Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem

22nd European Control Conference (ECC) - Stockholm, Sweden

Sibren Lagauw*, Lukas Vanpoucke, Bart De Moor
{sibren.lagauw, christof.vermeersch, bart.demoor}@esat.kuleuven.be

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS),
Department of Electrical Engineering (ESAT), KU Leuven, Belgium

June 28, 2024

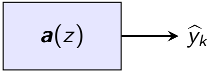
fwo Fonds Wetenschappelijk Onderzoek Vlaanderen
Opening new horizons

erc

Table of contents

Table of contents

Autonomous LTI model



- LTI dynamics of model-compliant data $\widehat{\boldsymbol{y}} \in \mathbb{R}^N$:

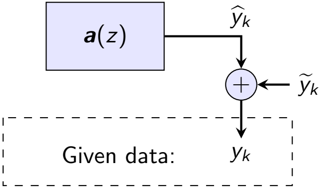
$$\widehat{y}_{k+n} + a_1 \widehat{y}_{k+n-1} + \cdots + a_n \widehat{y}_k = 0, \; \forall k{=}0,\ldots,N{-}n{-}1$$

- Kernel representation of *behavior*

$$\underbrace{\begin{bmatrix} a_n & \ldots & \ldots & a_1 & 1 & 0 & \ldots & 0 \\ 0 & a_n & \ldots & \ldots & a_1 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n & \ldots & \ldots & a_1 & 1 \end{bmatrix}}_{\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}} \begin{bmatrix} \widehat{y}_0 \\ \widehat{y}_1 \\ \vdots \\ \widehat{y}_{N-1} \end{bmatrix} = \boldsymbol{0}$$

- n unknown model parameters $\boldsymbol{a} \in \mathbb{R}^n$

Autonomous LTI model



Least-squares realization:

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y}-\widehat{\boldsymbol{y}}\|_2^2,$$

$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{N-n}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

- LTI dynamics of model-compliant data $\widehat{\boldsymbol{y}} \in \mathbb{R}^N$:

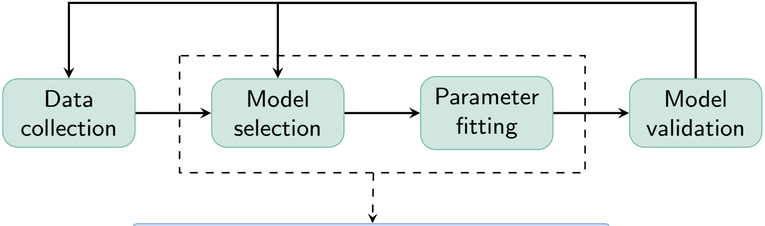
$$\widehat{y}_{k+n} + a_1\widehat{y}_{k+n-1} + \cdots + a_n\widehat{y}_k = 0, \ \forall k{=}0,\ldots,N{-}n{-}1$$

- Kernel representation of *behavior*

$$\underbrace{\begin{bmatrix} a_n & \ldots & \ldots & a_1 & 1 & 0 & \ldots & 0 \\ 0 & a_n & \ldots & \ldots & a_1 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n & \ldots & \ldots & a_1 & 1 \end{bmatrix}}_{\boldsymbol{T}^{\boldsymbol{a}}_{N-n}} \begin{bmatrix} \widehat{y}_0 \\ \widehat{y}_1 \\ \vdots \\ \widehat{y}_{N-1} \end{bmatrix} = \boldsymbol{0}$$

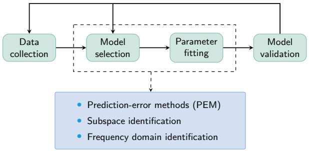
- n unknown model parameters $\boldsymbol{a} \in \mathbb{R}^n$

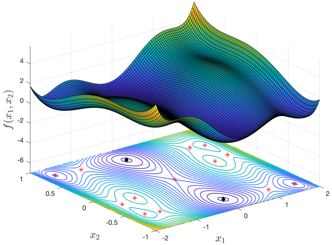
System identification for autonomous LTI models



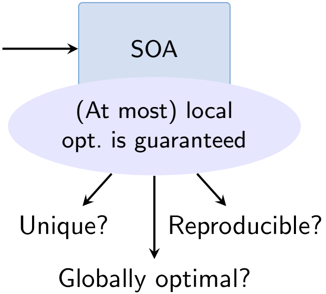
- Prediction-error methods (PEM)
- Subspace identification
- Frequency domain identification

