A Gröbner Basis Approach to Solve a Rank Minimization Problem Arising in 2D-identification

P. Rapisarda *

* *Vision, Learning and Control Group, School of Electronics and Computer Science, University of Southampton, UK,*
(e-mail: `pr3@ecs.soton.ac.uk`*)*

Abstract: The problem of state-space modelling of 2D-trajectories from exponential data can be solved using a duality approach. Finding a minimal complexity model, i.e. one having the minimal number of state variables among those unfalsified by the data, can be transformed to a rank-minimization problem involving constant matrices computed from the data. We illustrate a Gröbner basis approach to solve such problem.

Keywords: 2D systems, 2D-identification, Roesser models, 2D matrix Sylvester equation, rank-minimization problems

1. INTRODUCTION

Consider N $2D$ $\mathtt{w}$-dimensional vector-exponential trajectories $w_i(\bullet,\bullet)$, whose value at $(t_1,t_2) \in \mathbb{R}\times\mathbb{R}$ is

$$w_i(t_1,t_2) = \overline{w}_i e^{\lambda_1^i t_1} e^{\lambda_2^i t_2} \ , \tag{1}$$

where $\lambda_j^i \in \mathbb{C}$, $i = 1,\ldots,N$, $j = 1,2$, and we use the overline notation $\overline{w}_i \in \mathbb{C}^{\mathtt{w}}$ to indicate vectors. In this paper we assume that an *input-output partition* $w = \text{col}(u,y)$ of the variables is known. We also assume that the system generating the data is *controllable* in the behavioral sense (see sect. 4 of Zerz (2004) for a definition), so that its input-output behavior, described by the *transfer function*, uniquely determines the system. In the following we use such crucial assumption to conclude that the data-generating system has a well-defined *dual*, defined in sect. 2.1 of this paper. Finally, we assume that every entry of the $\mathtt{p}\times\mathtt{m}$ rational matrix describing the transfer function is of the form $\frac{n(s_1,s_2)}{d(s_1,s_2)} = \frac{\sum_{i=0}^m n_i(s_1)s_2^i}{\sum_{j=0}^n d_j(s_1)s_2^j}$ for some nonnegative integers m and n (dependent on the entry) with the univariate polynomials $n_m(s_1), d_n(s_1) \neq 0$ satisfying the following three properties:

(1) $m \leq n$
(2) $\deg(d_n(s_1) \geq \deg(n_i(s_1))$, $i = 0,\ldots,m-1$
(3) $\deg(d_n(s_1)) \geq \deg(d_i(s_1))$, $i = 0,\ldots,n$.

Using the terminology of the discrete case, we call rational functions of this type "*quarter-plane causal*".

Under such assumptions, the data-generating system can be represented by a *Roesser representation*:

$$\begin{bmatrix} \frac{\partial}{\partial t_1} x_1 \\ \frac{\partial}{\partial t_2} x_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u$$

$$y = [C_1 \ \ C_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + Du \ , \tag{2}$$

where $A \in \mathbb{R}^{(\mathtt{n}_1+\mathtt{n}_2)\times(\mathtt{n}_1+\mathtt{n}_2)}$, $B := \text{col}(B_1,B_2) \in \mathbb{R}^{(\mathtt{n}_1+\mathtt{n}_2)\times\mathtt{m}}$, $C := [C_1 \ C_2] \in \mathbb{R}^{\mathtt{p}\times(\mathtt{n}_1+\mathtt{n}_2)}$, $D \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}$, and the external variable $w := \text{col}(u,y)$. Such representations were introduced in Roesser (1975) in the discrete case; in the continuous case they can be used to describe physical phenomena such as the Darboux equation arising in gas absorption and water evaporation; transmission lines; and thermal processes (see Kaczorek (1985)).

The trajectory $\text{col}(x_1,x_2)$ in (2) is called the *state trajectory* corresponding to w. We call the dimension $\mathtt{n}_1+\mathtt{n}_2$ of the state variable $x = \text{col}(x_1,x_2)$ the *complexity* of the model (2).

The problem of *state-modelling* of (1) consists in finding matrices A, B, C and D such that (2) are satisfied by some trajectories $x_i : \mathbb{R}\times\mathbb{R} \to \mathbb{R}^{\mathtt{n}_1+\mathtt{n}_2}$ and the data $w_i = \text{col}(u_i,y_i)$, $i = 1,\ldots,N$ in (1).

In the paper Rapisarda and Antoulas (2016), the authors proposed a duality-based approach to solve such identification problem, whose main features we illustrate in sect. 2.3 of this paper. A crucial step in such approach is the solution of a *2D matrix Sylvester equation* involving a constant matrix $\mathbb{L}$ derived from (1) and its "dual" data, and diagonal matrices M_i, Λ_i, $i = 1,2$ associated with the frequencies μ_i^k and λ_i^k, $i = 1,2$, $k = 1,\ldots,N$ of the primal and dual data:

$$\mathbb{L} = M_1^* S_1 + S_1\Lambda_1 + M_2^* S_2 + S_2\Lambda_2 \ , \tag{3}$$

where S_i, $i = 1,2$ are the unknown matrices. Factorizing in a rank-revealing way the solutions S_i of (3), *state trajectories* can be computed corresponding to the data. The matrices A, B, C and D can then be computed in a straightforward way solving a system of linear equations.

