Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 54-9 (2021) 159–165

Multiparameter Eigenvalue Problems and Shift-invariance ⋆

Katrien De Cock * **Bart De Moor, Fellow IEEE and SIAM** *

* *KU Leuven, Department of Electrical Engineering (ESAT), Leuven, Belgium*
STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics
e-mail: {katrien.decock;bart.demoor}@esat.kuleuven.be

Abstract: We discuss four eigenvalue problems of increasing generality and complexity: rooting a univariate polynomial, solving the polynomial eigenvalue problem, rooting a set of multivariate polynomials and solving multi-parameter eigenvalue problems. In doing so, we provide a unifying framework for solving these eigenvalue problems, where we exploit properties of (block-) (multi-) shift-invariant subspaces and use multi-dimensional realization algorithms.

Keywords: multiparameter eigenvalue problems, shift-invariance, realization theory

1. INTRODUCTION

The multiparameter eigenvalue problem (MEVP) is a generalization of the standard eigenvalue problem. It involves more than one eigenvalue $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ and can have many appearances, e.g.,

$$(A_0 + A_1\lambda_1 + \cdots + A_n\lambda_n)x = 0 \qquad (A_0, \ldots, A_n \in \mathbb{R}^{l\times m},\ x \in \mathbb{C}^m)\ . \quad (1)$$

Other manifestations are MEVPs containing products of eigenvalues like the following three-parameter quadratic eigenvalue problem:

$$(A_{000} + A_{100}\lambda_1 + A_{010}\lambda_2 + A_{001}\lambda_3 + A_{200}\lambda_1^2 + A_{110}\lambda_1\lambda_2 + A_{101}\lambda_1\lambda_3 + A_{020}\lambda_2^2 + A_{011}\lambda_2\lambda_3 + A_{002}\lambda_3^2)x = 0\ , \quad (2)$$

and sets of MEVPs:

$$\begin{cases} A(\lambda_1,\lambda_2,\lambda_3)x = 0 \\ B(\lambda_1,\lambda_2,\lambda_3)y = 0 \\ C(\lambda_1,\lambda_2,\lambda_3)z = 0 \end{cases} \quad (3)$$

where we look for the common eigen-triplets of three matrix pencils, for different eigenvectors x, y and z.

Despite early work by Carmichael (1921), Atkinson (1972) and others (see, e.g. Volkmer (1988)) and a recent renewed interest (Hochstenbach et al. (2019)), it is clear that the MEVP has been less studied than the standard eigenvalue problem.

We show how multiparameter eigenvalue problems can be solved by exploiting a shift-invariance property of the null space of a block-Macaulay matrix. In order to explain this, we start with simpler problems that give rise to matrices with a less intricate structure than the block-Macaulay matrix and, step by step, increase the complexity to end up with the multiparameter eigenvalue problem. Each new case adds an additional layer of complexity and provides us with new insights so that we end up with a unifying framework to understand and solve multiparameter eigenvalue problems.

For each case, the same steps are taken to go from the seed problem to its solution: first we generate additional equations by multiplying the given equation by monomials of increasing degree. This process is called the Forward Shift Recursion (FSR). It creates a structured matrix. Next, the null space of the structured matrix is computed, which for each case exhibits a specific type of shift-invariance property. The shift-invariance leads to a system-theoretic interpretation and via realization theory we obtain the solutions of the seed problem.

It will be clear that better methods exist to solve univariate polynomials and polynomial eigenvalue problems. Our presentation of the problems and their solution highlights their role in our general framework. For rooting multivariate polynomial systems, dedicated symbolic and numerical algorithms have been developed. There is a huge literature with several schools: multi-resultant-based approaches (Dickenstein and Emiris (2005)), methods using Gröbner bases (Lazard (2009); Sturmfels (2002)), homotopy methods as in Morgan (2009); Sommese and Wampler (2006). Those algorithms can also be applied to solve the MEVP, since the MEVP can be formulated as a set of multivariate polynomial equations. Indeed, the MEVPs shown in (1–3), express the fact that the matrix pencils need to be rank deficient. Algebraically, this is equivalent with the requirement that all minors of these matrix pencils of certain dimensions be zero. Such a set of 'secular equations' are multivariate polynomials in the eigen-tuples

⋆ This work was supported by (1) KU Leuven: Research Fund (projects C16/15/059, C3/19/053, C24/18/022, C3/20/117), Industrial Research Fund (Fellowships 13-0260, IOF/16/004) and several Leuven Research and Development bilateral industrial projects; (2) Flemish Government Agencies: (a) FWO: EOS Project G0F6718N (SeLMA), SBO project S005319N, Infrastructure project I013218N, TBM Project T001919N, PhD Grants (SB/1SA1319N, SB/1S93918, SB/1S1319N), (b) EWI: the Flanders AI Research Program, (c) VLAIO: Baekeland PhD (HBC.20192204) and Innovation mandate (HBC.2019.2209); (3) European Commission: European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC Adv. Grant grant agreement No 885682); (4) Other funding: Foundation 'Kom op tegen Kanker', CM (Christelijke Mutualiteit).

10.1016/j.ifacol.2021.06.071

(in that sense comparable to the notion of a characteristic equation of a matrix).

We approach the problem of rooting polynomial systems and the MEVP from the linear algebra point of view. We do not require Gröbner bases or symbolic computations, allowing us to work in finite-precision arithmetic, deploying the full power of numerical linear algebra algorithms for the singular value and eigenvalue decomposition.

2. CASE 1: UNIVARIATE POLYNOMIAL EQUATION

2.1 Seed problem

The first problem we analyze is the univariate polynomial equation:

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n = 0$$
$$(\alpha_0, \ldots, \alpha_n \in \mathbb{R},\ \alpha_n \neq 0,\ x \in \mathbb{C}) \ , \quad (4)$$

where we want to find the n roots $x_i \in \mathbb{C},\ i = 1, \ldots, n$.