Nonlinear optimization problem(s)

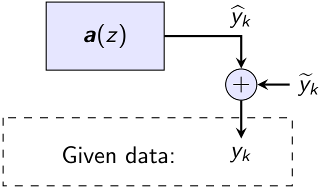




Non-convex
opt. problem

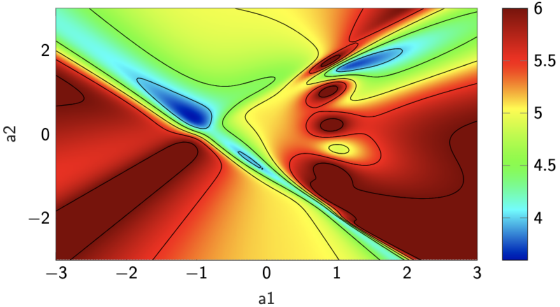


Example



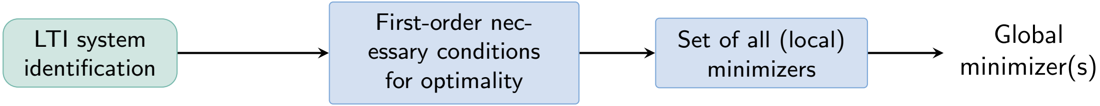
Least-squares realization:

$$\begin{aligned}\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad & \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2, \\ \text{s.t.} \quad & \boldsymbol{T}_{N-n}^{\boldsymbol{a}}\widehat{\boldsymbol{y}} = \boldsymbol{0}.\end{aligned}$$



(Objective function $\|\widetilde{\boldsymbol{y}}\|_2^2$ for random data $\boldsymbol{y} \in \mathbb{R}^{16}$ and $n = 2$)

Global optimality



Global optimality

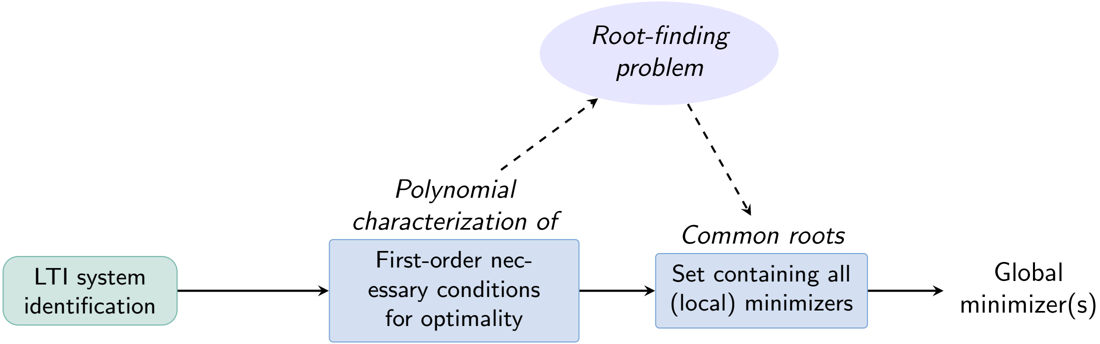


Table of contents

Least-squares realization

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2,$$

$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{N-n}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

$$\downarrow$$

$$\mathcal{L}(\boldsymbol{a}, \widehat{\boldsymbol{y}}, \boldsymbol{l}) = \tfrac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2 + \boldsymbol{l}^{\mathsf{T}}\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\widehat{\boldsymbol{y}},$$

$$\downarrow$$

$$\begin{aligned}
\partial\mathcal{L}/\partial\widehat{\boldsymbol{y}} &= \widehat{\boldsymbol{y}} - \boldsymbol{y} + \left(\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\right)^{\mathsf{T}}\boldsymbol{l} = \boldsymbol{0},\\
\partial\mathcal{L}/\partial\boldsymbol{a} &= \left(\widehat{\boldsymbol{Y}}_{N-n}\right)^{\mathsf{T}}\boldsymbol{l} - \boldsymbol{e}\lambda = \boldsymbol{0},\\
\partial\mathcal{L}/\partial\boldsymbol{l} &= \boldsymbol{T}^{\boldsymbol{a}}_{N-n}\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{Y}}_{N-n}\boldsymbol{a} = \boldsymbol{0}
\end{aligned}$$

Least-squares realization

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y}-\widehat{\boldsymbol{y}}\|_2^2,$$
$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

