KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Authors: Sarthak De, Bart De Moor Title: State Elimination in Polynomial Models: A Linear Algebraic Approach Conference: IFAC Journal of Systems and Control Vol: 36 Pages: NA
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 ☐ Other
Published version	`https://doi.org/10.1016/j.ifacsc.2026.100413`
Copyright statement	©2026 De, S. and De Moor, B.. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Author contact	sarthak.de@esat.kuleuven.be

State Elimination in Polynomial Models: A Linear Algebraic Approach ⋆

Sarthak De * **, Bart De Moor,** ***Fellow, IFAC, IEEE & SIAM*** *

* *Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS), Dept. of Electrical Engineering (ESAT), KU Leuven. Kasteelpark Arenberg 10, B-3001 Leuven, Belgium (e-mail: {sarthak.de,bart.demoor}@esat.kuleuven.be).*

Abstract: We derive a method to obtain the difference equation for discrete-time, autonomous, time-invariant, single-output polynomial state-space models, by proposing a linear algebraic framework for state elimination based on the novel "model matrix". For strongly locally observable models we show that the left null space of the model matrix is spanned by the rows of the Macaulay matrix associated with the corresponding output difference polynomial. We propose a state elimination algorithm that uses the singular value decomposition to compute a basis of the left null space and directly recovers the output difference polynomial. The method naturally accommodates floating-point coefficients. Numerical examples demonstrate the effectiveness of the proposed framework.

Keywords: Nonlinear system identification, Realization theory, Data driven control theory

1. INTRODUCTION

This paper considers the problem of state elimination, going from a given polynomial state space model to a non-trivial polynomial relation among consecutive model outputs. We will consider the class of discrete-time, autonomous, time-invariant, single-output polynomial state-space models described as [1]

$$\Sigma : \begin{cases} \boldsymbol{x}[k+1] &= \boldsymbol{f}(\boldsymbol{x}[k]) \\ y[k] &= g(\boldsymbol{x}[k]) \end{cases}, \tag{1}$$

where $\boldsymbol{x}[k] = \begin{bmatrix} x_1[k] \cdots x_{n_f}[k] \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{n_f}$ is the state variable at time $k \in \mathbb{N}$ and the state update function $\boldsymbol{f} = \begin{bmatrix} f_1 \cdots f_{n_f} \end{bmatrix}^{\mathrm{T}}$, $f_i \in \mathbb{R}[\boldsymbol{x}[k]]$ with $\mathbb{R}[\boldsymbol{x}[k]]$ the ring of polynomials over $\mathbb{R}$ in the components of the state variables $\boldsymbol{x}[k]$. Here, $g \in \mathbb{R}[\boldsymbol{x}[k]]$ is the observation function and $y[k] \in \mathbb{R}$ is the observed model output.

Example 1. Consider the discretized predator-prey model Lotka (1920), which is an instance of the model class Σ in (1),

$$\Sigma_{\mathrm{PP}} : \begin{cases} x_1[k+1] = 3x_1[k] - 0.01x_1[k]x_2[k], \\ x_2[k+1] = -0.1\, x_2[k] + 0.01x_1[k]x_2[k], \\ y[k] = x_2[k], \end{cases} \tag{2}$$

where the components of the state variable $\boldsymbol{x}[k] = [x_1[k]\ x_2[k]]^{\mathrm{T}}$ denote the population sizes of the prey and the predator, respectively. △

We denote the length $n_y + 1$ window of outputs starting at time k by $\boldsymbol{y}[k : k+n_y] = [y[k] \cdots y[k+n_y]]^{\mathrm{T}} \in \mathbb{R}^{n_y+1}$ where $n_y \in \mathbb{N}$ is the *order bound* . The output monomials are denoted by $\boldsymbol{y}[k : k+n_y]^{\boldsymbol{\omega}} = \prod_{i=0}^{n_y} y[k+i]^{\omega_i}$ where $\boldsymbol{\omega} \in \mathbb{N}^{n_y+1}$. Difference equation(s) are nontrivial linear combinations of output monomials which vanish over the window $\boldsymbol{y}[k : k+n_y]$ of an output sequence $\boldsymbol{y} = [y[0] \cdots y[\mathrm{N}]] \in \mathbb{R}^{\mathrm{N}+1}$ generated by Σ. The difference equation $\phi(\boldsymbol{y}[k : k+n_y]) = 0$ is of the form,

$$\phi = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}}\, \boldsymbol{y}[k{:}k{+}n_y]^{\boldsymbol{\omega}} \ = \ 0. \tag{3}$$

where $\phi \in \mathbb{R}[\boldsymbol{y}[k : k+n_y]]$ is the output difference polynomial in the components of $\boldsymbol{y}[k : k+n_y]$. The order of ϕ ($\mathrm{ord}(\phi)$) is the largest difference between the shifts of the outputs in (3). It is at most n_y. Here, $\boldsymbol{\omega} = (\omega_0, \ldots, \omega_{n_y}) \in \Omega$ is a multi-index where $\Omega = \{\boldsymbol{\omega} \in \mathbb{N}^{n_y+1} : c_{\boldsymbol{\omega}} \in \mathbb{R} \setminus \{0\}\}$, where $c_{\boldsymbol{\omega}}$ are the coefficients of the output monomials. The degree of ϕ is given as $\deg(\phi) = \max_{\boldsymbol{\omega} \in \Omega} \|\boldsymbol{\omega}\|_1$.

Example 1. (continued). The difference equation for the predator-prey model (2) is given by $\phi_{\mathrm{PP}}(\boldsymbol{y}[k : k+2]) = 0$ where the output difference polynomial,

$$\begin{aligned} \phi_{\mathrm{PP}} = -20y[k]y[k+1] + 100y[k]y[k+2] - 300y[k+1]^2 \\ + 0.10y[k]^2y[k+1] + y[k]y[k+1]^2, \end{aligned} \tag{4}$$

is of order two and degree three. △

Problem Definition We study *state elimination*: *given* a single-output polynomial state-space model Σ with n_f state components, *find* a nontrivial polynomial relation among consecutive outputs (a difference equation) that holds along all output sequences of Σ. Equivalently, we convert the

⋆ This work was supported in part by the KU Leuven Research Fund (IDN/24/013, iBOF/23/064), the Industrial Research Fund and several LRD bilateral industrial projects; in part by FWO (TBM T001725N; PhD grant 11K5623N); in part by EWI via the Flanders AI Research Program; in part by VLAIO (HBC.2021.0076); in part by FOD VVVL (AI-943); and in part by the ERC under the EU Horizon 2020 programme (Advanced Grant 885682).

[1] Throughout, square brackets $[k]$ indicate discrete-time indexing of the concerned variable (scalar or vector); we do not repeat "at time $k \in \mathbb{N}$" when the index is explicit.

model from its *image* representation, the state–space form in (1), where the output is a polynomial of the state, to its *kernel* representation, the difference equation in (3), which specifies a vanishing polynomial relation across successive outputs.

For models in Σ, the first $n_y + 1$ model outputs at time instant k are evaluated by multivariate polynomial functions in the components of $\boldsymbol{x}[k]$ as

$$y[k + n_y] = g(\boldsymbol{f}^{n_y}(\boldsymbol{x}[k])), n_y \geq 0, \tag{5}$$

where $\boldsymbol{f}^{n_y}(\boldsymbol{x}[k]) = \boldsymbol{f}(\boldsymbol{f}^{n_y-1}(\boldsymbol{x}[k]))$ and $\boldsymbol{f}^0(\boldsymbol{x}[k]) = \boldsymbol{x}[k]$.

Example 1. (continued). The first three model outputs of Σ_{PP} (2) at time instant k as multivariate polynomial functions in the components of $\boldsymbol{x}[k]$ are,

$$\begin{cases} y[k] = x_2[k], \\ y[k+1] = -0.1\, x_2[k] + 0.01\, x_1[k]x_2[k], \\ y[k+2] = 0.01\, x_2[k] - 0.004\, x_1[k]x_2[k] + 3 \times 10^{-4} x_1^2[k]x_2[k] \\ \qquad + 1 \times 10^{-5} x_1[k]x_2^2[k] - 1 \times 10^{-6} x_1^2[k]x_2^2[k]. \end{cases} \tag{6}$$

We know from (6) that $x_2[k] = y[k]$ and subsequently $x_1[k] = 100y[k+1]/y[k] + 10$. Substitute the state variables as functions of the model outputs $y[k]$,$y[k+1]$ in $y[k+2]$ and clear out the denominators. The states are now eliminated and we have recovered the difference equation $\phi_{\text{PP}} = 0$ where ϕ_{PP} is given by (4). △

While in Example 1 elimination by hand is transparent in Σ_{PP}, it becomes impractical as the state dimension n_f grows, the required window length n_y increases, the polynomial degree rises, or the dynamics exhibit dense cross-term coupling. One option is to iteratively compute pairwise *resultants* in (6), and eliminate one component of the state in each iteration. Resultant based eliminants may contain extraneous factors, requiring post-processing to isolate the factor(s) that vanish on all output sequences generated by the model (Dong et al. (2023)). A second route to compute the eliminant of minimal degree is by computing the Gröbner basis of the polynomial ideal generated by (6). Gröbner basis computation for elimination proceed by repeated normal form computations, where intermediate polynomials are successively reduced via multivariate polynomial division with respect to a monomial order that prioritizes the variables to be eliminated (Cox et al. (2010)). While this is naturally formulated in exact arithmetic, the resulting computations are sensitive to perturbations. These limitations motivate an alternative methodology for state elimination that can benefit from numerically stable linear algebraic computations.