It can be shown that the dimension $\mathtt{n}$ of the state variable $x = \text{col}(x_1,x_2)$ of an unfalsified model computed by our procedure equals $\text{rank}(S_1) + \text{rank}(S_2)$. Consequently, if a *low*, ideally *minimal complexity model* for (1) is sought, the problem arises of how to find solutions S_i, $i = 1,2$ to

(3) such that rank(S_1)+rank(S_2) is low (or minimal). This is an *affine rank minimization* problem:

$$\begin{aligned}&\text{Minimize rank } S_1 + \text{rank } S_2\\&\text{subject to } \mathcal{A}(S_1,S_2) = b\ ,\end{aligned} \tag{4}$$

where $\mathcal{A}$ is a linear map obtained in a straightforward way from the right-hand side of (3), and b is a vector obtained from the Loewner matrix L^*JR.

In this paper we illustrate a procedure based on symbolic algebra and Gröbner basis computations to solve this problem, which computes a parametrization of *all* solutions (S_1, S_2) to (3) with a given rank, opening up the possibility of exploring such parameter space in order to compute unfalsified models with special properties.

Notation

We denote by $\mathbb{C}^{m\times n}$ the set of all $m \times n$ matrices with entries in $\mathbb{C}$. $\mathbb{C}^{\bullet\times n}$ denotes the set of matrices with n columns and an unspecified (finite) number of rows. Given $A \in \mathbb{C}^{m\times n}$, we denote by A^* its conjugate transpose and by $A^\dagger$ its Moore-Penrose inverse. If A, B are matrices with the same number of columns, col(A, B) is the matrix obtained stacking A on top of B. $e^{\lambda_1\bullet}e^{\lambda_2\bullet}$ denotes the function from $\mathbb{R}^2$ to $\mathbb{C}$ whose value at (t_1, t_2) is $e^{\lambda_1 t_1}e^{\lambda_2 t_2}$.

2. FROM DATA TO STATE MODEL

2.1 Duality and divergence of fields of state variables

We associate to a Roesser representation (2) its *dual* one, defined by the equations:

$$\begin{bmatrix}\frac{\partial}{\partial \mathtt{t}_1}x_1'\\ \frac{\partial}{\partial \mathtt{t}_2}x_2'\end{bmatrix} = -A^\top \begin{bmatrix}x_1'\\x_2'\end{bmatrix} + \begin{bmatrix}C_1^\top\\C_2^\top\end{bmatrix}u'$$

$$y' = \begin{bmatrix}B_1^\top & B_2^\top\end{bmatrix}\begin{bmatrix}x_1'\\x_2'\end{bmatrix} - D^\top u'\ . \tag{5}$$

It can be shown that the set $\mathfrak{B}'$ of external trajectories $w' := \text{col}(u', y')$ corresponding to the state-representation (5) is also controllable (see Rapisarda and Antoulas (2016)).

Theorem 1. Let $\mathfrak{B}, \mathfrak{B}'$ be the input-output behaviors corresponding to the state representations (2) and (5), respectively. Let $w = \text{col}\,(u, y) \in \mathfrak{B}$ and $w' = \text{col}\,(u', y') \in \mathfrak{B}'$, with associated state trajectories $x = \text{col}(x_1, x_2)$ and $x' = \text{col}(x_1', x_2')$, respectively. Then

$$[u^*\ \ y^*]\underbrace{\begin{bmatrix}0_{\mathtt{m}\times\mathtt{p}} & I_{\mathtt{m}}\\ I_{\mathtt{p}} & 0_{\mathtt{p}\times\mathtt{m}}\end{bmatrix}}_{J:=}\begin{bmatrix}u'\\y'\end{bmatrix} = \frac{\partial}{\partial \mathtt{t}_1}(x_1^* x_1') + \frac{\partial}{\partial \mathtt{t}_2}(x_2^* x_2')\ . \tag{6}$$

Proof. The claim is a matter of straightforward verification using the equations (2) and (5). □

Th. 1 is a crucial result: the J-inner product of *input-output* trajectories of the primal and the dual system is the *divergence* of a field, whose components are the inner products of the *state* trajectories of the primal and the dual system in the first- and second independent variable.

2.2 The 2D matrix Sylvester equation

In the rest of the paper we assume that a set of N dual trajectories is known; such assumption is not restrictive, since dual trajectories can be readily computed from primal ones (see Rem. 1 below). Consider the primal and dual data:

$$\begin{aligned}w_i(\bullet,\bullet) &= \text{col}(u,y)(\bullet,\bullet) = \begin{bmatrix}\overline{u_i}\\ \overline{y_i}\end{bmatrix}e^{\lambda_1^i\bullet}e^{\lambda_2^i\bullet} \in \mathfrak{B}\ ,\\ w_i'(\bullet,\bullet) &= \text{col}(u',y')(\bullet,\bullet) = \begin{bmatrix}\overline{u_i'}\\ \overline{y_i'}\end{bmatrix}e^{\mu_1^i\bullet}e^{\mu_2^i\bullet} \in \mathfrak{B}'\ ,\end{aligned} \tag{7}$$

$i = 1,\ldots,N$. It is easy to verify from (2), (5) that to such trajectories correspond vector-exponential state trajectories

$$\overline{x}_i e^{\lambda_1^i\bullet}e^{\lambda_2^i\bullet}\ ,\ \overline{x}_j' e^{\mu_1^j\bullet}e^{\mu_2^j\bullet}\ , \tag{8}$$

where $\overline{x}_i', \overline{x}_i \in \mathbb{C}^{\mathtt{n}}$, $i, j = 1,\ldots,N$. We partition $\overline{x}_i' =: \text{col}(\overline{x}_{i,1}', \overline{x}_{i,2}')$ and $\overline{x}_i =: \text{col}(\overline{x}_{i,1}, \overline{x}_{i,2})$ according to the partition of the state trajectories in (5) and (2).