↓

$$\mathcal{L}(\boldsymbol{a},\widehat{\boldsymbol{y}},\boldsymbol{l}) = \tfrac{1}{2}\|\boldsymbol{y}-\widehat{\boldsymbol{y}}\|_2^2 + \boldsymbol{l}^{\mathsf{T}}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}},$$

↓

$$\partial\mathcal{L}/\partial\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{y}} - \boldsymbol{y} + (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}}\boldsymbol{l} = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{a} = (\widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}})^{\mathsf{T}}\boldsymbol{l} - \boldsymbol{e}\lambda = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{l} = \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}}\boldsymbol{a} = \boldsymbol{0}$$

Theorem (De Moor, 2020)

The minimal norm misfit $\widetilde{\boldsymbol{y}} = \boldsymbol{y} - \widehat{\boldsymbol{y}} \in \mathbb{R}^N$ corresponds to the orth. projection of $\boldsymbol{y}$ onto $\operatorname{row}\left(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\right)$,

$$\widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}}\left(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}}\right)^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\boldsymbol{y}.$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Least-squares realization

$$\min_{\boldsymbol{a},\widehat{\boldsymbol{y}}} \quad \frac{1}{2}\|\widetilde{\boldsymbol{y}}\|_2^2 = \frac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2,$$
$$\text{s.t.} \quad \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \boldsymbol{0}.$$

$$\downarrow$$

$$\mathcal{L}(\boldsymbol{a}, \widehat{\boldsymbol{y}}, \boldsymbol{l}) = \tfrac{1}{2}\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2 + \boldsymbol{l}^{\mathrm{T}}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}},$$

$$\downarrow$$

$$\partial\mathcal{L}/\partial\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{y}} - \boldsymbol{y} + (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}}\boldsymbol{l} = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{a} = (\widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}})^{\mathrm{T}}\boldsymbol{l} - \boldsymbol{e}\lambda = \boldsymbol{0},$$
$$\partial\mathcal{L}/\partial\boldsymbol{l} = \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}}\boldsymbol{a} = \boldsymbol{0}$$

Theorem (De Moor, 2020)

The minimal norm misfit $\widetilde{\boldsymbol{y}} = \boldsymbol{y} - \widehat{\boldsymbol{y}} \in \mathbb{R}^N$ *corresponds to the orth. projection of* $\boldsymbol{y}$ *onto* $\mathrm{row}\left(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\right)$,

$$\widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}}\left(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}}\right)^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\boldsymbol{y}.$$

Theorem (Lagauw, Vanpoucke, et al., 2024)

If $\boldsymbol{a}$ *is a (local) minimizer, then* $\exists\, \boldsymbol{g} \in \mathbb{R}^{N-2n}$ *s.t.,*

$$\widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathrm{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}})^{\mathrm{T}}\boldsymbol{g},$$

where $\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}} \in \mathbb{R}^{(N-2n)\times(N-n)}$ *is a banded Toeplitz matrix defined similarly to the matrix* $\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}} \in \mathbb{R}^{(N-n)\times N}$.

System of multivariate polynomial equations

- If $\boldsymbol{a}$ is a (local) minimizer, then $\exists\, \boldsymbol{g} \in \mathbb{R}^{N-2n}$ s.t.,

$$\widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}})^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} = (\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g}.$$
$$\iff (\boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}})^{-1}\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - (\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0},$$
$$\iff \boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0}.$$

- Define the algebraic variety,

$$\mathcal{V}_{\mathbb{R}} = \left\{(\boldsymbol{a}, \boldsymbol{g}) \in \mathbb{R}^{N-n} \;:\; \boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0}\right\}.$$

$\longrightarrow$ $\mathcal{V}_{\mathbb{R}}$ contains all minimizers $\boldsymbol{a}$ of the identification problem ✓

- square system of $N-n$ polynomial equations
- degree of the polynomials is at most *quartic*

Numerical example[1] I

- Find globally optimal first-order ($n = 1$) autonomous LTI realization
- Given output data $\boldsymbol{y} = [4, 3, 2, 1]^{\mathsf{T}}$ ($N = 4$)

$$\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0},$$

$$\iff \begin{cases} 0 = 4a_1 - 2a_1g_1 - a_1^2g_2 - a_1^3g_1 + 3, \\ 0 = 3a_1 - g_1 - 2a_1g_2 - 2a_1^2g_1 - a_1^3g_2 + 2, \\ 0 = 2a_1 - g_2 - a_1g_1 - 2a_1^2g_2 + 1. \end{cases}$$