To start the analysis, we use the following monic polynomial of degree 5:

$$p(x) = 4 - x - 3x^2 + 2x^3 - 3x^4 + x^5 \ . \quad (5)$$

The univariate polynomial equation $p(x) = 0$ can be written as $(4\ -1\ -3\ 2\ -3\ 1)\left(1\ x\ x^2\ x^3\ x^4\ x^5\right)^T = 0$.

2.2 Toeplitz matrix

Starting from the seed equation $p(x) = 0$, we can generate new equations by multiplying $p(x)$ by consecutive positive integer powers of x: $p(x)x^k = 0$. This process is called the Forward Shift Recursion (FSR). We then solve the system of equations, consisting of the seed equation and the additional equations. The FSR creates a structured matrix, in this case, a banded Toeplitz matrix.

When we apply the FSR for powers $k = 1, 2, 3$ to our example, we obtain

$$\underbrace{\begin{pmatrix} 4 & -1 & -3 & 2 & -3 & 1 & 0 & 0 & 0 \\ 0 & 4 & -1 & -3 & 2 & -3 & 1 & 0 & 0 \\ 0 & 0 & 4 & -1 & -3 & 2 & -3 & 1 & 0 \\ 0 & 0 & 0 & 4 & -1 & -3 & 2 & -3 & 1 \end{pmatrix}}_{T \in \mathbb{R}^{4\times 9}} \begin{pmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^8 \end{pmatrix} = 0 \ , \quad (6)$$

where the matrix T in Eq. (6) is a banded Toeplitz matrix (its elements are constant along the diagonals).

Let the five roots of the polynomial in (5) be denoted by $x_1, \ldots, x_5$, then we find in matrix notation

$$T \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 & x_5 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & x_5^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^8 & x_2^8 & x_3^8 & x_4^8 & x_5^8 \end{pmatrix}}_{K_T \in \mathbb{C}^{9\times 5}} = 0 \quad (7)$$

and the matrix K_T, whose columns span the null space of T, is a Vandermonde matrix. The fact that there is a choice of basis in the null space of an underdetermined banded Toeplitz matrix that can be in a Vandermonde structure, is a general statement, provided the roots of the generating polynomial are distinct. If they are not, a confluent Vandermonde matrix (Gautschi (1962)) is required (some columns of which contain derivatives with respect to the roots).

2.3 Backward-shift-invariance of the null space of the Toeplitz matrix

Forward- and backward-shift-invariance of a subspace is usually defined for infinite matrices (operators), see, e.g., Garcia et al. (2016). We will adapt the definition for shift-invariance to finite dimensional vector spaces.

Let $\mathrm{R}(M)$ be the range of a matrix $M \in \mathbb{C}^{m\times n}$. The backward-shift-invariance property of $\mathrm{R}(M)$ is defined as follows:

$$\mathrm{R}(M) \text{ is backward-shift-invariant iff } \mathrm{R}(\overline{M}) \subseteq \mathrm{R}(\underline{M}) \ ,$$

where $\overline{M}$ is the matrix M without its first row and $\underline{M}$ is the matrix M without its last row. The shift-invariance of $\mathrm{R}(M)$ can also be expressed as

$$\exists A \in \mathbb{C}^{n\times n} : \underline{M}A = \overline{M} \ ,$$

where $\mathrm{R}(\overline{M}) = \mathrm{R}(\underline{M})$ if A is nonsingular and $\mathrm{R}(\overline{M}) \subsetneq \mathrm{R}(\underline{M})$ otherwise.

We now return to the Toeplitz matrix T in Eq. (6). The backward-shift-invariance of the null space of T is easily verified because $\ker(T) = \mathrm{R}(K_T)$ and

$$\underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_5 \\ \vdots & \vdots & & \vdots \\ x_1^7 & x_2^7 & \cdots & x_5^7 \end{pmatrix}}_{\underline{K}_T} \underbrace{\begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & x_5 \end{pmatrix}}_{\Lambda} = \underbrace{\begin{pmatrix} x_1 & x_2 & \cdots & x_5 \\ x_1^2 & x_2^2 & \cdots & x_5^2 \\ \vdots & \vdots & & \vdots \\ x_1^8 & x_2^8 & \cdots & x_5^8 \end{pmatrix}}_{\overline{K}_T} \ , \quad (8)$$

which shows that $\mathrm{R}(\overline{K}_T) \subseteq \mathrm{R}(\underline{K}_T)$, with equality if there is no zero root in the seed equation (4), so when $\alpha_0 \neq 0$. For roots with multiplicity larger than 1, the matrix K_T has a confluent Vandermonde structure and instead of the diagonal matrix Λ, a Jordan form is needed.

2.4 Realization of single-output LTI system

Equation (8) also shows how the polynomial equation can be solved. Let $Z_T \in \mathbb{R}^{9\times 5}$ be a matrix whose columns span the null space of T, obtained from, e.g., the SVD of T. This implies that

$$\mathrm{rank}(Z_T) = 5 \ . \quad (9)$$

Since $\mathrm{R}(Z_T) = \ker(T) = \mathrm{R}(K_T)$, we know that $Z_T V = K_T$, where V is a nonsingular matrix. Furthermore, $\overline{Z}_T V = \overline{K}_T$ and $\underline{Z}_T V = \underline{K}_T$. Then, from Eq. (8), it follows that

$$\underline{Z}_T \underbrace{V \Lambda V^{-1}}_{A} = \overline{Z}_T \ , \quad (10)$$

where A is a 5×5 matrix whose eigenvalues are the roots of the polynomial.