Define from (7), (8) the matrices

$$\begin{aligned}&L := [w_1'\ \ldots\ w_N'] \in \mathbb{C}^{\mathtt{w}\times N}\ ,\ R := [w_1\ \ldots\ w_N] \in \mathbb{C}^{\mathtt{w}\times N}\\ &\Lambda_i := \text{diag}(\lambda_i^k)_{k=1,\ldots,N}\ M_i := \text{diag}(\mu_i^k)_{k=1,\ldots,N}\ i = 1,2\\ &X' := [\overline{x_1}'\ \ldots\ \overline{x_{N'}}] =: \begin{bmatrix}X_1'\\X_2'\end{bmatrix}\ ,\ X := [\overline{x_1}\ \ldots\ \overline{x_N}] =: \begin{bmatrix}X_1\\X_2\end{bmatrix}\ .\end{aligned} \tag{9}$$

The next result shows that matrices computed from the external data and matrices computed from the internal (i.e. state) ones are related.

Proposition 2. Define the matrices L, R, Λ_i, M_i, X', X by (9). Define $\mathbb{L} := L^*JR$; then

$$\mathbb{L} = M_1^* X_1'^* X_1 + X_1'^* X_1 \Lambda_1 + M_2^* X_2'^* X_2 + X_2'^* X_2 \Lambda_2\ . \tag{10}$$

Proof. Compute the value at $(0, 0)$ of equation (6) on the external data and their associated state trajectories.

Remark 1. Given a set of external trajectories of the primal, it is straightforward to compute a set of external trajectories of the dual, as the following result shows.

Proposition 3. Let $\mathfrak{B}$ be a 2D controllable behavior, and $J \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$ be an involution. Let $\overline{w}e^{\lambda_1\bullet}e^{\lambda_2\bullet} \in \mathfrak{B}$, and let $\overline{v} \in \mathbb{C}^{\mathtt{w}}$ satisfy $\overline{v}^*\overline{w} = 0$. Then $J\overline{v}e^{-\lambda_1^*\bullet}e^{-\lambda_2^*\bullet}$ is an external trajectory of the dual system.

Proof. See Prop. 5.5 of Rapisarda and Antoulas (2016).

Prop. 3 is at the heart of the *mirroring* technique already applied successfully to $1D$ interpolation and modelling problems, see Kaneko and Rapisarda (2003, 2007); Rapisarda and Willems (1997). □

2.3 From data to state trajectories

The following result follows from Prop. 2.

Theorem 4. Let $\mathfrak{B}, \mathfrak{B}'$ be the external behaviors corresponding to the state representations (2) and (5). Let data (7) be given. Define $L, R,\ \Lambda_i, M_i,\ i = 1, 2$ by (9) and $\mathbb{L} := L^*JR$. Define

$$\begin{aligned}&U := [\overline{u_1}\ \ldots\ \overline{u_N}] \in \mathbb{C}^{\mathtt{m}\times N}\ ,\ Y := [\overline{y_1}\ \ldots\ \overline{y_N}] \in \mathbb{C}^{\mathtt{p}\times N}\\ &U' := [\overline{u_1'}\ \ldots\ \overline{u_N'}] \in \mathbb{C}^{\mathtt{p}\times N}\ ,\ Y' := [\overline{y_1'}\ \ldots\ \overline{y_N'}] \in \mathbb{C}^{\mathtt{m}\times N}\end{aligned} \tag{11}$$

There exist $\mathtt{n}_i \in \mathbb{N}$, and matrices $X_i, X_i' \in \mathbb{C}^{\mathtt{n}_i \times N}$ $i = 1, 2$, such that (10) holds. Moreover, there exist $A_{ij} \in \mathbb{R}^{\mathtt{n}_i \times \mathtt{n}_j}$, $i = 1, 2$, $C_i \in \mathbb{R}^{\mathtt{p} \times \mathtt{n}_i}$, $B_i \in \mathbb{R}^{\mathtt{n}_i \times \mathtt{m}}$, $i = 1, 2$ such that the following equations hold:

$$\begin{bmatrix} X_1\Lambda_1 \\ X_2\Lambda_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} U$$
$$Y = [C_1 \ C_2] \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + DU \ ,$$
$$\begin{bmatrix} X_1'M_1 \\ X_2'M_2 \end{bmatrix} = -\begin{bmatrix} A_{11}^\top & A_{21}^\top \\ A_{12}^\top & A_{22}^\top \end{bmatrix} \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} + \begin{bmatrix} C_1^\top \\ C_2^\top \end{bmatrix} U'$$
$$Y' = \begin{bmatrix} B_1^\top & B_2^\top \end{bmatrix} \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} - D^\top U' \ . \tag{12}$$

Given such matrices A_{ij}, B_i, C_i, $i, j = 1, 2$, equations (2) and (5) define unfalsified Roesser models for the data (7).

Proof. Denote by $x_i = \mathrm{col}(x_{1,i}, x_{2,i})$, $x_i' = \mathrm{col}(x_{1,i}', x_{2,i}')$, the state trajectories associated in the state representations of the primal, respectively dual system, with w_i, respectively w_i', $i = 1, \ldots, N$. Now consider the value at $(0, 0)$ of (2) and (5) with such external- and state trajectories, and define

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} := \begin{bmatrix} x_{1,1}(0,0) & \ldots & x_{1,N}(0,0) \\ x_{2,1}(0,0) & \ldots & x_{2,N}(0,0) \end{bmatrix} ,$$

and $\begin{bmatrix} X_1' \\ X_2' \end{bmatrix}$ analogously. Evidently the state-, input- and output matrices of the primal and the dual state representations satisfy (12). This argument proves the first part of the theorem and the equations (12). The last part of the claim is straightforward. □