[1] Toy problem from (De Moor, 2020)

Numerical example[1] I

- Find globally optimal first-order ($n = 1$) autonomous LTI realization
- Given output data $\boldsymbol{y} = [4, 3, 2, 1]^\mathsf{T}$ ($N = 4$)

$$\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^\mathsf{T}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}})^\mathsf{T}\boldsymbol{g} = \boldsymbol{0},$$

$$\Longleftrightarrow \begin{cases} 0 = 4a_1 - 2a_1g_1 - a_1^2g_2 - a_1^3g_1 + 3, \\ 0 = 3a_1 - g_1 - 2a_1g_2 - 2a_1^2g_1 - a_1^3g_2 + 2, \\ 0 = 2a_1 - g_2 - a_1g_1 - 2a_1^2g_2 + 1. \end{cases}$$

- 7 common-roots, one of which in $\mathcal{V}_\mathbb{R}$

$\lVert\tilde{\boldsymbol{y}}\rVert_2^2$	a_1	g_1	g_2
0.1486	-0.6764	-0.2525	-0.2734
/	$-0.1589 \mp 0.808j$	$1.3577 \pm 3.8194j$	$1.8359 \pm 3.3491j$
/	$0.4209 \pm 0.6233j$	$3.0425 \mp 2.9959j$	$-0.0785 \pm 1.2013j$
/	$1.3261 \pm 2.0058j$	$-0.2739 \mp 0.6279j$	$0.3793 \mp 0.3849j$

[1] Toy problem from (De Moor, 2020)

Multiparameter eigenvalue problem

$$\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\boldsymbol{g} = \boldsymbol{0}. \qquad (1)$$

- Rewrite (1) as a *cubic n-parameter eigenvalue problem* in the parameters $a_1, \ldots, a_n$,

$$\underbrace{\begin{bmatrix}\boldsymbol{T}^{\boldsymbol{a}}_{N-n}\boldsymbol{y} & \boldsymbol{T}^{\boldsymbol{a}}_{N-n}(\boldsymbol{T}^{\boldsymbol{a}}_{N-n})^{\mathsf{T}}(\boldsymbol{T}^{\boldsymbol{a}}_{N-2n})^{\mathsf{T}}\end{bmatrix}}_{\mathcal{M}(\boldsymbol{a})}\begin{pmatrix}-1\\ \boldsymbol{g}\end{pmatrix} = \boldsymbol{0}.$$

- $\mathcal{M}(\boldsymbol{a}) = \sum_{\{\boldsymbol{\alpha}\}} \boldsymbol{M}_{\boldsymbol{\alpha}}\boldsymbol{a}^{\boldsymbol{\alpha}}$, is a matrix polynomial in the monomials $\boldsymbol{a}^{\boldsymbol{\alpha}} = a_1^{\alpha_1}\ldots a_n^{\alpha_n}$, with coefficient matrices $\boldsymbol{M}_{\boldsymbol{\alpha}} \in \mathbb{R}^{(N-n)\times(N-2n+1)}$.
- The values $\boldsymbol{a} \in \mathbb{C}^n$ for which $\mathcal{M}(\boldsymbol{a})$ becomes rank-deficient, such that there exists a vector $\boldsymbol{g} \in \mathbb{C}^{N-2n}$ for which these equations are satisfied, are the *affine eigentuples*[2] of this MEP.

[2] (De Cock and De Moor, 2021)

Numerical Example I (continued)

- Given output data $\boldsymbol{y} = [4, 3, 2, 1]^\mathsf{T}$ $(N = 4)$

$$\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\boldsymbol{y} - \boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^\mathsf{T}(\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}})^\mathsf{T}\boldsymbol{g} = \boldsymbol{0},$$

$$\iff \begin{cases} 0 = 4a_1 - 2a_1g_1 - a_1^2g_2 - a_1^3g_1 + 3, \\ 0 = 3a_1 - g_1 - 2a_1g_2 - 2a_1^2g_1 - a_1^3g_2 + 2, \\ 0 = 2a_1 - g_2 - a_1g_1 - 2a_1^2g_2 + 1. \end{cases}$$