In order to find the matrix A, the matrix Z_T has to satisfy the following rank conditions. We already know that Z_T is of full column rank, see (9). The matrix $\underline{Z}_T$ has to be of full column rank too:

$$\mathrm{rank}(\underline{Z}_T) = \mathrm{rank}(Z_T) \ , \quad (11)$$

which ensures the unicity of the solution A. This rank condition is in fact the partial realization criterion, which will become clear in Section 2.4. In addition, $\mathrm{R}(Z_T)$ needs to be shift-invariant:

$$\mathrm{rank}\left(\underline{Z}_T\ \overline{Z}_T\right) = \mathrm{rank}(\underline{Z}_T) \ .$$

So, once we have the matrix Z_T that satisfies the rank conditions $\mathrm{rank}\left(\underline{Z}_T\ \overline{Z}_T\right) = \mathrm{rank}(\underline{Z}_T) = \mathrm{rank}(Z_T) = 5$, we calculate the matrix A as

$$A = \underline{Z}_T^{\dagger} \overline{Z}_T \ ,$$

and the eigenvalues of A are the solutions of the polynomial equation. For the polynomial in (5) this results in

$$x_1=1,\ x_2=2.6450,\ x_3=-0.8184,\ x_{4,5}=0.0867\pm1.3566i\ . \quad (12)$$

Note that for the seed problem of this section, the rank conditions are automatically met if the leading coefficient of the polynomial, α_n, is not equal to zero. This is true for polynomials with distinct roots as well as polynomials with roots of multiplicity greater than one.

Relation to realization theory The null space of T can be interpreted as the range of the observability matrix of an nth order single-output LTI system. Indeed, from Eq. (10) we see that the second row of Z_T is equal to the first row of Z_T multiplied by A, the third row of Z_T is equal to the second row multiplied by A, etc. Consequently, we can write Z_T as

$$Z_T=\left(C^T\ (CA)^T\ \cdots\ (CA^8)^T\right)^T\ ,$$

where $C\in\mathbb{R}^{1\times5}$, $A\in\mathbb{R}^{5\times5}$. This is the observability matrix of a single-output autonomous linear time-invariant system of order 5, the poles of which are the roots of the polynomial.
Note that the fact that Z_T is of full column rank, as seen in (9), is equivalent to the model being observable. The second rank condition (11) is the partial realization condition, required for a unique solution for A.

Roots at infinity In order to discuss the implications of roots at infinity, which will become more important for the subsequent cases in Sections 3–5, we now interpret the polynomial of Eq. (5) as a polynomial of degree 7 with the two leading coefficients $\alpha_6=\alpha_7=0$ and the first nonzero coefficient $\alpha_5=1$. This means that the condition $\alpha_n\neq0$ in (4) is no longer met. The resulting polynomial of degree 7 has five affine roots (as in (12)) and a root at infinity of multiplicity 2. The columns of K_T corresponding to the roots at infinity can be chosen as $(0\ \cdots\ 0\ 1\ 0)^T$ and $(0\ \cdots\ 0\ 0\ 1)^T$.
It can be shown that shift-invariance gets a new meaning, as the model of the null space is now an observability matrix of a descriptor system, see Moonen et al. (1992); Dreesen et al. (2018). This implies that the null space is generated by the union of two subspaces, one that is backward-shift-invariant (and represents the causal part of the underlying state space model with dynamics modeled by the affine zeros) and another one that is forward-shift-invariant (and represents the anti-causal part of the underlying state space model dynamics which are modeled by the zeros at infinity).
Since $Z_T=K_TV$, we see that only the last two rows of Z_T are affected by the roots at infinity. Extracting the backward-shift-invariant part of $\mathrm{R}(Z_T)$ is done using the following column compression procedure (Dreesen et al. (2018)). Let $Z_T=\begin{pmatrix}Z_1\\Z_2\end{pmatrix}$, where $Z_2\in\mathbb{R}^{2\times7}$ and $Z_1=(U_1\ U_2)\begin{pmatrix}\Sigma&0\\0&0\end{pmatrix}\begin{pmatrix}Q_1^T\\Q_2^T\end{pmatrix}$ is the SVD of Z_1, $\Sigma\in\mathbb{R}^{5\times5}$. Then, Z_TQ is equal to $\begin{pmatrix}U_1\Sigma&0\\Z_2Q_1&Z_2Q_2\end{pmatrix}$ and $\mathrm{R}(U_1\Sigma)$ is a backward-shift-invariant subspace from which the affine roots can be determined.

3. CASE 2: POLYNOMIAL EIGENVALUE PROBLEM

3.1 Seed problem

The second seed problem is the regular[1] polynomial eigenvalue problem (PEVP):

$$\underbrace{(A_0+A_1\lambda+A_2\lambda^2+\cdots+A_n\lambda^n)}_{M(\lambda)}x=0$$
$$(A_0,\ldots,A_n\in\mathbb{R}^{l\times l},\ A_n\neq0,\ \lambda\in\mathbb{C},\ x\in\mathbb{C}^l)\ , \quad (13)$$

where we want to find scalars λ and nonzero vectors x satisfying $M(\lambda)x=0$.
Let r be the degree of the polynomial $q(\lambda)=\det M(\lambda)$. The r roots of $q(\lambda)$ are called the affine eigenvalues. Note that

$$q(\lambda)=\det(A_n)\lambda^{ln}+\ \text{lower order terms}\ .$$

If A_n is nonsingular, then $r=nl$, but if $\det(A_n)=0$, then $\deg(q(\lambda))=r<ln$ and besides r affine eigenvalues, there are $ln-r$ eigenvalues at infinity.
The standard way to solve the PEVP in Eq. (13) is to linearize it to a pencil of $nl\times nl$ matrices and solve the generalized eigenvalue problem (Higham et al. (2009)).
The example of this section has $n=3$, $l=2$ and the following matrices

$$A_0=\begin{pmatrix}4&1\\1&5\end{pmatrix},A_1=\begin{pmatrix}-2&3\\3&-1\end{pmatrix},A_2=\begin{pmatrix}1&-5\\-5&0\end{pmatrix},A_3=\begin{pmatrix}3&-4\\5&1\end{pmatrix}.$$

The matrix A_3 is of full rank and consequently, we will find six affine eigenvalues and corresponding eigenvectors.