Based on Th. 4, in Rapisarda and Antoulas (2016) the following $2D$ identification procedure was proposed. The first step consists in constructing the Loewner matrix $\mathbb{L}$ from the data (7). In the second step, one computes a pair (S_1, S_2) of solutions to the linear matrix equation (3). The third step consists in the rank-revealing factorization of $S_k = F_k'^\top F_k$, i.e. $\mathrm{rank}(S_k) = \mathrm{rank}(F_k) = \mathrm{rank}(F_k')$, $k = 1, 2$. Then, defining U and Y as in (11), an unfalsified Roesser module for the data is computed solving for A_{ij}, B_i, C_i, $i, j = 1, 2$ and D in

$$\begin{bmatrix} F_1\Lambda_1 \\ F_2\Lambda_2 \\ Y \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ C_1 & C_2 & D \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ U \end{bmatrix} . \tag{13}$$

Note that the dimension $\mathtt{n}_i$ of the x_i variable of the model computed with such procedure equals $\mathrm{rank}(S_i)$, $i = 1, 2$. It follows that to compute *low-complexity* unfalsified models for the primal data it is crucial to find low-rank solutions S_i, $i = 1, 2$ to equation (3).

In order to compute *real* models (2), (5), in the rest of this paper we will assume that the data sets $\mathfrak{D}' := \{w_i'\}_{i=1,\ldots,N}$ and $\mathfrak{D} := \{w_i\}_{i=1,\ldots,N}$ are *closed under conjugation*, i.e. that

$$w_i(\bullet, \bullet) = \overline{w}_i \lambda_{1,i}^\bullet \lambda_{2,i}^\bullet \in \mathfrak{D} \Longrightarrow w_i^*(\bullet, \bullet) = \overline{w}_i^* \lambda_{1,i}^{*\bullet} \lambda_{2,i}^{*\bullet} \in \mathfrak{D}$$
$$w_i'(\bullet, \bullet) = \overline{w}_i' \mu_{1,i}^\bullet \mu_{2,i}^\bullet \in \mathfrak{D}' \Longrightarrow w_i'^*(\bullet, \bullet) = \overline{w}_i'^* \mu_{1,i}^{*\bullet} \mu_{2,i}^{*\bullet} \in \mathfrak{D}' \ .$$

Such assumptions guarantee that the least squares solution of (13) consists of real matrices.

Remark 2. Ramos and his collaborators pioneered an identification approach based on the idea of computing 2D state-trajectories from data and subsequently using them, together with the input-output trajectories, to compute state-equations (see Farah et al. (2014); Ramos (1994)). Such approach is fundamentally based on shift-invariance, while we exploit the *pairing* relation (6), as in the $1D$ *Loewner framework*. See Antoulas and Rapisarda (2015); Rapisarda and Antoulas (2015) for more details. □

3. THE 2D MATRIX SYLVESTER EQUATION

The 2D matrix Sylvester equation (3) is linear in the unknowns S_i, $i = 1, 2$. It follows that we can find all solutions to (3) by computing a particular one, and summing to it the general form of the corresponding homogeneous 2D matrix Sylvester equation.

Using standard results in the theory of matrix Sylvester equations (see e.g. Peeters and Rapisarda (2006)) it can be shown that if the spectra of Λ_i and $-M_i^*$ are disjoint, then for every choice of the right-hand side Q there exist solutions $\overline{S}_i$ to the i-th *1D matrix Sylvester equation*

$$M_i^* \overline{S}_i + \overline{S}_i \Lambda_i = Q \ , \tag{14}$$

$i = 1, 2$. If $Q = \frac{1}{2}\mathbb{L}$ and if $\overline{S}_i$ solves the 1D Sylvester equation (14), $i = 1, 2$, then summing up the two 1D Sylvester equations it follows that the pair $(\overline{S}_1, \overline{S}_2)$ solves (3). From such discussion it follows that under the assumption $\sigma(\Lambda_i) \cap \sigma(M_i) = \emptyset$, a special solution of the 2D matrix Sylvester equation can always be computed solving two 1D Sylvester equations. Note that such spectral assumption is non-restrictive as long as the frequencies of the experimental data can be freely chosen.

To find all solutions of (3), consider the *homogeneous 2D matrix Sylvester equation*

$$0 = M_1^* S_1' + S_1' \Lambda_1 + M_2^* S_2' + S_2' \Lambda_2 \ , \tag{15}$$

in the unknowns S_i', $i = 1, 2$. The fact that for given Λ_i, M_i matrices, $i = 1, 2$, equation (15) has nontrivial (i.e. nonzero) solutions can be verified using standard linear algebra arguments [1] .

The following result is a straightforward consequence of this discussion.

Proposition 5. Assume that $\lambda_1^i + \mu_1^{j*} \neq 0$ and $\lambda_2^i + \mu_2^{j*} \neq 0$, $i, j = 1, \ldots, N$. There exist solutions $\overline{S}_i$, $i = 1, 2$, to (14) when $Q = \frac{1}{2}\mathbb{L}$. Define

$$\mathcal{S} := \{(S_1', S_2') \mid S_1' \text{ and } S_2' \text{ solve (15)}\} \ . \tag{16}$$

$(S_1, S_2) \in \mathbb{R}^{N \times N} \times \mathbb{R}^{N \times N}$ is a solution to (3) if and only if there exists $(\overline{S}_1', \overline{S}_2') \in \mathcal{S}$ such that $S_i = \overline{S}_i + \overline{S}_i'$, $i = 1, 2$.

A parametrization of $\mathcal{S}$ defined in (16) is straightforward to derive from the following result.