- Cubic 1-parameter eigenvalue problem:

$$\left[\underbrace{\begin{pmatrix} 3 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}}_{\boldsymbol{M}_0} + \underbrace{\begin{pmatrix} 4 & -2 & 0 \\ 3 & 0 & -2 \\ 2 & -1 & 0 \end{pmatrix}}_{\boldsymbol{M}_1} a_1 + \underbrace{\begin{pmatrix} 0 & 0 & -1 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}}_{\boldsymbol{M}_2} a_1^2 + \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}}_{\boldsymbol{M}_3} a_1^3\right]\begin{pmatrix} 1 \\ g_1 \\ g_2 \end{pmatrix} = \boldsymbol{0}.$$

Numerical Example II

- Find globally optimal second-order ($n = 2$) autonomous LTI realization
- Given output data[3] $\boldsymbol{y}$ ($N = 16$),

$$\boldsymbol{y} = \boldsymbol{y}_{3\text{rd}} + 0.05 * \texttt{randn(N,1)},$$

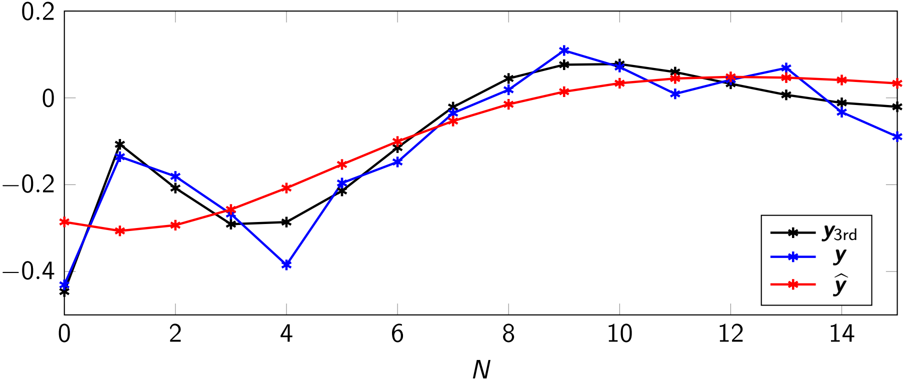
where $\boldsymbol{y}_{3\text{rd}}$ is generated by a third-order autonomous LTI model with poles $(0.2, 0.7 \pm 0.4j)$

- 739 affine common-roots, 9 of which are real-valued.

$\|\widetilde{\boldsymbol{y}}\|_2^2$	a_1	a_2	p_1	p_2
0.1327	−1.6255	0.7167	$0.8127 + 0.2369j$	$0.8127 - 0.2369j$
0.1514	−0.0752	−0.5850	0.8033	−0.7282
0.1606	−14.076	10.433	13.291	0.7849
0.5386	−0.7127	1.8381	$0.3564 + 1.3081j$	$0.3564 - 1.3081j$
⋮	⋮	⋮	⋮	⋮
0.5492	−1.3053	1.0564	$0.6527 + 0.7940j$	$0.6527 - 0.7940j$

[3] The considered instance of $\boldsymbol{y}$ has $\|\boldsymbol{y}\|_2^2 = 0.5509$.

Numerical Example II



Properties of the MEP

- One alternative MEP $\mathcal{M}_1$ described in the literature[4]
- Computational complexity of solving an MEP depends on:
  - the highest degree of the parameters
  - the number of parameters
  - the size of the coefficient matrices
- Size of the coefficient matrices the MEPs for several (N, n):

(N, n)	size($\mathcal{M}_1$)	size($\mathcal{M}_{\text{new}}$)
$(4, 1)$	7×7	3×3
$(16, 6)$	76×71	10×5
$(50, 8)$	386×379	42×35
$(200, 15)$	2975×2961	185×171

[4] (De Moor, 2020)

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Properties of the MEP

- One alternative MEP $\mathcal{M}_1$ described in the literature[4]
- Computational complexity of solving an MEP depends on:
  - the highest degree of the parameters
  - the number of parameters
  - the size of the coefficient matrices
- Size of the coefficient matrices the MEPs for several (N, n):

(N, n)	size$(\mathcal{M}_1)$	size$(\mathcal{M}_{\text{new}})$
$(4, 1)$	7×7	3×3
$(16, 6)$	76×71	10×5
$(50, 8)$	386×379	42×35
$(200, 15)$	2975×2961	185×171

$$\longrightarrow \quad \text{size}(\mathcal{M}_{\text{new}}) \approx \frac{\text{size}(\mathcal{M}_1)}{(n+1)}$$

[4] (De Moor, 2020)