Contributions Since output monomials can be written as polynomials in $\boldsymbol{x}[k]$ (*cf.* (5)), the existence of a difference equation (3) implies a nontrivial linear dependence among these polynomials. We track these linear relations via the novel "model matrix" and show that the problem of state elimination in models of class Σ boils down to computing the left null space of this model matrix (Section 2, Definition 1). The left null space of the model matrix captures all vector representations of the output difference polynomials in the difference equation representation of the system's observable dynamics (Section 3, Theorem 1). When Σ is strongly locally observable (Nijmeijer (1982)), we show that the left null space of the model matrix is spanned by the rows of the Macaulay matrix (Batselier et al. (2014)) corresponding to a unique, irreducible output difference polynomial (Section 3.1, Theorem 5). We establish an orthogonal complementarity between the image representation (via the model matrix) and the kernel representation (via the Macaulay matrix) of the dynamical model (Section 3.2, Lemma 3). We conclude Section 3 with Algorithm 1, which deploys the singular value decomposition (SVD) to compute a numerical basis of the left null space of the model matrix, ensuring stable computations with floating-point coefficients for state elimination. The notation used in this paper can be consulted in Table 1.

Historical perspective The derivation of input-output (I/O) representations from state-space models through state elimination is a fundamental problem in system analysis and control theory. For continuous-time systems described by differential equations, differential algebra has emerged as a key tool for computing I/O relations (Glad (1989)). This approach has proven instrumental for applications such as parameter estimation (Denis-Vidal et al. (2003)) and identifiability analysis (Dong et al. (2023); Ljung and Glad (1994)). By contrast, nonlinear discrete-time models have received comparatively less attention for state elimination. An early algebraic geometry analysis of polynomial discrete-time systems Σ is presented in (Sontag (1979)). Constructive approaches to state elimination are presented in Forsman (1993a,b) and Ovchinnikov et al. (2020), relying on Gröbner basis computations. Further developments include the adaptation of the differential algebraic Ritt algorithm (Ljung and Glad (1994)) to the discrete-time case (Lyzell et al. (2011)). While differential algebra and algebraic geometry provide powerful tools for analyzing polynomial dynamical systems, their reliance on symbolic algorithms can make them challenging to apply to problems involving real data where measurement errors are inevitable. In the spirit of Forsman (1993b) "*in accordance with the golden rule of science known as Occam's razor, one should never use more machinery than necessary*" this work revisits state elimination in discrete-time, autonomous, time-invariant, single-output polynomial state-space models through a linear algebraic framework.

2. MODEL MATRIX

To tackle the problem of state elimination in Σ using a linear algebraic framework, we will consider the vector representation of multivariate polynomials. To represent and manipulate polynomial vectors effectively, we fix a monomial order that determines the arrangement of monomials within the coefficient vector. We use the graded negative lexicographic order (*grnlex*) to order the monomials (Batselier et al. (2014)). However, the remainder of the paper is valid for any graded multivariate monomial ordering. The monomials in the components of $\boldsymbol{y}[k : k+2]$ are arranged in the *grnlex* order as,

$$1 < y[k] < y[k+1] < y[k+2] < y[k]^2 < y[k]y[k+1] < y[k]y[k+2] < y[k+1]^2 < \ldots.$$

We can now represent polynomials using coefficient row vectors whose components are ordered according to the *grnlex* monomial order ascending from left to right.

Example 1. (continued). The vector representation of $y[k+1] = -0.1x_2[k] + 0.01x_1[k]x_2[k]$ (6) in $\mathcal{C}_2^2$ is,

Table 1. Notation summary.

Notation	Description
a, $\boldsymbol{a}$, a_i, $\boldsymbol{A}$,A_{ij}	Scalar (lowercase), tuple/vector (bold lowercase), ith component of $\boldsymbol{a}$, matrix (bold uppercase), ijth component of $\boldsymbol{A}$.
$\dim(\cdot)$	Dimension of the vector space.
$\mathcal{L}_{\text{null}}(\cdot)$	Left null space.
$\text{row}(\cdot)$, $\text{col}(\cdot)$	Row space, column space.
$\mathcal{K}[\boldsymbol{a}]$	Polynomial ring over field $\mathcal{K}$ (e.g., $\mathbb{R}$ or $\mathbb{C}$) in the components of $\boldsymbol{a}$.
$\boldsymbol{a}^{\boldsymbol{\omega}}$	Monomial $\prod_{i=0}^{n} a_i^{\omega_i}$, $\boldsymbol{a} \in \mathbb{C}^{n+1}$, $\boldsymbol{\omega} \in \mathbb{N}^{n+1}$. The degree is $\deg(\boldsymbol{a}^{\boldsymbol{\omega}}) = \|\boldsymbol{\omega}\|_1$.
$\boldsymbol{a}^d$	Column vector with components that are monomials in $\boldsymbol{a}$ up to degree d.
$\mathcal{C}_r^d$	Space of r-variate polynomials of degree $\leq d$
$\text{LM}(\cdot)$	Leading monomial
n_y	The order bound.
$\boldsymbol{y}[k:k+n_y]$	Window of model outputs, $\boldsymbol{y}[k:k+n_y] = \begin{bmatrix} y[k] & \cdots & y[k+n_y] \end{bmatrix}^{\text{T}} \in \mathbb{R}^{n_y+1}$.
n_f	The number of state variables in Σ.
$\boldsymbol{\mathcal{M}}_{n_y}^d$	Model matrix constructed using $g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ where $i \in \{0,\ldots,n_y\}$ and their monomial functions up to degree d.
m	Max. degree of the state monomials in $g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ where $i \in \{0,\ldots,n_y\}$ used to construct $\boldsymbol{\mathcal{M}}_{n_y}^d$.
$\boldsymbol{W}_{n_y}^d$	Basis matrix for the left null space of $\boldsymbol{\mathcal{M}}_{n_y}^d$.
d_{suff}	Degree at which the number of rows (p) of $\boldsymbol{\mathcal{M}}_{n_y}^d$ exceeds the number of columns (q).
d_{min}	Smallest degree for which $\dim(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)) > 0$.
ϕ	Output difference polynomial.
$\text{ord}(\phi)$, $\deg(\phi)$	Max. difference between the shifts of the outputs in ϕ, max. degree of the output monomials in ϕ.
$\boldsymbol{M}_\phi^d$	Macaulay matrix of degree d generated by ϕ
σ_i,τ_{std},τ_{gap}, Δ_{gap}	ith singular value, standard tolerance, gap based tolerance, spectral gap (see Remark 3)

$$\begin{matrix} 1 & x_1[k] & x_2[k] & x_1[k]^2 & x_1[k]x_2[k] & x_2[k]^2 \end{matrix}$$
$$\begin{bmatrix} 0 & 0 & -0.10 & 0 & 0.01 & 0 \end{bmatrix}.$$

△

In a polynomial the highest ranked monomial according to the chosen monomial ordering with a non-zero coefficient is the leading monomial (LM). The $\text{LM}(y[k+1]) = x_1[k]x_2[k]$.

Remark 1. Depending on the context, we will use the same non-bold lower case letter for both the polynomial and its vector representation. In (6) we refer to the polynomial $y[k+1] \in \mathbb{R}[x_1[k], x_2[k]]$ whereas in $y[k+1] \in \mathcal{C}_2^2$ we refer to its vector representation (*cf.* Example 1).