Proposition 6. Denote the (k, j)-th entry of S_i' in (15) by $S_i'(k, j)$, $i = 1, 2$, $k, j = 1, \ldots, N$. A pair (S_1', S_2') is a solution of (15) if and only if

$$S_1'(k, j) \left(\lambda_1^j + \mu_1^{k*}\right) + S_2'(k, j) \left(\lambda_2^j + \mu_2^{k*}\right) = 0 \ ,$$

[1] One can reach the same conclusion observing that (3) arises from (6), and that there exist non-zero fields that have zero divergence.

$k, j = 1, \ldots, N$. Assume that $\lambda_1^j + \mu_1^{k*} \neq 0$, $k, j = 1, \ldots, N$; then (S_1', S_2') solves (15) if and only if

$$S_1'(k, j) = -S_2'(k, j) \frac{\lambda_2^j + \mu_2^{k*}}{\lambda_1^j + \mu_1^{k*}} = 0 \,, \qquad (17)$$

$k, j = 1, \ldots, N$.

Proof. Straightforward.

The following consequence of Prop. 6 is instrumental in our procedure for computing low-rank solutions to equation (3).

Corollary 7. Assume that $\lambda_1^j + \mu_1^{k*} \neq 0$, $k, j = 1, \ldots, N$. Denote by $\overline{S}_i$, $i = 1, 2$, the unique solution to (14) when $Q = \frac{1}{2}\mathbb{L}$. Then (S_1, S_2) is a solution of (3) if and only if there exist $S_2'(k, j) \in \mathbb{R}$, $k, j = 1, \ldots, N$, such that

$$\begin{aligned} S_1(k, j) &= \overline{S}_1(k, j) - S_2'(k, j) \frac{\lambda_2^j + \mu_2^{k*}}{\lambda_1^j + \mu_1^{k*}} \\ S_2(k, j) &= \overline{S}_2(k, j) - S_2'(k, j) \frac{\lambda_2^j + \mu_2^{k*}}{\lambda_1^j + \mu_1^{k*}} \,, \end{aligned} \qquad (18)$$

$k, j = 1, \ldots, N$.

4. GRÖBNER BASES AND LOW-RANK SOLUTIONS TO THE 2D MATRIX SYLVESTER EQUATION

In Cor. 7 we obtained a parametrization of all solutions to (3) by choosing S_2' freely in $\mathbb{R}^{N \times N}$. Defining the polynomial ring $\mathbb{R}[\{\xi_{kj}\}_{k,j=1,\ldots,N}]$ and identifying ξ_{kj} with the free parameter $S_2'(k, j)$, the matrices S_1 and S_2 defined in (18) are associated in a natural way with polynomial matrices in the ring $\mathbb{R}[\{\xi_{kj}\}_{k,j=1,\ldots,N}]^{N \times N}$.

To state our procedure to find minimal rank solutions to (3) we use the following notation. If $\mathcal{S}$ is a subset of a polynomial ring, we denote by $\mathcal{I}(\mathcal{S})$ the ideal generated by $\mathcal{S}$; and if $\mathcal{I}$ is an ideal, we denote by $\mathcal{G}(\mathcal{I})$ the Gröbner basis (in some well-ordering) of $\mathcal{I}$. Moreover, we denote by $\mathcal{V}(\mathcal{I})$ the algebraic variety of a polynomial ideal $\mathcal{I}$.

Algorithm

For $n = 1, \ldots, N$ do
  For all $(\mathtt{n}_1, \mathtt{n}_2) \in \{(\mathtt{n}_1, \mathtt{n}_2) \in \mathbb{N} \times \mathbb{N} \mid \mathtt{n}_1 + \mathtt{n}_2 = n\}$ do
    Compute $\mathcal{M}_i := \{(\mathtt{n}_i + 1)\text{-minors of } S_i\}$, $i = 1, 2$
    Compute $\mathcal{G}_i := \mathcal{G}(\mathcal{I}(\mathcal{M}_i))$, $i = 1, 2$
    If $\overline{\mathcal{G}} := \mathcal{G}(\mathcal{G}_1 \cup \mathcal{G}_2) \neq \{1\}$ then
      Choose $v \in \mathcal{V}(\mathcal{I}(\overline{\mathcal{G}}))$
      Return $(S_1(v), S_2(v))$
    Else $n := n + 1$
  End for
End for

Proposition 8. Assume that $\lambda_1^j + \mu_1^{k*} \neq 0$ and $\lambda_2^j + \mu_2^{k*} \neq 0$, $k, j = 1, \ldots, N$. Then the algorithm provides a solution pair (S_1, S_2) to (3) such that $\text{rank}(S_1) + \text{rank}(S_2)$ is minimal.

Proof. A matrix has rank less than or equal to $\mathtt{n}_i$ if and only if all its minors of order $\mathtt{n}_i + 1$ or higher are zero. Consequently, values for the free parameters $S_2'(k, j)$ such that the matrices defined in (18) have rank less than or equal to $\mathtt{n}_i$ exist if and only if the two systems of polynomial equations in the indeterminates ξ_{kj}, $k, j = 1, \ldots, N$ obtained from such minors condition have a common solution.

It follows from standard Gröbner basis considerations that such systems of polynomial equations have a common solution if and only if the ideal generated by the basis $\overline{\mathcal{G}}$ is nontrivial, equivalently if and only if $\overline{\mathcal{G}} \neq \{1\}$.

To conclude the proof, note that there exist solutions of rank N to (3), since the disjoint spectra condition guarantees the existence of solutions to the 1D Sylvester equation (14) with right-hand side $Q = \mathbb{L}$; from any such solution S_1 one can construct a pair $(S_1, 0_{N \times N})$ to (3). Thus the procedure terminates after a finite number of steps. □

We conclude this section with some remarks.