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Table of contents

H_2-norm model reduction of SISO LTI models

Minimize H_2-norm of approximation error $E(z) = H(z) - \widehat{H}(z)$:

$$\widehat{H}(z) \in \underset{\widehat{H}(z) \in \mathcal{M}}{\operatorname{argmin}} \; J^2, \tag{2}$$

where,

$$J^2 = \|E(z)\|_{H_2}^2 = \sum_{k=0}^{\infty} \left(h_k - \widehat{h}_k \right)^2,$$

with $\{h_k\}_{k=0,\ldots,\infty}$ and $\{\widehat{h}_k\}_{k=0,\ldots,\infty}$ the impulse responses of $H(z)$, $\widehat{H}(z)$, respectively, and,

$$\widehat{H}(z) = \frac{\widehat{b}(z)}{\widehat{a}(z)} = \frac{\widehat{b}_{m-1} z^{m-1} + \cdots + \widehat{b}_1 z + \widehat{b}_0}{z^m + \widehat{a}_{m-1} z^{m-1} + \cdots + \widehat{a}_1 z + \widehat{a}_0},$$

Model reduction as a limiting case ($N \to \infty$)

- Take $\boldsymbol{y}$ the impulse response of a stable mth order SISO model, then, for a given $n < m$, the LS realization problem finds the optimal $\widehat{\boldsymbol{y}}$ for which,
  - $\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|_2^2$ is minimal and,
  - $\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}}\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}}\boldsymbol{a} = 0$.

- by Kronecker's Theorem[5] the rank-deficiency of the Hankel matrix $\widehat{\boldsymbol{Y}}_{\boldsymbol{N-n}} \in \mathbb{R}^{\infty\times(n+1)}$ implies that $\widehat{\boldsymbol{y}}$ is the impulse response of an nth order SISO model.

$\longrightarrow$ for $\boldsymbol{y} \in \mathbb{R}^N$ with $N \to \infty$,

autonomous LTI LS realization $\approx$ SISO H_2-norm model reduction

[5] (Kronecker, 1881)

Interpolatory conditions for optimality

Theorem (Meier and Luenberger, 1967)

Given a stable SISO model $H(z) \in \mathcal{M}$ *of order* n*, let* $\widehat{H}(z)$ *of order* m $(m<n)$ *be a stationary point of the* H_2*-norm model reduction problem in* (2)*. Then for all poles* p_i *of* $\widehat{H}(z)$*,*

$$H(1/p_i)^{(j)} = \widehat{H}(1/p_i)^{(j)}, \quad j = 0, \ldots, d_i,$$

where d_i *is the multiplicity of the pole* p_i *and the superscript* j *denotes the* j*th derivative wrt.* z*.*

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Interpolatory conditions for optimality

Theorem (Meier and Luenberger, 1967)

Given a stable SISO model $H(z) \in \mathcal{M}$ of order n, let $\widehat{H}(z)$ of order m ($m<n$) be a stationary point of the H_2-norm model reduction problem in (2). Then for all poles p_i of $\widehat{H}(z)$,

$$H(1/p_i)^{(j)} = \widehat{H}(1/p_i)^{(j)}, \quad j = 0, \ldots, d_i,$$

where d_i is the multiplicity of the pole p_i and the superscript j denotes the jth derivative wrt. z.

Theorem (Regalia, 1995)

Given a stable SISO model $H(z) \in \mathcal{M}$ of order n, let $\widehat{H}(z)$ of order m with $m<n$ be a stationary point of the model reduction problem (2). Then for all $z \in \mathbb{C}$:

$$H(z) - \widehat{H}(z) = \frac{b(z)}{a(z)} - \frac{\widehat{b}(z)}{\widehat{a}(z)} = [z^m \widehat{a}(1/z)]^2 \frac{G(z)}{\widehat{a}(z)},$$

with $G(z)$ the z-transform of some real-valued, stable and causal signal.