Definition 1. (*model matrix*). Given an order bound n_y, consider the set of n_y+1 polynomials $g(\boldsymbol{x}[k]), \ldots, g(\boldsymbol{f}^{n_y}(\boldsymbol{x}[k])) \in \mathbb{R}[\boldsymbol{x}[k]]$ which evaluate the model outputs $\boldsymbol{y}[k:k+n_y] \in \mathbb{R}^{n_y+1}$ at some state $\boldsymbol{x}[k] \in \mathbb{R}^{n_f}$ for a model which belongs to Σ. The *model matrix* of degree $d \in \mathbb{N}$ denoted by $\boldsymbol{\mathcal{M}}_{n_y}^d \in \mathbb{R}^{p\times q}$, is defined as the matrix whose rows are the vector representations of the polynomial functions $g(\boldsymbol{x}[k])^{\alpha_0} \cdots g(\boldsymbol{f}^{n_y}(\boldsymbol{x}[k]))^{\alpha_{n_y}}$, that evaluate the output monomials $\boldsymbol{y}[k:k+n_y]^{\boldsymbol{\alpha}} = y[k]^{\alpha_0} \cdots y[k+n_y]^{\alpha_{n_y}} = g(\boldsymbol{x}[k])^{\alpha_0} \cdots g(\boldsymbol{f}^{n_y}(\boldsymbol{x}[k]))^{\alpha_{n_y}}$, for $\boldsymbol{\alpha} \in \mathcal{A}$, with $\mathcal{A} = \left\{\boldsymbol{\alpha} \in \mathbb{N}^{n_y+1} : \|\boldsymbol{\alpha}\|_1 \leq d\right\}$. In compact form, the model matrix is expressed as

	1	$x_1[k]$	$x_2[k]$	$x_1[k]^2$	$x_1[k]x_2[k]$	$x_2[k]^2$	$x_1[k]^3$	$x_1[k]^2x_2[k]$	$x_1[k]x_2[k]^2$	$\cdots$
1	1	0	0	0	0	0	0	0	0	$\cdots$
$y[k]$	0	0	1	0	0	0	0	0	0	$\cdots$
$y[k+1]$	0	0	−0.1	0	0.01	0	0	0	0	$\cdots$
$y[k+2]$	0	0	0.01	0	−0.004	0	0	0.0003	1×10^{-5}	$\cdots$
$y[k]^2$	0	0	0	0	0	1	0	0	0	$\cdots$
$y[k]y[k+1]$	0	0	0	0	0	−0.1	0	0.01	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Fig. 1. A part of the model matrix $\boldsymbol{\mathcal{M}}_2^2$. Each row corresponds to an output monomial, expressed as its polynomial function in the components of $\boldsymbol{x}[k]$. The vertical and horizontal lines partition monomials by degree.

$$\boldsymbol{\mathcal{M}}_{n_y}^d = \begin{bmatrix} 1 & g(\boldsymbol{x}[k])^{\text{T}} & \cdots & g(\boldsymbol{f}^n(\boldsymbol{x}[k]))^{\text{T}} & \cdots & \left(g(\boldsymbol{x}[k]^d)\right)^{\text{T}} & \cdots & \left(g(\boldsymbol{f}^{n_y}(\boldsymbol{x}[k]))^d\right)^{\text{T}} \end{bmatrix}^{\text{T}}.$$

The row space of the model matrix, $\text{row}(\boldsymbol{\mathcal{M}}_{n_y}^d) \in \mathcal{C}_{n_f}^{md}$, where $m = \max_{i\in\{0,\ldots,n_y\}} \left(\deg\left(g(\boldsymbol{f}^i(\boldsymbol{x}[k]))\right)\right)$ satisfies the relation

$$\boldsymbol{y}[k:k+n_y]^d = \boldsymbol{\mathcal{M}}_{n_y}^d \boldsymbol{x}[k]^{md}. \tag{7}$$

Example 1. (continued). A part of the model matrix $\boldsymbol{\mathcal{M}}_2^2$ corresponding to the predator-prey model (2) is shown in Fig. 1. $\boldsymbol{\mathcal{M}}_2^2$ is further continued in Fig. 2 which shows the spy plot of $\boldsymbol{\mathcal{M}}_2^4$, with $\boldsymbol{\mathcal{M}}_2^1$ to $\boldsymbol{\mathcal{M}}_2^3$ outlined within. △

Lemma 1. The size of the model matrix $\boldsymbol{\mathcal{M}}_{n_y}^d \in \mathbb{R}^{p\times q}$ with respect to the order bound n_y and degree d, corresponding to a model which belongs to Σ, is:

$$p = \binom{d+n_y+1}{n_y+1} \text{ and } q = \binom{dm+n_f}{n_f}.$$

There exists a sufficiently large $d = d_{\text{suff}}$, s.t. $\boldsymbol{\mathcal{M}}_{n_y}^d$ has more rows, p, than columns, q, $\forall\, d \geq d_{\text{suff}}$, *iff* $n_y \geq n_f$.

Proof. Using $\binom{d+r}{r} \sim d^r/r!$ as $d \to \infty$, we get $p \sim d^{n_y+1}/(n_y+1)!$ and $q \sim (md)^{n_f}/n_f!$, so $p/q \sim (n_f! m^{-n_f} d^{n_y+1-n_f})/(n_y+1)!$. Hence, there exists a sufficiently large $d = d_{\text{suff}}$, such that $p > q\ \forall\, d \geq d_{\text{suff}}$ when $n_y \geq n_f$. If $n_y + 1 = n_f$ then $p/q \to m^{-n_f}$ (so $p = q$ for $m = 1$ and $p < q$ for $m > 1$) and if $n_y < n_f$ then $p/q \to 0$ i.e., $p < q\, \forall\, d$. □

3. STATE ELIMINATION

Lemma 1 characterizes the size of the model matrix $\boldsymbol{\mathcal{M}}_{n_y}^d$ based on its degree d and order bound n_y. The size of the model matrix, influences the structure of its fundamental subspaces. We are interested in the left null space of the model matrix denoted by $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)$. The columns of the basis matrix $\boldsymbol{W}_{n_y}^d \in \mathbb{R}^{p\times\ell}$ span $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)$, where the dimension of the left null space is denoted as $\dim\left(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)\right) = \ell$. The columns of $\boldsymbol{W}_{n_y}^d$ provide the vector representation of the output difference polynomial (Theorem 1). Theorem 5 demonstrates that for strongly locally observable systems, the left null space is entirely defined by the unique, irreducible output difference polynomial (of order at most n_f) (Section 3.1). This allows us to establish an orthogonal complementarity between the kernel and image representations of the dynamical model (Lemma 3). We formalize the aforementioned procedure for state elimination in Algorithm 1 (Section 3.2).

Example 1. (continued). For the model matrix corresponding to the predator-prey model (2), $\dim\big(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_2^d)\big) = 0$ for $d \leq 2$. However, for $d = 3$ $\dim\big(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_2^3) = 1$. There exists $\phi_{\text{PP}}^{\text{T}} \in \mathcal{L}_{\text{null}}\big(\boldsymbol{\mathcal{M}}_2^3\big)$ such that upon pre multiplication with (7) we have the difference equation, $\phi_{\text{PP}}(\boldsymbol{y}[k : k+2]) = \phi_{\text{PP}}\boldsymbol{y}[k : k+2]^3 = \big(\phi_{\text{PP}}\boldsymbol{\mathcal{M}}_2^3\big)\boldsymbol{x}[k]^{12} = 0$. The expression $\phi_{\text{PP}}\boldsymbol{y}[k : k+2]^3$ is the output difference polynomial in $\mathbb{R}[\boldsymbol{y}[k : k+2]]$ and $\phi_{\text{PP}} \in \mathcal{C}_3^3$ is its vector representation given by

$$\begin{bmatrix} \overset{\dots}{\mathbf{0}} & \overset{y[k]y[k+1]}{-20} & \overset{y[k]y[k+2]}{100} & \overset{y[k+1]^2}{-300} & \overset{\dots}{\mathbf{0}} & \overset{y[k]y[k+1]^2}{0.10} & \overset{y[k]y[k+2]^2}{0} & \overset{y[k]y[k+1]^2}{1} & \overset{\dots}{\mathbf{0}} \end{bmatrix}.$$

Observe that, ϕ_{PP} is *irreducible*, it cannot be factored as $\phi_{\text{PP}} = rs$ with nonconstant $r, s \in \mathbb{R}[\boldsymbol{y}[k{:}k{+}2]]$. We focus on irreducible polynomials as they ensure a minimal–degree difference polynomial. △

We now present Theorem 1, which gives a sufficient condition for the existence of difference equations for models in Σ, followed by Lemma 2, which provides conditions for irreducibility.

Theorem 1. (*Existence*). Given a discrete-time, autonomous, time-invariant, single-output polynomial state-space model Σ, with n_f components in the state variable, there exists a difference equation $\phi(\boldsymbol{y}[k : k+n_y]) = 0$, where the output difference polynomial $\phi \in \mathbb{R}[\boldsymbol{y}[k : k+n_y]]$ in the model outputs is of order at most n_y, where $n_y \geq n_f$.

