boldsymbol{\beta})_1$ positions and assign '2'; and so on.

V. Conclusion and future work

It was shown that the Generalized Cayley-Hamilton Theorem of [13] can be used to construct a difference equation representation of a given state-space representation of a single-output, mD, linear, shift-invariant, causal, autonomous system with simple eigentuples. Additionally, the proposed methodology allows for the parameterization of a difference equation representation in terms of a general, finite set of n affine eigentuples of multiplicity precisely one — or equivalently, the parameterization of a set of multivariate polynomial equations in terms of its n solutions with multiplicity one. This generality distinguishes it from the existing approaches in the literature [5], [6], which make a choice of initial states that depends on the precise eigentuples at hand, leading to discontinuities for parameterizations when considering these tuples as model parameters [7]. Additionally, we derived a closed-form expression for the resulting set of equations, expanding each equation in terms of its monomials and the corresponding coefficients.

Since general systems of multivariate polynomials can have both so-called *solutions at infinity* in projective space and solutions with some multiplicity structure [4], we aim to extend the presented methodology to deal with these more general cases in future work. Additionally, we hope to extend this approach to multiple-output systems and systems with inputs. Furthermore, we aim to apply this parameterization to the data-driven estimation of the difference equations [8], [19], To this end, it is imperative that we investigate the numerical accuracy and scalability of this approach.

References

[1] S. Attasi, "Modelling and recursive estimation for double indexed sequences," in *System Identification Advances and Case Studies*. Amsterdam, The Netherlands: Elsevier, 1976, vol. 126, pp. 289–348.

[2] J. A. Ball and V. Vinnikov, "Overdetermined multidimensional systems: State space and frequency domain methods," in *Mathematical Systems Theory in Biology, Communications, Computation, and Finance*. New York, NY, USA: Springer, 2003, pp. 63–119.

[3] E. Fornasini, P. Rocha, and S. Zampieri, "State space realization of 2-D finite-dimensional behaviours," *SIAM J. Control Optim*, vol. 31, no. 6, pp. 1502–1517, 1993.

[4] P. Dreesen, K. Batselier, and B. De Moor, "Multidimensional realisation theory and polynomial system solving," *Int. J. Control*, vol. 91, no. 12, pp. 2692–2704, 2018.

[5] K. Batselier and N. Wong, "Computing the state difference equations for discrete overdetermined linear mD systems," *Automatica*, vol. 64, pp. 254–261, 2016.

[6] H. M. Möller and B. Buchberger, "The construction of multivariate polynomials with preassigned zeros," in *Computer Algebra*. Berlin, Germany: Springer, 1982, pp. 24–31.

[7] H. J. Stetter, *Numerical Polynomial Algebra*. Philadelphia, PA, USA: SIAM, 2004.

[8] B. De Moor, "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem," *Commun. Inf. Syst.*, vol. 20, no. 2, p. 163–207, 2020.

[9] P. Rocha and J. C. Willems, "State for 2-D systems," *Linear Algebra Appl.*, vol. 122-124, pp. 1003–1038, Sep.–Nov. 1989.

[10] E. Zerz, *Topics in Multidimensional Linear Systems Theory*. London, U.K.: Springer, 2000.

[11] U. Oberst, "Multidimensional constant linear systems," *Acta Applicandae Mathematicae*, vol. 20, pp. 1–175, Jul. 1990.

[12] P. E. Hydon, *Difference Equations by Differential Equation Methods*. Cambridge, U.K.: Cambridge Univ., 2014.

[13] B. G. Mertzios and M. A. Christodoulou, "On the generalized Cayley-Hamilton theorem," *IEEE Trans. Autom. Control*, vol. 31, no. 2, pp. 156–157, Feb. 1986.

[14] T. Kaczorek, "Generalization of Cayley-Hamilton theorem for n-D polynomial matrices," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 671–674, May 2005.

[15] K. Batselier, P. Dreesen, and B. De Moor, "On the null spaces of the Macaulay matrix," *Linear Algebra Appl.*, vol. 460, pp. 259–289, Nov. 2014.

[16] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge U.K.: Cambridge Univ., 2013.

[17] R. M. Corless, P. M. Gianni, and B. M. Trager, "A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots," in *Proc. Int. Symp. Symb. Algebraic Comput.*, 1997, p. 133–140.

[18] B. H. Dayton and Z. Zeng, "Computing the multiplicity structure in solving polynomial systems," *Proc. Int. Symp. Symb. Algebraic Comput.*, pp. 116–123, 2005.

[19] C. Vermeersch, "The (block) Macaulay matrix: Solving systems of multivariate polynomial equations and multiparameter eigenvalue problems," Ph.D. dissertation, KU Leuven, Leuven, Belgium, 2023.