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Intermezzo: globally optimal H_2-norm SISO MOR

Theorem (Lagauw, Agudelo, et al., 2023)

Let f_k be the coefficient corresponding to s^k in the polynomial,

$$l(s) = b(s)\widehat{a}(s) - a(s)\widehat{b}(s) - [\widehat{a}(-s)]^2 \widetilde{G}(s),$$

and define the algebraic variety,

$$\mathcal{V}_{\mathbb{R}} = \left\{ (\widehat{\boldsymbol{a}}, \widehat{\boldsymbol{b}}, \boldsymbol{g}) \in \mathbb{R}^{m+n} \ : \ f_k(\widehat{\boldsymbol{a}}, \widehat{\boldsymbol{b}}, \boldsymbol{g}) = 0, \quad \forall k = 0, \ldots, m+n-1 \right\}.$$

If $\mathcal{V}'_{\mathbb{R}} \subseteq \mathcal{V}_{\mathbb{R}}$ contains the $(\widehat{\boldsymbol{a}}, \widehat{\boldsymbol{b}}, \boldsymbol{g}) \in \mathcal{V}_{\mathbb{R}}$ for which $\widehat{H}(s)$ is stable, then $\mathcal{V}'_{\mathbb{R}}$ contains all minimizers $\widehat{H}(s)$.

- square system of $m+n$ polynomial equations
- degree of the polynomials is at most *cubic*

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Model reduction as a limiting case ($N \to \infty$)

Theorem (Regalia, 1995)

Given a stable SISO model $H(z) \in \mathcal{M}$ of order n, let $\widehat{H}(z)$ of order m with $m<n$ be a stationary point of the model reduction problem (2). Then for all $z \in \mathbb{C}$:

$$H(z) - \widehat{H}(z) = \frac{b(z)}{a(z)} - \frac{\widehat{b}(z)}{\widehat{a}(z)} = [z^m \widehat{a}(1/z)]^2 \frac{G(z)}{\widehat{a}(z)},$$

with $G(z)$ the z-transform of some real-valued, stable and causal signal.

Theorem (Lagauw, Vanpoucke, et al., 2024)

If $\boldsymbol{a}$ is a (local) minimizer, then $\exists\, \boldsymbol{g} \in \mathbb{R}^{N-2n}$ s.t.,

$$\boldsymbol{y} - \widehat{\boldsymbol{y}} = \widetilde{\boldsymbol{y}} = (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}})^{\mathsf{T}} (\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}})^{\mathsf{T}} \boldsymbol{g},$$

where $\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-2n}} \in \mathbb{R}^{(N-2n)\times(N-n)}$ is a banded Toeplitz matrix defined similarly to the matrix $\boldsymbol{T}^{\boldsymbol{a}}_{\boldsymbol{N-n}} \in \mathbb{R}^{(N-n)\times N}$.

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Table of contents

Summary & future work

- Globally optimal LS realization of autonomous LTI models is an MEP
- LS realization of autonomous LTI models $\leftrightarrow$ SISO H_2-norm model reduction

Future work:

- Proof zero-dimensionality of the (affine) solution set
- Particular interest in the *real-valued* common-roots
- Exploit structure of MEPs in computations

S. Lagauw, L. Vanpoucke, et al. (June 2024). "Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem". In: *Proc. of the 22nd European Control Conference (ECC)*. Stockholm, Sweden, pp. 3439–3444

S. Lagauw, O. M. Agudelo, et al. (2023). "Globally Optimal SISO H_2-Norm Model Reduction Using Walsh's Theorem". In: *IEEE Control Systems Letters* 7, pp. 1670–1675

Questions?

The work of Sibren Lagauw was supported by the
FWO Fundamental Research PhD Fellowship under Grant 11K5623N.

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

References I

- Agudelo, O. M., C. Vermeersch, and B. De Moor (2021). "Globally Optimal H_2-Norm Model Reduction: A Numerical Linear Algebra Approach". In: *IFAC PapersOnLine (Special Issue: Proceedings of MTNS 2020)* 54.9, pp. 564–571.
- Collins, E.G., D.J. Phillips, and D.C. Hyland (1991). "Robust decentralized control laws for the ACES structure". In: *IEEE Control Systems Magazine* 11.3, pp. 62–70.
- De Cock, K. and B. De Moor (2021). "Multiparameter Eigenvalue Problems and Shift-Invariance". In: *IFAC-PapersOnLine (Special Issue: Proceedings of MTNS 2020)* 54.9, pp. 159–165.
- De Moor, B. (2020). "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2, pp. 163–207.
- Kronecker, L. (1881). "Zur Theorie der Elimination einer Variablen aus zwei algebraischen Gleichungen". In: *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin)*. Reprinted as pp. 113–192 in Mathematische Werke, vol. 2, B.G. Teubner, Leipzig, 1897., pp. 535–600.
- Lagauw, S., O. M. Agudelo, and B. De Moor (2023). "Globally Optimal SISO H_2-Norm Model Reduction Using Walsh's Theorem". In: *IEEE Control Systems Letters* 7, pp. 1670–1675.
- Lagauw, S., L. Vanpoucke, and B. De Moor (June 2024). "Exact Characterization of the Global Optima of Least Squares Realization of Autonomous LTI Models as a Multiparameter Eigenvalue Problem". In: *Proc. of the 22nd European Control Conference (ECC)*. Stockholm, Sweden, pp. 3439–3444.
- Meier, L. and D.G. Luenberger (Oct. 1967). "Approximation of Linear Constant Systems". In: *IEEE Transactions on Automatic Control* 12.5, pp. 585–588.
- Regalia, Phillip A. (1995). *Adaptive IIR Filtering in Signal Processing and Control*. Marcel Dekker Inc., New York, NY.