Proof. Given a state-space model of the form Σ with n_f components in the state variable. We know from Lemma 1, when the order bound $n_y \geq n_f$ there exists a degree d_{suff} such that for all $d \geq d_{\text{suff}}$ the model matrix $\boldsymbol{\mathcal{M}}_{n_y}^d \in \mathbb{R}^{p\times q}$ has more rows than columns, confirming the existence of a non-trivial left null space. The basis matrix of $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)$ is $\boldsymbol{W}_{n_y}^d$, which upon pre-multiplication with (7) results in,

$$\big(\boldsymbol{W}_{n_y}^d\big)^{\text{T}}\boldsymbol{y}[k : k+n_y]^d = \Big(\big(\boldsymbol{W}_{n_y}^d\big)^{\text{T}}\boldsymbol{\mathcal{M}}_{n_y}^d\Big)\,\boldsymbol{x}[k]^{md} = \mathbf{0}.$$

The column vectors of the basis matrix $\boldsymbol{W}_{n_y}^d = \begin{bmatrix}\phi_1^{\text{T}} \cdots \phi_\ell^{\text{T}}\end{bmatrix} \in \mathbb{R}^{p\times\ell}$ are the vector representations $\phi_i \in \mathcal{C}_{n_y+1}^d$ of the output difference polynomials $\phi_i \in \mathbb{R}[\boldsymbol{y}[k : k+n_y]]$. Additionally, observe that $\phi_i\boldsymbol{y}[k : k+n_y]^d = 0$. Hence, there exists a difference equation $\phi_i(\boldsymbol{y}[k : k+n_y]) = 0$ where $\phi_i \in \mathbb{R}[\boldsymbol{y}[k : k+n_y]]$ and is of order at most n_y. □

Remark 2. We denote d_{suff} as the smallest degree for which $\boldsymbol{\mathcal{M}}_{n_y}^d$ has more rows than columns, ensuring a non-trivial left null space (Theorem 1). However, a non-trivial left null space may exist at a lower degree $d_{\min} < d_{\text{suff}}$, depending on the coefficients of $g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ for $i \in \{0, \dots, n_y\}$.

Lemma 2. (*Irreducible output difference polynomials*). In the setting of Thm. 1, suppose for some $d_{\min}$ $\dim\big(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^{d_{\min}})\big) = \ell \geq 1$, and $\dim\big(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)\big) = 0$ for $d < d_{\min}$. Each ϕ_i in $\boldsymbol{W}_{n_y}^{d_{\min}} = \begin{bmatrix}\phi_1^{\text{T}} \cdots \phi_\ell^{\text{T}}\end{bmatrix}$, satisfies the difference equation $\phi_i(\boldsymbol{y}[k : k+n_y]) = 0$. Then, each $\phi_i \in \mathbb{R}[\boldsymbol{y}[k : k+n_y]]$ is irreducible.

Proof. Suppose ϕ_i is reducible, i.e. $\phi_i = rs$ with $r, s \in \mathbb{R}[\boldsymbol{y}[k : k+n_y]]$. Then at least one of the factors must vanish, either $r = 0$ or $s = 0$. Since the degrees of r, s are necessarily smaller than that of ϕ_i, either r or s should already be recovered in a basis matrix of $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)$ for some $d < d_{\min}$. However, this contradicts our assumption that $\dim\big(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}_{n_y}^d)\big) = 0$ for $d < d_{\min}$. Thus, each ϕ_i is irreducible. □

Example 1. (continued). For the model matrix corresponding to the predator-prey model it is only at $d_{\text{suff}} = 43$, where $p > q$ i.e., $\text{size}\big(\boldsymbol{\mathcal{M}}_2^{43}\big) = 15180 \times 15051$, whereas a non-trivial left null space exists for $d_{\min} = 3$ where $\text{size}\big(\boldsymbol{\mathcal{M}}_2^3\big) = 20 \times 91$. △

3.1 Strongly locally observable models

Strongly locally observable models in the class Σ satisfy the nonlinear observability rank condition (Theorem 3). This ensures that the rows corresponding to the output monomials of $\boldsymbol{y}[k : k+n_f-1] = [y[k] \cdots y[k+n_f-1]]^{\text{T}} \in \mathbb{R}^{n_f}$ in $\boldsymbol{\mathcal{M}}_{n_f}^d$ are mutually linearly independent (Corollary 1), resulting in the difference equation $\phi(\boldsymbol{y}[k : k+n_f]) = 0$ where $\phi \in \mathbb{R}[\boldsymbol{y}[k : k+n_f]]$ is the unique, irreducible output difference polynomial of order at most n_f (Theorem 5).

Definition 2. (*Strongly locally observable models (SLO)*, Nijmeijer (1982)). Consider states $\boldsymbol{x}[k], \bar{\boldsymbol{x}}[k] \in \mathbb{R}^{n_f}$. A model Σ is strongly locally observable at $\boldsymbol{x}[k] \in \mathbb{R}^{n_f}$ if there exists a neighborhood $\mathcal{U} \subseteq \mathbb{R}^{n_f}$ of $\boldsymbol{x}[k]$ so that for any $\bar{\boldsymbol{x}}[k] \in \mathcal{U}$, $g(\boldsymbol{f}^i(\bar{\boldsymbol{x}}[k])) = g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ for $i \in \{0, \dots, n_f-1\}$, implies $\bar{\boldsymbol{x}}[k] = \boldsymbol{x}[k]$.

We now restate three known theorems without proof; the interested reader may consult the references for their proofs.

Theorem 2. (*Nonlinear observability rank condition (NORC)*, Nijmeijer (1982)). Given the model Σ, consider the vector function $\boldsymbol{b}$ whose components are the functions in $g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ with $i \in \{0, \dots, n_f-1\}$. Let $\text{Jac}_{\boldsymbol{x}[k]}(\boldsymbol{b})$ denote the Jacobian of $\boldsymbol{b}$ with respect to $\boldsymbol{x}[k]$. If $\text{rank}\left(\text{Jac}_{\boldsymbol{x}[k]}(\boldsymbol{b})\right) = n_f$ for some $\boldsymbol{x}[k] \in \mathbb{R}^{n_f}$, then the model is SLO at $\boldsymbol{x}[k]$.

Theorem 3. (*Converse of the NORC* Nijmeijer (1982)). If the model is SLO then the NORC is satisfied on an open and dense subset of the state space $\mathcal{X} \subseteq \mathbb{R}^{n_f}$.

Example 1. (continued). For the predator-prey model in (2), $\boldsymbol{b} = [x_2[k]\ \ 0.1x_2[k] + 0.01x_1[k]x_2[k]]^{\text{T}}$,

$$\text{Jac}_{\boldsymbol{x}[k]}(\boldsymbol{b}) = \begin{bmatrix} 0 & 1 \\ 0.01x_2[k] & 0.1 + 0.01x_1[k] \end{bmatrix},$$

and $\det\left(\text{Jac}_{\boldsymbol{x}[k]}(\boldsymbol{b})\right) = -0.01x_2[k]$. The $\text{rank}(\text{Jac}_{\boldsymbol{x}[k]}(\boldsymbol{b})) = 2$, $\forall\, \boldsymbol{x}[k] \in \mathcal{X} = \{[x_1[k]\ \ x_2[k]]^{\text{T}} \in \mathbb{R}^2 : x_2[k] \neq 0\}$, i.e., the model (2) is SLO $\forall\, \boldsymbol{x}[k] \in \mathcal{X}$.

Theorem 4. (*Jacobian criterion*). (Beecken et al., 2013, Theorem 8) Let $\boldsymbol{r} = [r_1 \cdots r_n]^{\text{T}}$ where $r_i \in \mathbb{R}[\boldsymbol{x}]$, where $\boldsymbol{x} = [x_1 \cdots x_n]^{\text{T}} \in \mathbb{R}^n$. Then the polynomials r_i are algebraically independent, i.e., there exists no non-trivial $\phi \in \mathbb{R}[\boldsymbol{x}]$ s.t., $\phi(r_1, \dots, r_n) \equiv 0$, *iff* $\text{rank}\left(\text{Jac}_{\boldsymbol{x}}(\boldsymbol{r})\right) = n$.

Corollary 1. If Σ is SLO, then $\boldsymbol{\mathcal{M}}_{n_f-1}^d$ has full row rank, i.e., $\text{rank}\big(\boldsymbol{\mathcal{M}}_{n_f-1}^d\big) = p\ \forall\, d \in \mathbb{N}$.

Proof. We know from Theorem 3, that if Σ is SLO, then the polynomials $g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ with $i \in \{0, \dots, n_f-1\}$ satisfy the NORC on an open and dense subset of the state space.