Remark 3. The innermost **For** loop iteration can be performed in a *total-degree lexicographic* way, i.e. searching the space of possible model complexities $(\mathtt{n}_1, \mathtt{n}_2)$ as in Fig. 1, and giving priority to models with lower dimension of the x_1-variable.

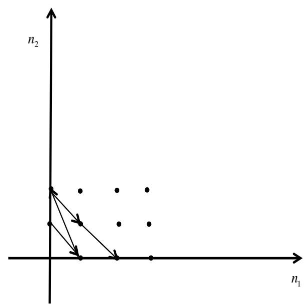

Fig. 1. Total degree lexicographic search in complexity space

Remark 4. Once a pair (S_1, S_2) has been computed through the above procedure, the (generically satisfied) sufficient conditions stated in Th. 5.8 of Rapisarda and Antoulas (2016) can be checked:

$$\begin{aligned} \text{im } Y'^{*} \cap \text{im } U'^{*} &= \{0\} \\ \text{im } S_1 \cap \text{im } S_2 &= \{0\} \\ [S_1 \; S_2] \cap \text{im } U'^{*} &= \{0\} \,, \end{aligned}$$

where Y' and U' are defined in (11). An unfalsified model for the data can be computed as follows.

Let $S_i = X_i'^{*} X_i$, $i = 1, 2$ be rank-revealing factorizations.; it is shown in Th. 5.8 of Rapisarda and Antoulas (2016) that there exist a left inverse $\left[X_1'^{*} \; X_2'^{*}\right]^{\dagger}$ of $\left[X_1'^{*} \; X_2'^{*}\right]$ and $F \in \mathbb{C}^{\mathtt{p} \times N}$ such that $\left[X_1'^{*} \; X_2'^{*}\right]^{\dagger} U'^{*} = 0_{N \times \mathtt{p}}$ and $F\left[Y'^{*} \; U'^{*}\right] = [0_{\mathtt{p} \times \mathtt{m}} \; I_{\mathtt{p}}]$. The i/s/o matrices are computed solving (13) in the least squares sense:

$$A := -\left[X_1'^{*} \; X_2'^{*}\right]^{\dagger} \left[M_1^{*} X_1'^{*} \; M_2^{*} X_2'^{*}\right] \,, \; B := \left[X_1'^{*} \; X_2'^{*}\right]^{\dagger} Y'^{*}$$

$$C := F\left(I_N - \left[X_1'^{*} \; X_2'^{*}\right] \left[X_1'^{*} \; X_2'^{*}\right]^{\dagger}\right) \left[M_1^{*} X_1'^{*} \; M_2^{*} X_2'^{*}\right]$$

$$D := F\left[X_1'^{*} \; X_2'^{*}\right] \left[X_1'^{*} \; X_2'^{*}\right]^{\dagger} Y'^{*} \,.$$

Remark 5. We have implemented the above procedure in `Mathematica`. Anecdotal evidence obtained dealing with

N larger than five or six suggests that it is difficult to apply it to larger scale problems, since the use of Gröbner bases is computationally rather intensive. Using heuristic affine rank minimization algorithms could be more widely applicable. The advantage of a Gröbner basis approach is that it computes a parametrization of *all* solutions (S_1, S_2) to (3) with a given complexity. It consequently opens up the possibility of exploring such parameter space, for example in order to computed unfalsified models with desired properties. □

5. EXAMPLE

The generating system is described by the SISO transfer function $G(s_1, s_2) = \frac{1}{s_1 s_2 - 1}$, considered in Ex. 2.2 of Antoulas et al. (2012). We choose the interpolation points

$$\left\{(\lambda_1^i, \lambda_2^i)\right\}_{i=1,\dots,4} = \left\{(4,3),(5,4),(9,\frac{1}{4}),(3,2)\right\}$$
$$\left\{(\mu_1^i, \mu_2^i)\right\}_{i=1,\dots,4} = \left\{(2,-\frac{1}{2}),(2,-\frac{3}{2}),(2,-1),(\frac{1}{2},-1)\right\} .$$

Such interpolation point correspond to the directions

$$\{\overline{w}_i\}_{i=1,\dots,4} = \left\{\begin{bmatrix}11\\1\end{bmatrix},\begin{bmatrix}19\\1\end{bmatrix},\begin{bmatrix}5\\4\end{bmatrix},\begin{bmatrix}5\\1\end{bmatrix}\right\}$$
$$\{\overline{w}_i'\}_{i=1,\dots,4} = \left\{\begin{bmatrix}-2\\-1\end{bmatrix},\begin{bmatrix}-4\\-1\end{bmatrix},\begin{bmatrix}-3\\-1\end{bmatrix},\begin{bmatrix}-3\\-2\end{bmatrix}\right\} .$$

From such data we compute the Loewner matrix:

$$\mathbb{L} = \begin{bmatrix} -13 & -21 & -13 & -7 \\ -15 & -23 & -21 & -9 \\ -14 & -22 & -17 & -8 \\ -25 & -41 & -22 & -13 \end{bmatrix} .$$