Numerical example (MOR) I

- Consider the third order model ($n = 3$) used in Agudelo et al., 2021:

$$H(s) = \frac{s^2 + 9s - 10}{s^3 + 12s^2 + 49s + 78}.$$

- For $m = 1$, the polynomial $l(s)$ from (**??**) is given as,

$$
\begin{aligned}
l(s) = &\underbrace{(1 - g_1 - \widehat{b}_0)}_{f_3} s^3 + \underbrace{(\widehat{a}_0 - 12\widehat{b}_0 - g_0 + 2\widehat{a}_0 g_1 + 9)}_{f_2} s^2 \\
&+ \underbrace{(9\widehat{a}_0 - 49\widehat{b}_0 + 2\widehat{a}_0 g_0 - \widehat{a}_0^2 g_1 - 10)}_{f_1} s \\
&+ \underbrace{(-g_0 \widehat{a}_0^2 - 10\widehat{a}_0 - 78\widehat{b}_0)}_{f_0} 1.
\end{aligned}
$$

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Numerical example (MOR) I

Strategy: find all $(\widehat{\boldsymbol{a}}, \widehat{\boldsymbol{b}}, \boldsymbol{g})$ for which $I(s) = 0, \ \forall s \in \mathbb{C}$"

$$\iff \begin{cases} 0 = f_3 = 1 - g_1 - \widehat{b}_0, \\ 0 = f_2 = \widehat{a}_0 - 12\widehat{b}_0 - g_0 + 2\widehat{a}_0 g_1 + 9, \\ 0 = f_1 = 9\widehat{a}_0 - 49\widehat{b}_0 + 2\widehat{a}_0 g_0 - \widehat{a}_0^2 g_1 - 10, \\ 0 = f_0 = -g_0 \widehat{a}_0^2 - 10\widehat{a}_0 - 78\widehat{b}_0. \end{cases} \tag{3}$$

- Common roots of (3) contain all stationary points $\widehat{H}(s)$ of (2):

J	$\widehat{a}_0$	$\widehat{b}_0$	g_0	g_1	stable
8.9403	$-4.1639 + 0.9026j$	$24.930 - 6.5393j$	$-106.84 - 18.287j$	$-23.930 + 6.5393j$	✗
8.9403	$-4.1639 - 0.9026j$	$24.930 + 6.5393j$	$-106.84 + 18.287j$	$-23.930 - 6.5393j$	✗
0.3982	0.2671	−0.0437	10.349	1.0437	✓
0.2784	0.6914	9.6796	1.2799	−2.0986	✓
0.5232	−16.618	1.9264	0.0576	−0.9264	✗

Numerical example (MOR) II

We search for the globally optimal 4th order reduced model ($m = 4$) of the state-space model[6] ($n = 17$).

- System of 21 polynomial equations with $d_{\text{max}} = 3$
- $\mathcal{V}_{\mathbb{R}}$ contains 290 tuples, 69 remain in $\mathcal{V}'_{\mathbb{R}}$
- Four best-performing stationary points $\widehat{H}(s)$:

	J	$p_{1,2}$	$p_{3,4}$
*	9.14×10^{-3}	$-0.032 \pm 78.54j$	$-0.111 \pm 15.43j$
l_1	9.22×10^{-3}	$-0.032 \pm 78.54j$	$-5.713 \pm 52.57j$
l_2	1.03×10^{-2}	$-0.032 \pm 78.54j$	$-0.023 \pm 3.842j$
l_3	1.09×10^{-2}	$-0.032 \pm 78.54j$	$-4.663 \pm 15.88j$

[6] The model describes the interaction between a torque activator and a torsional rate sensor for the ACES structure Collins et al., 1991.

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

Numerical example (MOR) II