According to Theorem 4, there exists no non-trivial $\phi \in \mathbb{R}[\boldsymbol{y}[k:k+n_f-1]]$ s.t. $\phi\left(g(\boldsymbol{x}[k]),\ldots,g\left(\boldsymbol{f}^{n_f-1}(\boldsymbol{x}[k])\right)\right)\equiv 0$, which implies that there exists no linear relation between the rows of $\boldsymbol{\mathcal{M}}^d_{n_f-1}\ \forall\, d$ and $\operatorname{rank}\big(\boldsymbol{\mathcal{M}}^d_{n_f-1}\big)=p\ \forall\, d\in\mathbb{N}$. □

Example 1. (continued). The mutually linearly independent rows of $\boldsymbol{\mathcal{M}}^4_1$ corresponding to the output monomials in $\boldsymbol{y}[k:k+1]$ of the predator-prey model (2) are covered with a pink band over the model matrix $\boldsymbol{\mathcal{M}}^4_2$ in Fig. 2. △

Theorem 5. (*Uniqueness*). Given a discrete-time, autonomous, time-invariant, single-output polynomial state-space model Σ, with n_f components in the state variable and which is strongly locally observable. There exists a difference equation $\phi(\boldsymbol{y}[k:k+n_f])=0$ where the irreducible output difference polynomial $\phi\in\mathbb{R}[\boldsymbol{y}[k:k+n_f]]$ in the model outputs is unique (up to a nonzero scalar multiple) and of order at most n_f.

Proof. We know from Theorem 1, that there exists a difference equation $\phi=0$ s.t. $\phi\in\mathbb{R}[\boldsymbol{y}[k:k+n_f]]$, which implies that the numerical basis $\boldsymbol{W}^d_{n_f}\neq\boldsymbol{0}\ \forall\ d\geq d_{\text{suff}}\in\mathbb{N}$. Additionally, we know from Corollary 1 that there exists no linear combination between rows corresponding to the output monomials of the polynomials in $\boldsymbol{y}[k:k+n_f-1]^d\,\forall\, d\in\mathbb{N}$, therefore any $\phi_i\in\big(\boldsymbol{W}^d_{n_f}\big)^{\mathrm{T}}$ must contain nonzero coefficients in at least one monomial with $y[k+n_f]$. Assume that for some $d_{\min}$, $\dim\big(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d_{\min}}_{n_f})\big)=\ell=2$ (the following approach can be generalized for $\ell>2$) and $\dim\big(\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d}_{n_f})\big)=0\ \forall\, d<d_{\min}$ so that the basis matrix of $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d_{\min}}_{n_f})$, $\boldsymbol{W}^d_{n_f}=\begin{bmatrix}\phi_r^{\mathrm{T}}\ \phi_s^{\mathrm{T}}\end{bmatrix}\in\mathcal{C}^{d_{\min}}_{n_f+1}$ where ϕ_r and ϕ_s are two distinct irreducible output difference polynomials each satisfying the equation $\phi_r(\boldsymbol{y}[k:k+n_f])=0$ and $\phi_s(\boldsymbol{y}[k:k+n_f])=0$. We can split ϕ_r,ϕ_s as $\phi_r=ry[k+n_f]+h_r=0$ and $\phi_s=sy[k+n_f]+h_s=0$, where $r,s\in\mathbb{R}[\boldsymbol{y}[k:k+n_f]]$, $h_r,h_s\in\mathbb{R}[\boldsymbol{y}[k:k+n_f-1]]$ so that $h_r,h_s\not\equiv 0$ (Corollary 1), thus $r,s\not\equiv 0$. Now multiply, ϕ_r and ϕ_s with s and r, such that $s\phi_r=(rs)\,y[k+n_f]+sh_r=0$ and $r\phi_s=(rs)\,y[k+n_f]+rh_s=0$. Subtracting these equations, we obtain $sh_r=rh_s$ and since $r\not\equiv 0$, we conclude $h_s=(s/r)\,h_r$ Substituting this into ϕ_s we have,

$$\phi_s=s/r\,(ry[k+n_f]+h_r)=(s/r)\,\phi_r. \tag{8}$$

If s/r is rational, then the right-hand side is not a polynomial since ϕ_r is irreducible, ϕ_s cannot be in $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d_{\min}}_{n_f})$. If s/r is polynomial, then $\deg(\phi_s)>d_{\min}$, again implying $\phi_s\notin\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d_{\min}}_{n_f})$. Thus, we must have $s=r$ and consequently $\phi_s=\phi_r$, contradicting the assumption that there are two distinct irreducible output difference polynomials of order at most n_f and degree $d_{\min}$. We can repeat the procedure for some $\phi_s\in\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d}_{n_f})$ s.t. $\deg(\phi_s)=d>d_{\min}$. Assume that ϕ_s is irreducible. We can rewrite ϕ_s as in (8). The factor s/r can only admit a polynomial of degree $d-d_{\min}$ s.t. ϕ_s forms an element of $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d}_{n_f})$, contradicting our assumption that ϕ_s is irreducible. Notice, that ϕ_s is a polynomial multiple of the irreducible output difference polynomial ϕ_r which is unique up to a nonzero scalar multiple which defines the same difference equation, since $c\phi_r(\boldsymbol{y}[k:k+n_f])=0$ *iff* $\phi_r(\boldsymbol{y}[k:k+n_f])=0$ for any $c\in\mathbb{R}\setminus\{0\}$. □

Example 1. (continued). The difference equation corresponding to the SLO predator–prey model (2), $\phi_{\text{PP}}(\boldsymbol{y}[k:k+2])=0$, is described by an irreducible output difference polynomial ϕ_{PP} in (4), which is unique up to multiplication by a nonzero scalar. We fix a unique representative by normalizing ϕ_{PP} to be *monic* with respect to the *grnlex* order; i.e., we scale ϕ_{PP} so that the coefficient of its leading monomial equals 1. Under *grnlex*, $\text{LM}(\phi_{\text{PP}})=y[k]y[k+1]^2$. We follow this convention throughout the paper. △

3.2 Orthogonal complementarity

We will now establish the orthogonal complementarity between the image representation (via the model matrix) and the kernel representation (via the Macaulay matrix) for SLO models in the class Σ.

Definition 3. (*Macaulay matrix*, Batselier et al. (2014)). Given a set of polynomials $r_1,\ldots,r_s$ where $r_i\in\mathbb{C}[x_1,\ldots,x_n]$, of degree d_i resepectively, the Macaulay matrix of degree d, denoted by $\boldsymbol{M}^d$ contains in its rows the vector representations of

$$\boldsymbol{M}^d=\begin{bmatrix}r_1^{\mathrm{T}}\ (x_1r_1)^{\mathrm{T}}\ \cdots\ (x_n^{d-d_1}r_1)^{\mathrm{T}}\ \cdots\ r_s^{\mathrm{T}}\ (x_1f_s)^{\mathrm{T}}\ \cdots\ (x_n^{d-d_s}r_s)^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}},$$

where each polynomial r_i is multiplied with all monomials from degree 0 to degree $d-d_i$.

Example 1. (continued). We will denote the Macaulay matrix of degree 4 generated using ϕ_{PP} (4) by $\boldsymbol{M}^4_\phi$, see top of Fig. 2. The sub matrix $\boldsymbol{M}^3_\phi$ is outlined within. △

Lemma 3. (*On the orthogonal complementarity*) In the setting of Theorem 5, the rows of a Macaulay matrix $\boldsymbol{M}^d_\phi\in\mathbb{R}^{\ell\times p}$ generated by the unique irreducible output difference polynomial ϕ, span the left null space of the model matrix $\boldsymbol{\mathcal{M}}^d_{n_f}\in\mathbb{R}^{p\times q}$. Then, for all $d\geq d_{\min}$ $\operatorname{row}\big(\boldsymbol{M}^d_\phi\big)\perp\operatorname{col}\big(\boldsymbol{\mathcal{M}}^d_{n_f}\big)$ and $\dim\big(\operatorname{row}\big(\boldsymbol{M}^d_\phi\big)\big)+\dim\big(\operatorname{col}\big(\boldsymbol{\mathcal{M}}^d_{n_f}\big)\big)=p$.