3.2 Block-Toeplitz matrix

The FSR consists in multiplying the seed equation $M(\lambda)x=0$ by consecutive positive integer powers of λ. This generates new equations $M(\lambda)\lambda^kx=0$ $(k=1,2,\ldots)$ and a banded block-Toeplitz matrix arises.
For two recursions $k=1,2$ this gives

$$\underbrace{\begin{pmatrix}A_0&A_1&A_2&A_3&0&0\\0&A_0&A_1&A_2&A_3&0\\0&0&A_0&A_1&A_2&A_3\end{pmatrix}}_{\mathcal{T}\in\mathbb{R}^{6\times12}}\begin{pmatrix}x\\\lambda x\\\lambda^2x\\\vdots\\\lambda^5x\end{pmatrix}=0\ .$$

3.3 Block backward-shift-invariance of the null space of the block-Toeplitz matrix

While the null space of the Toeplitz matrix T in Section 2.3 is scalar backward-shift-invariant, the null space of $\mathcal{T}$ is *block* backward-shift-invariant. This means that $\mathrm{R}(\overline{Z}_{\mathcal{T}})\subseteq\mathrm{R}(\underline{Z}_{\mathcal{T}})$, where $\overline{Z}_{\mathcal{T}}$ and $\underline{Z}_{\mathcal{T}}$ are now equal to the matrix $Z_{\mathcal{T}}$ without its first/last block row, respectively. For our example, it is easy to see that the null space of $\mathcal{T}$ is block backward-shift-invariant when using the Vandermonde-like basis for the null space. For the distinct affine eigenvalues $\lambda_1,\ldots,\lambda_6$, the null space of $\mathcal{T}$ is spanned by the columns of

$$K_{\mathcal{T}}=\begin{pmatrix}x_1&x_2&\cdots&x_6\\\lambda_1x_1&\lambda_2x_2&\cdots&\lambda_6x_6\\\vdots&\vdots&&\vdots\\\lambda_1^5x_1&\lambda_2^5x_2&\cdots&\lambda_6^5x_6\end{pmatrix}\in\mathbb{C}^{12\times6}\ .$$

[1] $M(\lambda)x=0$ is a regular PEVP if $\det M(\lambda)\not\equiv0$

The block backward-shift-invariance, shown here for all λ simple, is obvious:

$$\underline{K}_{\mathcal{T}} \underbrace{\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \lambda_6 \end{pmatrix}}_{\Lambda} = \overline{K}_{\mathcal{T}} \; . \tag{14}$$

3.4 Realization of multi-output LTI system

Obtaining the eigenvalues and eigenvectors of the PEVP goes in a completely similar way as finding the roots of a univariate polynomial, as explained in Section 2.4:

(1) Calculate a matrix $Z_{\mathcal{T}} \in \mathbb{R}^{12\times 6}$, the columns of which span the null space of $\mathcal{T}$. The number of rows should be large enough so that the partial realization condition is met: $\text{rank}(\underline{Z}_{\mathcal{T}}) = \text{rank}(Z_{\mathcal{T}}) = nl = 6$.
(2) Determine the matrix $A \in \mathbb{R}^{nl\times nl}$ that solves the set of linear equations $\underline{Z}_{\mathcal{T}}A = \overline{Z}_{\mathcal{T}}$.
(3) The matrix A is related to the diagonal matrix Λ in Eq. (14) by a similarity transformation. Consequently, the eigenvalues of A are the affine eigenvalues of the PEVP, $\lambda_1, \ldots, \lambda_6$.

The eigenvector x_i corresponding to the calculated eigenvalue λ_i can be determined by solving the homogeneous system of linear equations $(A_0+A_1\lambda_i+A_2\lambda_i^2+A_3\lambda_i^3)x_i = 0$ for each $i = 1, \ldots, 6$. Alternatively, one can construct the matrix $K_{\mathcal{T}}$ as $K_{\mathcal{T}} = Z_{\mathcal{T}}V$, where V contains the eigenvectors of A. The first block row of $K_{\mathcal{T}}$ contains the eigenvectors x_i of the PEVP. The solutions for the given problem are

λ_i	x_i
-1.6327	$(-0.0584\ \ -0.9983)^T$
-0.8661	$(0.5187\ \ 0.8550)^T$
$0.7105\pm0.7009i$	$(-0.7842\pm0.5811i\ \ -0.1963\pm0.0942i)^T$
$0.4085\pm0.6478i$	$(-0.9513\pm0.0538i\ \ 0.2940\pm0.0752i)^T$

Relation to realization theory The null space $\text{R}(Z_{\mathcal{T}}) = \text{R}(K_{\mathcal{T}})$ can be modeled as the range of the observability matrix of an LTI system of order 6 with two outputs:

$$Z_{\mathcal{T}} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^5 \end{pmatrix} ,$$

where $C \in \mathbb{R}^{2\times 6}$, $A \in \mathbb{R}^{6\times 6}$. In general, when we solve a PEVP of degree n with $l \times l$ matrices $A_0, \ldots, A_n$ (A_n nonsingular), then we obtain the observability matrix of an autonomous linear time-invariant system of order nl with l outputs, the poles of which are the eigenvalues of the PEVP.