To parametrize all solutions to the 2D Sylvester equation (3), we first compute solutions $\overline{S}_i$, $i = 1,2$ to the 1D Sylvester equation (14) with $Q = \frac{1}{2}\mathbb{L}$:

$$\overline{S}_1 = \begin{bmatrix} -\frac{13}{12} & -\frac{3}{2} & -\frac{13}{22} & -\frac{7}{10} \\ -\frac{5}{4} & -\frac{23}{14} & -\frac{21}{22} & -\frac{9}{10} \\ -\frac{7}{6} & -\frac{11}{7} & -\frac{17}{22} & -\frac{4}{5} \\ -\frac{25}{9} & -\frac{41}{11} & -\frac{22}{19} & -\frac{13}{7} \end{bmatrix}$$
$$\overline{S}_2 = \begin{bmatrix} -\frac{13}{5} & -3 & 26 & -\frac{7}{3} \\ -5 & -\frac{23}{5} & \frac{42}{5} & -9 \\ -\frac{7}{2} & -\frac{11}{3} & \frac{34}{3} & -4 \\ -\frac{25}{4} & -\frac{41}{6} & \frac{44}{3} & -\frac{13}{2} \end{bmatrix} .$$

It can be verified that $\text{rank}(\overline{S}_1) = 3$ and $\text{rank}(\overline{S}_2) = 4$. If state directions would be computed by factorizing $\overline{S}_i$, $i = 1,2$, they would have dimensions 3 and 2 respectively, and the total complexity of the model would be 3+2=5. We show that there exist lower complexity unfalsified models.

Let $g_{ij} \in \mathbb{R}$, $i,j = 1,\dots,4$ be to-be-determined parameters; it is a matter of straightforward verification to check that the general solutions S_1', S_2' of (15) are (see Prop. 6):

$$S_1'(g_{ij}) = \begin{bmatrix} -\frac{5}{12}g_{11} & -\frac{g_{12}}{2} & \frac{g_{13}}{44} & -\frac{3}{10}g_{14} \\ -\frac{g_{21}}{4} & -\frac{5}{14}g_{22} & \frac{5g_{23}}{44} & -\frac{g_{24}}{10} \\ -\frac{g_{31}}{3} & -\frac{3}{7}g_{32} & \frac{3g_{33}}{44} & -\frac{g_{34}}{5} \\ -\frac{4}{9}g_{41} & -\frac{6}{11}g_{42} & \frac{3g_{43}}{38} & -\frac{2}{7}g_{44} \end{bmatrix}$$
$$S_2'(g_{ij}) = [g_{ij}]_{i,j=1,\dots,4} . \tag{19}$$

From Cor. 7 it follows that any pair $(S_1(g_{ij}), S_2(g_{ij}))$ which is a solution of the 2D Sylvester equation (3) can be written as $(\overline{S}_1 + S_1'(g_{ij}), \overline{S}_2 + S_2'(g_{ij}))$ for some $S_1'(g_{ij})$, $S_2'(g_{ij})$ as in (19).

We apply the algorithm in sect. 4. For $n = 1$, the only pairs $(\mathtt{n}_1, \mathtt{n}_2) \in \mathbb{N}\times\mathbb{N}$ such that $\mathtt{n}_1 + \mathtt{n}_2 = n$ are respectively $(1,0)$ and $(0,1)$. When considering $(1,0)$ as possible complexity, we compute the 2×2 minors of $S_1(g_{ij})$ and the 1×1 minors of $S_2(g_{ij})$, and the respective Gröbner bases $\mathcal{G}_1$ and $\mathcal{G}_2$; it can be verified that the Gröbner basis of $\mathcal{G}_1 \cup \mathcal{G}_2$ is $\{1\}$. An analogous situation occurs for $(\mathtt{n}_1, \mathtt{n}_2) = (0,1)$. From this analysis it follows that there do not exist unfalsified models with complexity $n = 1$.

For $n = 2$, $(2,0)$, $(0,2)$, and $(1,1)$ are the only possible pairs of state dimensions. In the first case, we compute the 3×3 minors of $S_1(g_{ij})$ and the 1×1 minors of $S_2(g_{ij})$, and the respective Gröbner bases $\mathcal{G}_1$ and $\mathcal{G}_2$. It can be verified that the Gröbner basis of $\mathcal{G}_1 \cup \mathcal{G}_2$ is $\{1\}$; hence no unfalsified model exists with $\mathtt{n}_1 = 2$, $\mathtt{n}_2 = 0$. Similar computations lead us to conclude that no unfalsified models of complexity $(0,2)$ exist. To verify the existence of models of complexity $(1,1)$, we compute the 2×2 minors of $S_1(g_{ij})$ and the 2×2 minors of $S_2(g_{ij})$; it can be verified that the Gröbner basis of the ideal generated by such polynomials in the parameters g_{ij} $i,j = 1,\dots,4$ is

$$\begin{aligned}&\{15 - 32g_{44} + 4g_{44}^2, 1934 + 21g_{43} - 228g_{44},\\ &83 + 21g_{42} - 33g_{44}, 67 + 28g_{41} - 36g_{44},\\ &-9 + 14g_{34} - 10g_{44}, 1060 + 21g_{33} - 132g_{44},\\ &11 + 6g_{32} - 6g_{44}, 13 + 14g_{31} - 12g_{44}, -79 + 14g_{24} - 10g_{44},\\ &1664 + 35g_{23} - 220g_{44}, 9 + 10g_{22} - 10g_{44}, -4 + 7g_{21} - 6g_{44},\\ &43 + 42g_{14} - 30g_{44}, 456 + 7g_{13} - 44g_{44}, 5 + 2g_{12} - 2g_{44},\\ &64 + 35g_{11} - 30g_{44}\} .\end{aligned}$$