Proof. According to Theorem 5, the unique irreducible output difference polynomial ϕ spans the one-dimensional left null space $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d_{\min}}_{n_f})$. For $d>d_{\min}$ every element of $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d}_{n_f})$ is a polynomial multiple of ϕ, i.e., $r\phi$ with $r\in\mathbb{R}[\boldsymbol{y}[k:k+n_f]]$ of total degree at most $d-d_{\min}$. Writing $r=\sum_{\boldsymbol{\omega}\in\Omega}c_{\boldsymbol{\omega}}\boldsymbol{y}[k:k+n_f]^{\boldsymbol{\omega}}$ with $\Omega=\{\boldsymbol{\omega}\in\mathbb{N}^{n_f+1}:\|\boldsymbol{\omega}\|_1\leq d-d_{\min}\text{ and }c_{\boldsymbol{\omega}}\in\mathbb{R}\setminus\{0\}\}$, we have $r\phi=\sum_{\boldsymbol{\omega}\in\Omega}c_{\boldsymbol{\omega}}(\boldsymbol{y}[k:k+n_f]^{\boldsymbol{\omega}}\phi)$. The vector representations of $\boldsymbol{y}[k:k+n_f]^{\boldsymbol{\omega}}\phi$ have distinct leading monomials, so they are linearly independent and form the column vectors of the basis matrix $\boldsymbol{W}^d_{n_f}$ that span $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d}_{n_f})$. The rows of $\big(\boldsymbol{W}^d_{n_f}\big)^{\mathrm{T}}$ correspond exactly to the rows of the Macaulay matrix $\boldsymbol{M}^d_\phi$ (Definition 3) generated by the unique, irreducible output difference polynomial ϕ of the SLO state-space model in Σ (1). Therefore $\mathcal{L}_{\text{null}}(\boldsymbol{\mathcal{M}}^{d}_{n_f})=\operatorname{row}(\boldsymbol{M}^d_\phi)$, which implies $\operatorname{row}(\boldsymbol{M}^d_\phi)\perp\operatorname{col}(\boldsymbol{\mathcal{M}}^d_{n_f})$ and by the fundamental theorem of linear algebra $\dim\big(\operatorname{row}\big(\boldsymbol{M}^d_\phi\big)\big)+\dim\big(\operatorname{col}\big(\boldsymbol{\mathcal{M}}^d_{n_f}\big)\big)=p$. □

We conclude this section with Algorithm 1 for state elimination. When initialized with an SLO model in Σ and $n_y=n_f$, the algorithm returns the difference equation $\phi(\boldsymbol{y}[k:k+n_f])=0$, where $\phi\in\mathbb{R}[\boldsymbol{y}[k:k+n_y]]$, is the unique, irreducible output difference polynomial of $\deg(\phi)=d_{\min}$ and $\operatorname{ord}(\phi)\leq n_f$ (Theorem 5). For

$n_y > n_f$ and/or non-SLO models, Algorithm 1 may return multiple difference equations, described by irreducible output difference polynomials of minimal order(s) $\leq n_y$.

Algorithm 1 State elimination algorithm

Input: State-space model in Σ (with numerical coefficients), order bound $n_y \geq n_f$.

Output: Difference equation(s) $\phi(\boldsymbol{y}[k:k+n_y]) = 0$ where $\phi \in \mathbb{R}[\boldsymbol{y}[k:k+n_y]]$, that are irreducible, of $\text{ord}(\phi) \leq n_y$ and $\deg(\phi) = d_{\min}$.

1: Initialize $d \leftarrow 0$ and $\boldsymbol{W}^0_{n_y} \leftarrow 0$.
2: **while** $\boldsymbol{W}^d_{n_y} = \boldsymbol{0}$ **do**
3: $d \leftarrow d+1$ and construct the model matrix $\boldsymbol{\mathcal{M}}^d_{n_y}$.
4: Compute the SVD of $\boldsymbol{\mathcal{M}}^d_{n_y}$ and extract a numerical basis matrix of the left null space $\boldsymbol{W}^d_{n_y}$.
5: **end while**
6: $d_{\min} \leftarrow d$ and **return** $\boldsymbol{W}^{d_{\min}}_{n_y} = \begin{bmatrix}\phi_1^{\mathrm{T}} & \cdots & \phi_\ell^{\mathrm{T}}\end{bmatrix}$.

Remark 3. Consider the singular values of $\boldsymbol{\mathcal{M}}^d_{n_y} \in \mathbb{R}^{p\times q}$ as $\sigma_1 \geq \cdots \geq \sigma_{\min(p,q)}$. The numerical left null space is spanned by the left singular vectors associated with singular values below some tolerance $\tau \in \mathbb{R}$ i.e., those $\sigma_i \leq \tau$. As a baseline, we use the standard numerical rank tolerance $\tau_{\text{std}} = \epsilon\,\sigma_1 \max(p,q)$, here ϵ is the machine precision. However, in the presence of noise, the distinction between numerically negligible and informative singular values may not be clear, and the resulting numerical rank can be sensitive to the chosen tolerance. Therefore, when a pronounced spectral gap $\Delta_{\text{gap}} = \log_{10}\sigma_{i^\star} - \log_{10}\sigma_{i^\star+1}$ is observed, we instead use the tolerance $\tau_{\text{gap}} = \frac{1}{2}(\sigma_{i^\star}+\sigma_{i^\star+1})$, where $i^\star \in \arg\max_i\big(\log_{10}\sigma_i - \log_{10}\sigma_{i+1}\big)$ (Breiding et al. (2018)). For the numerical examples in the following section, the spectral gap based choice is more practical, and we therefore use τ_{gap} throughout the remainder of the paper.

Model extensions For notational simplicity and a didactical exposition, we restrict ourselves to the model class Σ. The framework extends, *mutatis mutandis*, to rational state-update and observation functions: Consider the common denominator across functions corresponding to the output monomials, Algorithm 1 seeks linear relations among the numerator polynomials via a modified model matrix to recover an output difference equation. The approach also extends to multiple outputs, either by treating each output as a separate instance of Σ, or by forming a joint model matrix whose left null space may include coupled relations between output streams; a systematic characterization of such coupled relations is left for future work. Finally, in models with inputs, the model matrix is input-dependent; however, for a known input sequence one can substitute the input values to obtain a numerical instance of the model matrix and proceed with Algorithm 1. Input-output difference equations valid for arbitrary inputs require modifications to the model matrix construction and are beyond the scope of this paper.

4. NUMERICAL EXAMPLES

Example 1. (continued on *spectral gap enhancement*). To illustrate how the relative magnitudes of the model coefficients can affect numerical rank decisions, we consider the predator-prey model and show that an equivalent model under a similarity transformation [2] can yield a more pronounced spectral gap in the model matrix. For Σ_{PP}, the separation between the last two singular values is given by $\Delta_{\text{gap}} = 6.6583$. We consider the state transformation $\tilde{\boldsymbol{x}}[k] = \boldsymbol{T}\boldsymbol{x}[k]$ where $\boldsymbol{T} \in \mathbb{R}^{n_f\times n_f}$ is an invertible transformation matrix which yields the equivalent model $\tilde{\boldsymbol{x}}[k+1] = \boldsymbol{T}\,\boldsymbol{f}(\boldsymbol{T}^{-1}\tilde{\boldsymbol{x}}[k]) = \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}[k]), y[k] = g\big(\boldsymbol{T}^{-1}\tilde{\boldsymbol{x}}[k]\big) = \tilde{g}(\tilde{\boldsymbol{x}}[k])$. Consider the transformation matrices parameterized as $\boldsymbol{T} = t\boldsymbol{I}_2$, where $t \in \{0.1, 0.01\}$ and $\boldsymbol{I}_n \in \mathbb{R}^{n\times n}$ is an identity matrix of size $n \in \mathbb{N}$, the transformed model is,

$$\Sigma_{\text{PPt}} : \begin{cases} \tilde{x}_1[k+1] = 3\tilde{x}_1[k] - (0.01/t)\tilde{x}_1[k]\tilde{x}_2[k], \\ \tilde{x}_2[k+1] = -0.1\,\tilde{x}_2[k] + (0.01/t)\tilde{x}_1[k]\tilde{x}_2[k], \\ y[k] = (1/t)\tilde{x}_2[k]. \end{cases} \quad (9)$$

For $t = 0.1$, the ratio between the largest and smallest coefficient in the state-update function of Σ_{PPt} is 30, compared with 300 in Σ_{PP}. Since $\tilde{\boldsymbol{f}}^i(\boldsymbol{x}[k])$ is formed by repeated composition, its coefficients accumulate as products of those in $\tilde{\boldsymbol{f}}(\boldsymbol{x}[k])$; reducing their relative magnitudes therefore mitigates coefficient growth/decay in $\tilde{g}(\tilde{\boldsymbol{f}}^i(\tilde{\boldsymbol{x}}[k]))$. Accordingly, for Σ_{PPt} the ratio between the largest and smallest coefficient in $\tilde{g}(\tilde{\boldsymbol{f}}^2(\tilde{\boldsymbol{x}}[k]))$ reduces to 10 (from 1×10^4 in (5)). Constructing $\boldsymbol{\mathcal{M}}^3_2$ from the transformed model yields a more pronounced spectral gap $\Delta_{\text{gap}} = 11.9666$, and we recover the same output difference polynomial ϕ_{PP} after normalization. For $t = 0.01$, the output polynomial functions exhibit larger relative coefficient magnitudes i.e., ratio on the order of 1×10^2 in $\tilde{g}(\tilde{\boldsymbol{f}}^2(\tilde{\boldsymbol{x}}[k]))$, where the spectral gap decreases to $\Delta_{\text{gap}} = 9.4714$. This motivates using equivalent models (via similarity transformations) that reduce coefficient disparities within the state update function. Developing systematic criteria for selecting such transformations is left for future work. △