Eigenvalues at infinity Where we needed $\alpha_n \neq 0$ in the Toeplitz case for backward-shift-invariance (see Section 2), in the block-Toeplitz case, we need A_n to be nonsingular for the null space to be block backward-shift-invariant.
When A_n is singular, the model of the null space is an observability matrix of a descriptor system as in Section 2.4, see Moonen et al. (1992); Dreesen et al. (2018), but now with $l > 1$ outputs. The null space is a union of two subspaces, one that is backward-shift-invariant and another one that is forward-shift-invariant. We can again apply the column compression procedure explained in Section 2.4 to extract a block backward-shift-invariant subspace, on which the procedure for affine eigenvalues can be applied (see Section 4 for a worked-out example with solutions at infinity).

4. CASE 3: SET OF MULTIVARIATE POLYNOMIAL EQUATIONS

4.1 Seed problem

The third problem is solving a set of multivariate polynomial equations (here shown for three variables x_1, x_2, x_3):

$$\begin{cases} \alpha_{000} + \alpha_{100}x_1 + \alpha_{010}x_2 + \alpha_{001}x_3 + \alpha_{200}x_1^2 + \alpha_{110}x_1x_2 + \cdots &= 0 \\ \beta_{000} + \beta_{100}x_1 + \beta_{010}x_2 + \beta_{001}x_3 + \beta_{200}x_1^2 + \beta_{110}x_1x_2 + \cdots &= 0 \\ \gamma_{000} + \gamma_{100}x_1 + \gamma_{010}x_2 + \gamma_{001}x_3 + \gamma_{200}x_1^2 + \gamma_{110}x_1x_2 + \cdots &= 0 \end{cases}$$

Instead of discussing the general case, we work again with a simple example to explain our method. The set of equations we want to solve is

$$p(x, y) = y^2 - x^3 + xy^2 = 0 \tag{15a}$$
$$q(x, y) = 6.25 + x^2 - y^2 = 0 \; . \tag{15b}$$

This set of equations has four real solutions (x_i, y_i) $i = 1, \ldots, 4$ and two solutions at infinity.
The method used in this section and its relation to realization theory have been described in more detail in Dreesen et al. (2018).

4.2 Macaulay matrix

Because we now have more than one variable, it is necessary to fix an order for the different monomials. We use the degree negative lexicographic ordering (see (Batselier et al., 2013, Definition 2.1)): $1 < x < y < x^2 < xy < y^2 < x^3 < x^2y < xy^2 < y^3 < \cdots$. The two equations in (15) can be put in matrix-vector form:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & 0 \\ 6.25 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & x & y & x^2 & xy & y^2 & x^3 & x^2y & xy^2 & y^3 \end{pmatrix}^T}_{v^T} = 0 \; . \tag{16}$$

The vector v is a multidimensional (mD) generalization of the Vandermonde columns in (7).
We generate new equations by applying the FSR. Whereas in the problems of Section 2 and Section 3, we multiplied the equations with increasing powers of a single variable, namely x and λ, respectively, we now multiply the equations by all monomials in the two variables x and y up to a certain degree. Besides the seed problem $p(x, y) = q(x, y) = 0$ we then obtain the extra equations $p(x, y)x^ky^l = 0$ and $q(x, y)x^my^n = 0$ (k, l, m, n are nonnegative integers).
It turns out that for the set of equations (15), we need to 'fill up' degrees 3 and 4 in order to be able to construct a multi-shift-invariant subspace for the shifts in x and y (fulfilling the rank conditions for the null space matrix, see below). This means that we have to multiply Eq. (15a), which is of degree 3, by the monomials of degree 1 (x, y) and Eq. (15b), of degree 2, with the monomials of degree 1 and 2 (x, y, x^2, xy, y^2). This creates seven new equations

and the resulting degree 4 Macaulay matrix is then a 9×15 matrix of rank 9, denoted by M:

$$M = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 & 0 \\ 6.25 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6.25 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6.25 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6.25 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6.25 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6.25 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix}$$

For more involved problems, one usually needs more FSRs to construct a Macaulay matrix whose null space satisfies the rank conditions.

4.3 Backward-multi-shift-invariance of the null space of the Macaulay matrix

Let Z_M be a matrix whose columns span the null space of the Macaulay matrix M, obtained, e.g., by computing the right singular vectors of M that correspond to its zero singular values. When we inspect the linearly independent rows of Z_M from the top of the matrix to the bottom, we notice that the rows $1, 2, 3$ are linearly independent. Row 4 is linearly dependent on the previous three rows, row 5 is again linearly independent of all previous rows. Then, it takes until rows 11 and 12 before there are again two linearly independent rows. In total there are six linearly independent rows, which is the rank of Z_M and the number of solutions of the equations in (15), in accordance with Bézout's theorem, see (Cox et al., 2015, p. 459), provided the variety is zero-dimensional.

The distribution of the linearly independent rows in Z_M is illustrated in Figure 1. The degree 3 block only contains rows that are dependent on the previous ones. This is called the 'mind-the-gap zone'. We need this gap in order to be able to find a shift-invariant subspace and it is this gap together with the rank conditions on Z_M that made us apply the FSR until we had all degree 4 equations. If we applied more FSRs and increased the number of equations even further, we would see the gap become wider.

The two linearly independent rows at the bottom of the matrix are caused by the solutions at infinity. This is similar to what happens when a univariate polynomial has solutions at infinity, as explained in Section 2.4. The null space is a union of a causal and anti-causal shift-invariant space and an appropriate basis to separate them needs to be found. This is done by using the column compression procedure that was described in Section 2.4. The transformed 10×4 submatrix is indicated by the shaded segment in Figure 1 and denoted by Z_{M_c}. The range of Z_{M_c} is backward-*multi*-shift-invariant because there is now more than one backward shift possible, the most straightforward shifts being the backward shift in x and the backward shift in y.