The variety associated with the ideal generated by such polynomials consists of precisely two points, corresponding respectively to

$$S_2' = \begin{bmatrix} -\frac{7}{5} & -2 & -62 & -\frac{2}{3} \\ 1 & -\frac{2}{5} & -\frac{222}{5} & 6 \\ -\frac{1}{2} & -\frac{4}{3} & -\frac{142}{3} & 1 \\ -\frac{7}{4} & -\frac{19}{6} & -\frac{260}{3} & \frac{1}{2} \end{bmatrix} \tag{20}$$

and

$$S_2' = \begin{bmatrix} \frac{23}{5} & 5 & -18 & \frac{13}{3} \\ 7 & \frac{33}{5} & -\frac{2}{5} & 11 \\ \frac{11}{2} & \frac{17}{3} & -\frac{10}{3} & 6 \\ \frac{29}{4} & \frac{47}{6} & -\frac{32}{3} & \frac{15}{2} \end{bmatrix},$$

in the parametrization (19). The corresponding S_1' matrices are computed as in the first equation in (19). Taking S_2' equal to (20) and computing the corresponding S_1', a solution pair (S_1, S_2) to (3) is

$$S_1 = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & -2 & -\frac{1}{2} \\ -\frac{3}{2} & -\frac{3}{2} & -6 & -\frac{3}{2} \\ -1 & -1 & -4 & -1 \\ -2 & -2 & -8 & -2 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} -4 & -5 & -36 & -3 \\ -4 & -5 & -36 & -3 \\ -4 & -5 & -36 & -3 \\ -8 & -10 & -72 & -6 \end{bmatrix}$$

Recall that S_1 and S_2 have rank 1 by construction.

It is a matter of straightforward verification that the sufficient conditions of Rem. 4 are satisfied. An SVD can be performed on S_1 and S_2 to obtain the factorizations

$$S_1 = \begin{bmatrix} 0.631 \\ 1.892 \\ 1.262 \\ 2.523 \end{bmatrix} \underbrace{\begin{bmatrix} -0.793 & -0.793 & -3.171 & -0.793 \end{bmatrix}}_{=:X_1}$$

$$S_2 = \begin{bmatrix} 3.724 \\ 3.724 \\ 3.724 \\ 7.448 \end{bmatrix} \underbrace{\begin{bmatrix} -1.0742 & -1.343 & -9.667 & -0.806 \end{bmatrix}}_{=:X_2}.$$

The matrices X_1 and X_2 define the first, respectively second component of the state directions associated with the given external trajectories.

The matrices A, B, C and D of a Roesser model can be computed through the formulas at the end of Rem. 4, or solving in the least squares sense the system of equations

$$\begin{bmatrix} X_1\Lambda_1 \\ X_2\Lambda_2 \\ Y \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ C_1 & C_2 & D \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ U \end{bmatrix}. \tag{21}$$

It can be verified that

$$A = \begin{bmatrix} 0 & 2.95164 \\ 0.338794 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ -0.268543 \end{bmatrix}$$

$$C = \begin{bmatrix} -1.2616 & 0 \end{bmatrix}$$

$$D = 0,$$

solve such system of equations. The transfer function associated with this model is

$$C \begin{bmatrix} s_1 - A_{11} & -A_{12} \\ -A_{21} & s_2 - A_{22} \end{bmatrix}^{-1} B = \frac{1}{s_1 s_2 - 1},$$

the one used to generate the data. □

6. CONCLUSION

We presented a Gröbner basis approach to the solution of a 2D matrix Sylvester equation arising in 2D identification of Roesser models from frequency data.

REFERENCES

Antoulas, A., Ionita, A., and Lefteriu, S. (2012). On two-variable rational interpolation. *Lin. Alg. Appl.*, 436(8), 2889 – 2915.

Antoulas, A. and Rapisarda, P. (2015). Bilinear differential forms and the Loewner framework for rational interpolation. In M. Belur, M. Çamlibel, and P. Rapisarda (eds.), *Mathematical Control Theory II: Lect. Notes Control, vol. 462*, chapter 2. Springer-Verlag.

Farah, M., Mercere, G., Ouvrard, R., Poinot, T., and Ramos, J. (2014). Identification of 2D Roesser models by using linear fractional transformations. In *Proc. 2014 European Control Conference (ECC)*, 382–387.

Kaczorek, T. (1985). *Two-dimensional linear systems.* Berlin: Springer.

Kaneko, O. and Rapisarda, P. (2003). Recursive exact H_∞-identification from impulse-response measurements. *Syst. Contr. Lett.*, 49(5), 323–334.

Kaneko, O. and Rapisarda, P. (2007). On the Takagi interpolation problem. *Lin. Alg. Appl.*, 425(2), 453–470.

Peeters, R. and Rapisarda, P. (2006). Solution of polynomial Lyapunov and Sylvester equations. In B. Hanzon and M. Hazewinkel (eds.), *Constructive algebra and systems theory*, 151–167. KNAW.

Ramos, J. (1994). A subspace algorithm for identifying 2-D separable in denominator filters. *IEEE Trans. Circ. Syst. II: Analog and Digital Signal Processing*, 41(1), 63–67.

Rapisarda, P. and Antoulas, A. (2015). A duality perspective on Loewner interpolation. In *Proc. 56th IEEE Conference on Decision and Control (CDC)*, 2096–2100.

Rapisarda, P. and Antoulas, A. (2016). State-space modelling of two-dimensional vector-exponential trajectories. *SIAM J. Contr. Opt.*, 54(5), 2734–2753.

Rapisarda, P. and Willems, J. (1997). The subspace Nevanlinna interpolation problem and the most powerful unfalsified model. *Systems & control letters*, 32(5), 291–300.

Roesser, R. (1975). A discrete state-space model for linear image processing. *IEEE Trans. Aut. Contr.*, 20(1), 1–10.

Zerz, E. (2004). Multidimensional behaviours: an algebraic approach to control theory for PDE. *Int. J. Contr.*, 77(9), 812–820.