Example 1. (continued on *model perturbations*). We next probe the robustness of Algorithm 1 under coefficient perturbations. Starting from the transformed predator–prey model Σ_{PPt} (9) with $t = 0.1$, we construct perturbed maps by injecting degree-one perturbations according to

$$\begin{bmatrix}\boldsymbol{f}_h(\tilde{\boldsymbol{x}}[k])\\ g_h(\tilde{\boldsymbol{x}}[k])\end{bmatrix} = \begin{bmatrix}\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}[k])\\ \tilde{g}(\tilde{\boldsymbol{x}}[k])\end{bmatrix} + 10^{-h}\,\boldsymbol{B}\begin{bmatrix}1\\ \tilde{x}_1[k]\\ \tilde{x}_2[k]\end{bmatrix}, \quad (10)$$

where $\boldsymbol{B} \in \mathbb{R}^{3\times3}$ is a random matrix whose entries are sampled from a uniform distribution on $(0,1)$. For each $h \in \{5,7,9\}$, we compute the irreducible output difference polynomial of minimal order, denoted by $\phi_{\text{PP}h}$, and compare it to the output difference polynomial ϕ_{PP} from the unperturbed model Σ_{PPt} using $\|\phi_{\text{PP}} - \phi_{\text{PP}h}\|$ in Table 2. The leading monomial of ϕ_{PPh} is the same as ϕ_{PP}, however additional lower-degree terms appear. The residual $\big|\phi_{\text{PP}h}(\boldsymbol{y}[0{:}2])\big|$ for all h, on the output sequence of Σ_{PP}, $\boldsymbol{y}[0{:}2] = [10\ 9\ 25.2]^{\mathrm{T}}$ generated from $\boldsymbol{x}[0] = [100\ 10]^{\mathrm{T}}$ is also shown in Table 2. Additionally, we report the spectral gap Δ_{gap} of the model matrix $\boldsymbol{\mathcal{M}}^3_2$ in Table 2. As the perturbation level increases, the spectral gap shrinks, which

[2] Polynomial state-space models are unique up to a similarity transformation, their output behavior is invariant under an invertible state transformation (Paduart et al. (2010)).

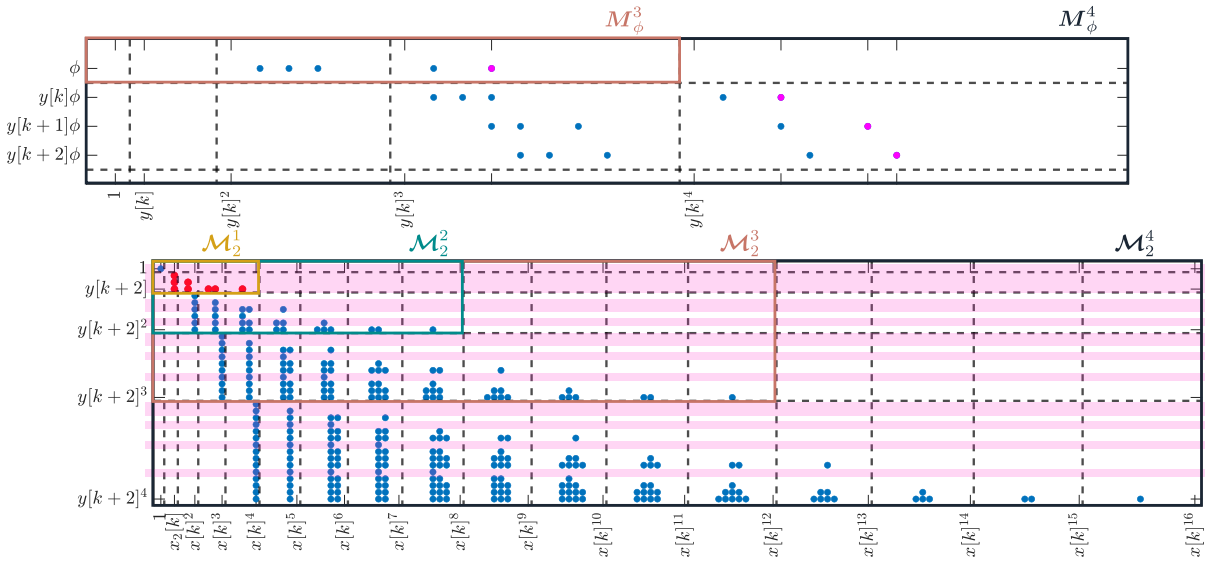


Fig. 2. **Bottom:** Model matrix $\mathcal{M}_2^4$ associated with the predator-prey model (2). Coefficients of polynomials $g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ for $i \in \{0,1,2\}$ are marked in (•). The linearly independent rows of $\mathcal{M}_1^4$ corresponding to the output monomials $\boldsymbol{y}[k:k+1]^4$ are covered with a pink band. Dashed vertical and horizontal lines partition monomials by degree. Colored outlines indicate the embedded sub matrices $\mathcal{M}_2^3$, $\mathcal{M}_2^2$ and $\mathcal{M}_2^1$. **Top:** Macaulay matrix M_ϕ^4. The first row represents the unique, irreducible output difference polynomial (4) ϕ_{PP}, which coincides with M_ϕ^3. Vertical dashed lines partition degree blocks; horizontal dashed lines group multiples by degree. Coefficients of $\text{LM}(\phi_{\text{PP}}) = y[k]y[k+1]^2$ and its monomial multiples are marked (•).

Table 2. Performance of Algorithm 1 in determining the unique, irreducible output difference polynomial corresponding to the perturbed predator-prey model (10). Reported are the deviation from the unperturbed difference polynomial $\|\phi_{\text{PP}} - \phi_{\text{PP}h}\|$, the residual $\left|\phi_{\text{PP}h}(\boldsymbol{y}[0{:}2])\right|$ computed on a model trajectory (Example 1) and the spectral gap Δ_{gap} corresponding to the model matrix $\mathcal{M}_2^3$.

h	$\|\phi_{\text{PP}} - \phi_{\text{PP}h}\|$	$\left\|\phi_{\text{PP}h}(\boldsymbol{y}[0{:}2])\right\|$	Δ_{gap}
9	2.1977×10^{-5}	1.3943×10^{-4}	8.6547
7	2.2×10^{-3}	1.3868×10^{-2}	6.6547
5	2.198×10^{-1}	1.3878	4.6546

complicates numerical rank determination. For instance, at $h = 5$ the standard threshold $\tau_{\text{std}} = 2.0310 \times 10^{-11}$ is lesser than $\sigma_{20} = 2.0467 \times 10^{-7}$ and is therefore not informative, whereas the pronounced gap preceding σ_{20} makes the gap-based choice $\tau_{\text{gap}} \approx 4.6 \times 10^{-3}$ (for all h) a more practical tolerance for identifying the numerical left null space. For comparison, we also computed the output difference polynomial via Gröbner basis computation in MACAULAY2 (Grayson and Stillman (2022)). For the perturbed model, the resulting polynomial $\phi_{\text{PP}g}$ contains nonzero coefficients on degree four monomials and exhibits severe coefficient growth (fractions with up to ∼105-digit integers in our experiments). Thus, while both approaches coincide in the exact (unperturbed) setting, under small perturbations the proposed numerical procedure tracks the dynamics much more closely. △