We can now continue with Z_{M_c} and concentrate on the affine roots $(x_1, y_1), \ldots, (x_4, y_4)$ only. The 10×4 matrix containing the Vandermonde-like basis vectors for $\mathrm{R}(Z_{M_c})$ is denoted by $K_M = (v_1\ v_2\ v_3\ v_4)$, where the v_i vectors are defined in Eq. (16), provided the roots are distinct (a confluent Vandermonde matrix would supply a basis otherwise). The first six rows of K_M constitute a matrix of full column rank (rank $= 4$). This submatrix will play the role of *top* matrix in the realization problem; in the previous two cases (Sections 2 and 3) the *top* matrix was denoted by $\underline{K}$. Because it is of full column rank, it satisfies the partial realization condition. It can be selected from K_M by multiplying by S_t, the *top* selection matrix $S_t = (I_6\ 0)$.

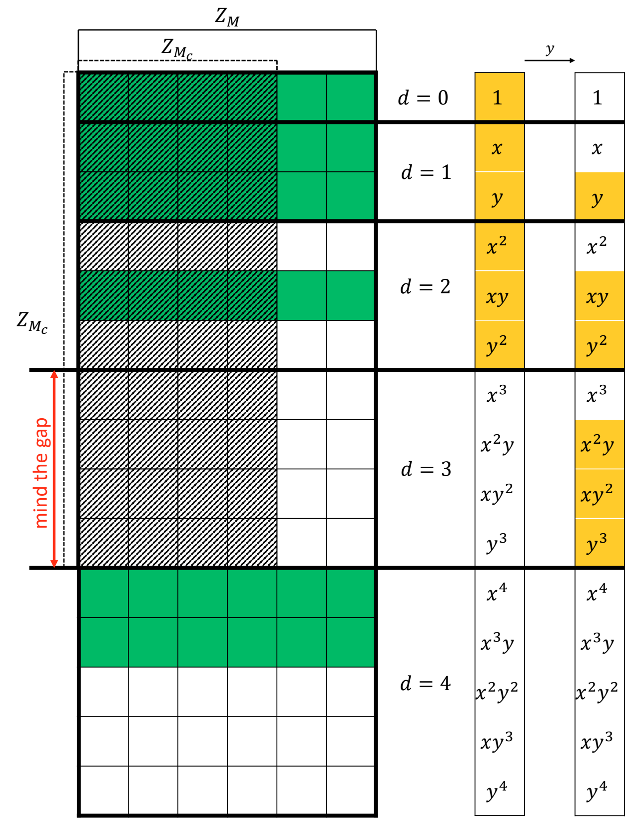

Fig. 1. Left: representation of the matrix Z_M with its linearly independent rows in green. All rows in the degree-3 block are linearly dependent on the previous rows. This is the mind-the-gap zone. The shaded columns symbolize the column compressed subspace that is multi-shift-invariant. Right: the rows in the degree $0, 1, 2$ blocks (yellow) are shifted by a y-shift to the corresponding rows.

The shifted matrix (previously denoted by $\overline{K}$), on the other hand, depends on the variable in which we want to do the shift. We choose to shift in the variable y, which is equivalent to multiplying the matrix K_M by $\Lambda_y = \begin{pmatrix} y_1 & 0 & 0 & 0 \\ 0 & y_2 & 0 & 0 \\ 0 & 0 & y_3 & 0 \\ 0 & 0 & 0 & y_4 \end{pmatrix}$. In Figure 1 we see how the rows of the degree $0, 1, 2$ blocks are shifted by y. The six rows that are the result of shifting the first six rows of K_M with y are shown on the right in yellow. They can be selected from K_M by applying the *bottom* selection matrix S_{b_y} to K_M, where $S_{b_y} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & I_3 \end{pmatrix}$. Consequently, we obtain this equation: $S_t K_M \Lambda_y = S_{b_y} K_M$. The role of the mind-the-gap zone, indicated by the red arrow in Figure 1, is apparent: it provides space to shift the rows without getting in the zone of the roots at infinity.

We can select the same rows from Z_{M_c}. Indeed, K_M and Z_{M_c} are related as $Z_{M_c} V = K_M$, where V is a nonsingular matrix. Since $S_t K_M \Lambda_y = S_{b_y} K_M$, we have $S_t Z_{M_c} \underbrace{V \Lambda_y V^{-1}}_{A_y} = S_{b_y} Z_{M_c}$ and we clearly see the backward-y-shift-invariance of the selected subspace. Because we can also find a backward-shift-invariance with other shifts, we call the null space backward-*multi*-shift-invariant.

4.4 Realization of single-output mD system

The matrix A_y, which realizes the shift in y, can be obtained as

$$A_y = (S_t Z_{M_c})^\dagger S_{b_y} Z_{M_c}$$

and the eigenvalues of A_y give us the y-values of the affine solutions. Moreover, the eigenvalue decomposition of A_y gives us the x-solutions too. Let $A_y = V \Lambda_y V^{-1}$ be the eigenvalue decomposition of A_y, then $Z_{M_c} V = K'_M$. By normalizing the columns of K'_M so that their first element is equal to 1, we obtain the complete Vandermonde-like matrix K_M and can read the affine roots (x_i, y_i) of Eq. (15) at the second and third row of K_M: $(-5, -5.5902)$, $(-5, 5.5902)$, $(-1.25, -2.7951)$, $(-1.25, 2.7951)$.

Relation to realization theory The compressed null space $\mathrm{R}(Z_{M_c}) = \mathrm{R}(K_M)$ can be interpreted as the range of the observability matrix of a single-output two-dimensional LTI system:

$$\left(C^T \ (CA_x)^T \ (CA_y)^T \ (CA_x^2)^T \ \cdots \ (CA_x A_y^2)^T \ (CA_y^3)^T\right)^T ,$$

where $A_x, A_y \in \mathbb{R}^{4\times 4}$ are commuting matrices and $C \in \mathbb{R}^{1\times 4}$.