Example 2. Consider the SLO Wiener model given by

$$\Sigma_{\text{W}} : \{x[k+1] = 2x[k], \quad y[k] = x[k]^2 + x[k].$$

Upon initializing Algorithm 1 with $n_y = 1$, we recover the difference equation $\phi_{\text{W1}}(\boldsymbol{y}[k:k+1]) = 0$ where the unique, irreducible output difference polynomial $\phi_{\text{W1}} = 4y[k] - 2y[k+1] + 16y[k]^2 - 8y[k]y[k+1] + y[k+1]^2$ is of $\text{ord}(\phi_{\text{W1}}) = 1$ and $\deg(\phi_{\text{W1}}) = 2$. If we initialize Algorithm 1 with the order bound $n_y = 2$, we find the difference equation $\phi_{\text{W2}}(\boldsymbol{y}[k:k+2]) = 0$ where the linear difference polynomial, $\phi_{\text{W2}} = y[k+2] - 6y[k+1] + 8y[k]$ and $\text{ord}(\phi_{\text{W2}}) = 2$. To relate ϕ_{W1} and ϕ_{W2}, note that both $\phi_{\text{W1}}(\boldsymbol{y}[k:k+1]) = 0$ and its one-step shift $\phi'_{\text{W1}} = \phi_{\text{W1}}(\boldsymbol{y}[k+1:k+2]) = 0$. Hence $\phi'_{\text{W1}} - 4\phi_{\text{W1}}$ also vanishes on the window $\boldsymbol{y}[k:k+2] \in \mathbb{R}^3$. Moreover, $\phi'_{\text{W1}} - 4\phi_{\text{W1}} = (y[k+2] - 2y[k+1] - 8y[k] - 2)\,\phi_{\text{W2}}$, so the right-hand side vanishes on $\boldsymbol{y}[k:k+2]$ as well. Among the two factors, the one that is identically zero for all consecutive model outputs generated by Σ_{W}, given by $y[k+i] = 2^i x[k] + 4^i x[k]^2$ where $i \in \mathbb{N}$ and $x[k] \in \mathbb{R}$, is ϕ_{W2}. This restriction is implicitly enforced in the construction of the model matrix, therefore, when $n_y = 2 > n_f$, Algorithm 1 initialized with Σ_{W} returns an irreducible, output difference polynomial ϕ_{W2} that is not of minimal order but minimal degree. This is consistent with Markovsky (2019), who shows that the behavior of an autonomous Wiener model is contained in that of a higher order autonomous linear model. While ϕ_{W2} is simpler, it is only a necessary condition for Σ_{W} and is not, by itself, sufficient for model discrimination: any sequence of the form $y[k+i] = 2^i x_1[k] + 4^i x_2[k]$ where $(x_1[k], x_2[k]) \in \mathbb{R}^2$ and $i \in \mathbb{N}$ satisfies $\phi_{\text{W2}}(\boldsymbol{y}[k:k+2]) = 0$, whereas sequences of Σ_{W} correspond to the restricted subclass $x_2[k] = x_1[k]^2$. Equivalently, Σ_{W} induces a nonlinear constraint on the initial output pair: $(y[k], y[k+1]) \in V_{\text{W}} := \{(y_0, y_1) \in \mathbb{R}^2 : \phi_{\text{W1}}(y_0, y_1) = 0\}$. Only for initial conditions in V_{W} does the

Table 3. Scaling of Algorithm 1 on (11) as n_f increases. Reported are the degree $d_{\min}$ yielding a nontrivial numerical left null space, the size of the corresponding model matrix, total computation time, and the maximum degree m of the state monomials in $g(\boldsymbol{f}^i(\boldsymbol{x}[k]))$ for $i \in \{1, \ldots, n_f\}$.

n_f	$d_{\min}$	m	$\text{size}\big(\boldsymbol{\mathcal{M}}_{n_f}^{d_{\min}}\big)$	Time (s)
1	2	2	6×5	0.0195
2	3	4	20×91	0.0274
3	5	8	126×12341	1.1968
4	≥ 6	16	$\geq 462 \times 3921225$	$> 2.8575 \times 10^3$

linear recurrence $\phi_{\text{W2}}(\boldsymbol{y}[k:k+2]) = 0$ reproduce sequences compatible with Σ_{W}. Whereas $\phi_{\text{W1}}(\boldsymbol{y}[k:k+1]) = 0$ can be used to generate real output sequences for all $y[k] \in [-\frac{1}{4}, \infty)$. In this sense, the minimal-order relation provides the tight characterization of the model output behavior, while higher-order linear relations are useful but require additional nonlinear constraints. △

Example 3. We will now assess how Algorithm 1 scales on a family of SLO models given by,

$$\Sigma_{\text{B}} : \{\boldsymbol{x}[k+1] = \boldsymbol{A}\boldsymbol{x}[k] + \boldsymbol{B}\,\boldsymbol{x}_c[k], \quad y[k] = x_1[k], \quad (11)$$

where $(\boldsymbol{x}_c[k])_i = x_i[k]\,x_{i+1}[k]$, $i = 1, \ldots, n_f$, $x_{n_f+1}[k] \equiv x_1[k]$, $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n_f \times n_f}$ are diagonal matrices, randomly initialized by sampling diagonal entries from a uniform distribution on $(0, 1)$. For each $n_f = \{1, 2, 3, 4\}$ we run Algorithm 1 with the order bound n_f. Table 3 reports m, $d_{\min}$, the size($\boldsymbol{\mathcal{M}}_{n_y}^{d_{\min}}$) and the computation time [3]. As n_f increases, m grows, which forces a sharp increase in the size of $\boldsymbol{\mathcal{M}}_{n_y}^{d}$ (Lemma 1). Already at $n_f = 4$, the explicit construction of $\boldsymbol{\mathcal{M}}_4^6$ for (11) requires storage which exceeds the available memory. This motivates structure-exploiting and/or recursive alternatives for larger-scale problems. △

5. CONCLUSION

We present a linear-algebraic framework for state elimination in discrete-time, autonomous, time-invariant, single-output polynomial state-space models. The output difference equation is recovered from the left null space of a model matrix. For strongly locally observable models, the left null space of the model matrix coincides with the row space of a Macaulay matrix, generated by the unique, irreducible output difference polynomial, revealing an orthogonal complementarity between image and kernel representations. Numerical examples demonstrate robustness of the proposed state elimination algorithm to model perturbations. Future research directions include: (1) leveraging the orthogonal complementarity between the Macaulay matrix and the model matrix to solve the reverse direction (state-space realization from difference equation); (2) precisely characterizing the left null space of the model matrix corresponding for non-SLO models and when $n_y > n_f$; and (3) developing efficient recursive algorithms that utilize matrix sparsity and structure for larger, more complex models.

[3] All experiments were run on an HP EliteBook 845 G8 laptop with a 6-core AMD Ryzen 7 PRO 5850U CPU (1.90 GHz) and 32 GB RAM, on a 64-bit operating system. All timings are averaged over 5 consecutive runs and numerical values are rounded for display.

REFERENCES

Batselier, K., Dreesen, P., and De Moor, B. (2014). On the null spaces of the Macaulay matrix. *Linear Algebra and its Applications*, 460, 259–289.

Beecken, M., Mittmann, J., and Saxena, N. (2013). Algebraic independence and blackbox identity testing. *Information and Computation*, 222, 2–19.

Breiding, P., Kališnik, S., Sturmfels, B., and Weinstein, M. (2018). Learning algebraic varieties from samples. *Revista Matemática Complutense*, 31, 545–593.

Cox, D.A., Little, J., and O'Shea, D. (2010). *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics. Springer, 3rd ed.

Denis-Vidal, L., Joly-Blanchard, G., Noiret, C., and Petitot, M. (2003). System identifiability (symbolic computation) and parameter estimation (numerical computation). *Numerical Algorithms*, 34, 283–292.

Dong, R., Goodbrake, C., Harrington, H.A., and Pogudin, G. (2023). Differential elimination for dynamical models via projections with applications to structural identifiability. *SIAM Journal on Applied Algebra and Geometry*, 7(1), 194–235.

Forsman, K. (1993a). POLYCON-computer algebra software for polynomial control systems. In *Proceedings of 32nd IEEE CDC*, 282–287, vol.1.

Forsman, K. (1993b). Some generic results on algebraic observability and connections with realization theory. In *Proc. of the 2nd European Control Conference*, vol. 3, 1185–1190. Groningen.

Glad, S.T. (1989). Nonlinear state space and input output descriptions using differential polynomials. In *New Trends in Nonlinear Control Theory*, vol. 122, 182–189. Springer, Berlin, Heidelberg.

Grayson, D.R. and Stillman, M.E. (2022). Macaulay2, a software system for research in algebraic geometry. Available at `http://www2.macaulay2.com`.

Ljung, L. and Glad, T. (1994). On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2), 265–276.

Lotka, A.J. (1920). Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences*, 6(7), 410–415.

Lyzell, C., Glad, T., Enqvist, M., and Ljung, L. (2011). Difference algebra and system identification. *Automatica*, 47(9), 1896–1904.

Markovsky, I. (2019). On the behavior of autonomous Wiener systems. *Automatica*, 110, 108601.

Nijmeijer, H. (1982). Observability of autonomous discrete time non-linear systems: A geometric approach. *International Journal of Control*, 36(5), 867–874.

Ovchinnikov, A., Pogudin, G., and Scanlon, T. (2020). Effective difference elimination and nullstellensatz. *Journal of the European Mathematical Society*, 22(8), 2419–2452.

Paduart, J., Lauwers, L., Swevers, J., Smolders, K., Schoukens, J., and Pintelon, R. (2010). Identification of nonlinear systems using polynomial nonlinear state space models. *Automatica*, 46(4), 647–656.

Sontag, E.D. (1979). *Polynomial Response Maps*. Springer Berlin, Heidelberg, ed. 1.