In general, when the variety is zero-dimensional (isolated roots) and there are zeros at infinity, then the complete null space of the Macaulay matrix is the column space of an observability matrix of a multi-dimensional descriptor system, which exhibits both causal and anti-causal behavior. Therefore, the null space is the union of a backward-shift-invariant null space with causal behavior and a forward-shift-invariant subspace with anti-causal behavior.

5. CASE 4: MULTIPARAMETER EIGENVALUE PROBLEM

5.1 Seed problem

The multiparameter eigenvalue problem is the last problem that we tackle and it is the most general one. Examples of different types of MEVPs were given in Eqs. (1)–(3). The method that we use, was introduced by De Moor (2019) and Vermeersch and De Moor (2019), where the authors showed that the global optimum for two identification problems can be obtained by solving an MEVP.

We explain the solution method by looking at the following two-parameter eigenvalue problem

$$(A_0 + A_1\lambda + A_2\mu)x = 0 \ , \tag{17}$$

where $A_0 = \begin{pmatrix} 2 & -5 \\ -2 & -1 \\ 5 & -1 \end{pmatrix}$, $A_1 = \begin{pmatrix} 3 & 0 \\ 3 & -1 \\ -3 & 2 \end{pmatrix}$, $A_2 = \begin{pmatrix} 2 & 2 \\ 3 & 2 \\ -2 & -4 \end{pmatrix}$.

We want to find all eigenvectors $x_i \in \mathbb{C}^2$ and the corresponding 2-tuples of eigenvalues (λ_i, μ_i), $i = 1, \ldots, 3$.

The MEVP of Eq. (17) will be denoted by $M(\lambda,\mu)x = 0$, where the polynomial matrix $M(\lambda,\mu) = A_0 + A_1\lambda + A_2\mu$.

5.2 Block-Macaulay matrix

The FSR consists in multiplying the seed equation $M(\lambda,\mu)x = 0$ by all monomials in the variables λ and μ of increasing degree. We again use the degree negative lexicographic ordering. The FSR generates new equations $M(\lambda,\mu)\lambda^k\mu^l x = 0$ (k, l are nonnegative integers).

The matrix-vector version of the MEVP (17) is

$$(A_0 \ A_1 \ A_2) \begin{pmatrix} x \\ \lambda x \\ \mu x \end{pmatrix} = 0$$

and consequently, the FSR creates a block-Macaulay matrix. This is a generalization of the Macaulay matrix of Section 4 and was first mentioned in De Moor (2019).

For the example in (17), the block-Macaulay matrix of degree 2 is equal to

$$\mathcal{M} = \begin{pmatrix} A_0 & A_1 & A_2 & 0 & 0 & 0 \\ 0 & A_0 & 0 & A_1 & A_2 & 0 \\ 0 & 0 & A_0 & 0 & A_1 & A_2 \end{pmatrix} .$$

The columns of the following block version of the Vandermonde-like matrix provide a basis for the null space of $\mathcal{M}$, assuming only distinct and affine solutions:

$$K_{\mathcal{M}} = \begin{pmatrix} x_1 & x_2 & x_3 \\ \lambda_1 x_1 & \lambda_2 x_2 & \lambda_3 x_3 \\ \mu_1 x_1 & \mu_2 x_2 & \mu_3 x_3 \\ \lambda_1^2 x_1 & \lambda_2^2 x_2 & \lambda_3^2 x_3 \\ \lambda_1\mu_1 x_1 & \lambda_2\mu_2 x_2 & \lambda_3\mu_3 x_3 \\ \mu_1^2 x_1 & \mu_2^2 x_2 & \mu_3^2 x_3 \end{pmatrix} . \tag{18}$$

5.3 Block backward multi-shift-invariance of the null space of the block-Macaulay matrix

The null space of the block-Macaulay matrix is very similar to the null space of the Macaulay matrix (Section 4), but the first 'row' in (18) is not a vector but a matrix. Again, we have more than one possible shift, so the null space is multi-shift-invariant. In our example, we can shift with λ, μ, $\lambda\mu$, ... However, while in the Macaulay case, we selected rows in Z_{M_c} that were 'hit' by the shift, in the block-Macaulay case, when we make a shift with a certain monomial, we need to select block rows instead. Therefore, the null space of the block-Macaulay matrix is *block* backward-multi-shift-invariant.

5.4 Realization of multi-output mD system

The following steps are taken:

(1) We calculate a matrix $Z_{\mathcal{M}}$, whose columns span the null space of $\mathcal{M}$. The null space dimension is the number of roots (affine and at infinity), provided the roots are isolated. Because (17) only has affine solutions, we do not need a mind-the-gap, nor a column compression.
(2) Using the selection matrix $S_t = (I_4 \ 0)$ we select the top part of $Z_{\mathcal{M}}$ as $S_t Z_{\mathcal{M}}$, making sure it is of full column rank to satisfy the partial realization condition, and we use the selection matrix $S_{b_\lambda} = \begin{pmatrix} 0 & I_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_2 & 0 & 0 \end{pmatrix}$ to select the λ-shifted part (the block rows affected by a λ-shift) as $S_{b_\lambda} Z_{\mathcal{M}}$.

(3) The block backward-λ-shift-invariance property of the null space ensures that there is a matrix $A_\lambda \in \mathbb{R}^{3\times 3}$ so that $S_t Z_{\mathcal{M}} A_\lambda = S_{b_\lambda} Z_{\mathcal{M}}$, which is obtained as $A_\lambda = (S_t Z_{\mathcal{M}})^\dagger S_{b_\lambda} Z_{\mathcal{M}}$.

(4) Let the eigenvalue decomposition of A_λ be $A_\lambda = V\Lambda V^{-1}$. The eigenvectors of A_λ can be used to transform the matrix $Z_{\mathcal{M}}$ into the matrix $K_{\mathcal{M}} = Z_{\mathcal{M}} V$, which then delivers the eigenvectors (first block row of $K_{\mathcal{M}}$) and the corresponding eigenvalue 2-tuples (λ_i, μ_i) for $i = 1, 2, 3$:

λ_i	μ_i	x_i
3.4536	1.1169	$\begin{pmatrix} 0.1862 \\ 0.9825 \end{pmatrix}$
$-0.2268 + 1.4608i$	$0.4415 - 0.7775i$	$\begin{pmatrix} -0.4946 + 0.5971i \\ -0.5972 + 0.2053i \end{pmatrix}$
$-0.2268 - 1.4608i$	$0.4415 + 0.7775i$	$\begin{pmatrix} -0.4946 - 0.5971i \\ -0.5972 - 0.2053i \end{pmatrix}$

Solutions at infinity For the MEVP case too, solutions at infinity are possible. We then have to do enough FSRs to ensure the mind-the-gap zone exists in the null space matrix and apply column compression as explained in Section 2.4 and applied in Section 4.3. This allows us to separate the affine eigenvalues from those at infinity. A block backward-multi-shift-invariant subspace can be extracted from which the affine solutions follow.

Relation to realization theory The column space of $K_{\mathcal{M}}$ or $Z_{\mathcal{M}}$ in our example, can be seen as the range of an observability matrix of a 2D commutative system of order 3 with two outputs:

$$\begin{pmatrix} C^T & (CA_\lambda)^T & (CA_\mu)^T & (CA_\lambda^2)^T & (CA_\lambda A_\mu)^T & (CA_\mu^2)^T \end{pmatrix}^T,$$

where $C \in \mathbb{R}^{2\times 3}$, $A_\lambda, A_\mu \in \mathbb{R}^{3\times 3}$ and $A_\lambda A_\mu = A_\mu A_\lambda$.
If there were solutions at infinity, then the whole column space could be modeled as the range of the observability matrix of an mD commutative descriptor system.

6. CONCLUSIONS

In this paper we presented four problems of increasing complexity (rooting a univariate polynomial, solving a polynomial eigenvalue problem, rooting a set of multivariate polynomials, solving a multiparameter eigenvalue problem) that can be solved using the same steps:

(1) create a structured matrix by generating new equations using the Forward Shift Recursion,
(2) calculate the null space of the matrix and check its shift-invariance property,
(3) apply realization theory to find the solutions by solving an eigenvalue problem.

We have shown that solving a multiparameter eigenvalue problem (MEVP) boils down to solving a standard eigenvalue problem. This has already led to new theoretical insights about globally optimal solutions to system identification problems in De Moor (2019, 2020) and Vermeersch and De Moor (2019). Many more optimization problems can be formulated as an MEVP. For all these problems, the global optimum can be obtained by solving a standard eigenvalue problem. Future work will also be concerned with making our algorithms faster and more feasible by exploiting the structure and sparsity of the matrices involved.

REFERENCES

Atkinson, F.V. (1972). *Multiparameter Eigenvalue Problems Volume 1: Matrices and Compact Operators*. Academic Press.

Batselier, K., Dreesen, P., and De Moor, B. (2013). The geometry of multivariate polynomial division and elimination. *SIAM Journal on Matrix Analysis and Applications*, 34(1), 102–125.

Carmichael, R.D. (1921). Boundary value and expansion problems: Algebraic basis of the theory. *American Journal of Mathematics*, 43(2), 69–101.

Cox, D.A., Little, J., and O'Shea, D. (2015). *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer.

De Moor, B. (2019). Least squares realization of LTI models is an eigenvalue problem. In *Proceedings of the European Control Conference (ECC 2019)*, 2270–2275.

De Moor, B. (2020). Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem. *Communications in Information and Systems*, 20(2), 163–207.

Dickenstein, A. and Emiris, I.Z. (eds.) (2005). *Solving Polynomial Equations: Foundations, Algorithms, and Applications*. Springer.

Dreesen, P., Batselier, K., and De Moor, B. (2018). Multidimensional realisation theory and polynomial system solving. *International Journal of Control*, 91(12), 2692–2704.

Garcia, S.R., Mashreghi, J., and Ross, W.T. (2016). *Introduction to Model Spaces and their Operators*. Cambridge University Press.

Gautschi, W. (1962). On inverses of Vandermonde and confluent Vandermonde matrices. *Numerische Mathematik*, 4(1), 117–123.

Higham, N.J., Mackey, D.S., and Tisseur, F. (2009). The conditioning of linearizations of matrix polynomials. *SIAM Journal on Matrix Analysis and Applications*, 28(4), 1005–1028.

Hochstenbach, M.E., Meerbergen, K., Mengi, E., and Plestenjak, B. (2019). Subspace methods for three-parameter eigenvalue problems. *Numerical Linear Algebra with Applications*, 26(4).

Lazard, D. (2009). Thirty years of polynomial system solving, and now? *Journal of Symbolic Computation*, 44(3), 222–231.

Moonen, M., De Moor, B., Ramos, J., and Tan, S. (1992). A subspace identification algorithm for descriptor systems. *Systems & Control Letters*, 19, 47–52.

Morgan, A. (2009). *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems*. SIAM.

Sommese, A.J. and Wampler, C.W. (2006). *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific.

Sturmfels, B. (2002). *Solving Systems of Polynomial Equations*. American Mathematical Society.

Vermeersch, C. and De Moor, B. (2019). Globally optimal least-squares ARMA model identification is an eigenvalue problem. *IEEE Control Systems Letters*, 3(4), 1062–1067.

Volkmer, H. (1988). *Multiparameter Eigenvalue Problems and Expansion Theorems*. Springer.