STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Subspace methods for the Block Macaulay Matrix Framework

42nd Benelux Meeting on Systems and Control, Elspeet, The Netherlands 2023
Sarthak De, Christof Vermeersch and Bart De Moor, *Fellow, IEEE & IFAC & SIAM*
sarthak.de@esat.kuleuven.be
March 22, 2023

Outline

System Identification

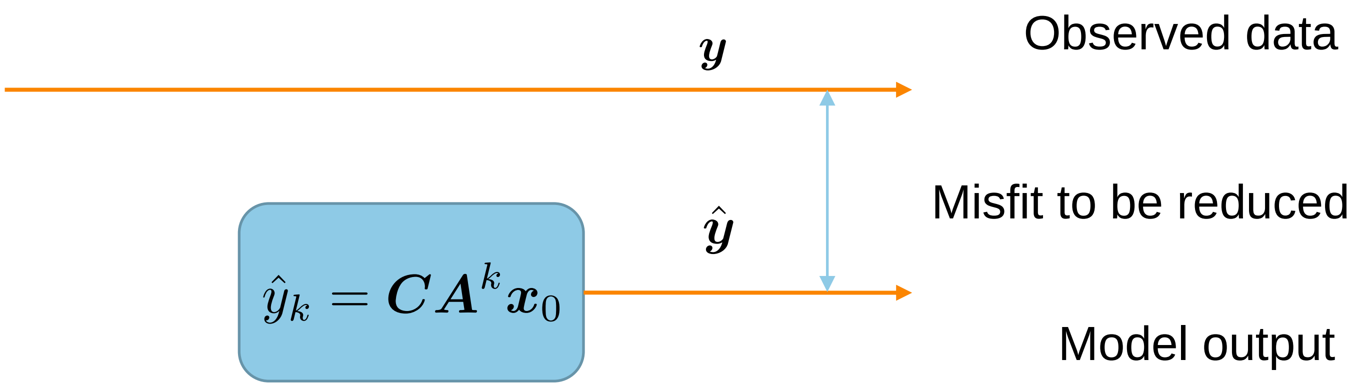
Rectangular Multiparameter Eigenvalue Problems

Subspace Methods

Conclusion and Future Work

System identification

We want to find a LTI autonomous model for observed data



Least-squares realization problem

Minimize the misfit

$$\begin{aligned} \text{min} \quad & \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|_2^2 \\ \text{subject to} \quad & \boldsymbol{T}_\alpha \hat{\boldsymbol{y}} = \boldsymbol{0} \end{aligned}$$

Toeplitz matrix

First order optimality conditions yield the multiparameter eigenvalue problem

$$\left(\boldsymbol{A}_{00} + \boldsymbol{A}_{10}\alpha_1 + \boldsymbol{A}_{01}\alpha_2 + \boldsymbol{A}_{20}\alpha_1^2 + \boldsymbol{A}_{11}\alpha_1\alpha_2 + \boldsymbol{A}_{02}\alpha_2^2\right)\boldsymbol{z} = 0$$

$$(\alpha_1, \alpha_2)$$

Poles of a second order SISO LTI Autonomous model !

(De Moor, B., 2019)

Multiparameter eigenvalue problem

The multiparameter eigenvalue problem $\boldsymbol{\mathcal{M}}(\lambda_1, \ldots, \lambda_n)\boldsymbol{z} = 0$ consists in finding all *n*-tuples $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$ and corresponding vectors $\boldsymbol{z} \in \mathbb{C}^{l \times 1} \backslash \{\boldsymbol{0}\}$, so that

$$\boldsymbol{\mathcal{M}}(\lambda_1, \ldots, \lambda_n)\,\boldsymbol{z} = \left(\sum_{\{\boldsymbol{\omega}\}} \boldsymbol{A}_{\boldsymbol{\omega}} \boldsymbol{\lambda}^{\boldsymbol{\omega}} \right) \boldsymbol{z} = \boldsymbol{0},$$

$$||\boldsymbol{z}|| = 1.$$

Coefficient matrices $\boldsymbol{A}_{\boldsymbol{\omega}} = \boldsymbol{A}_{(\omega_1, \ldots, \omega_n)} \in \mathbb{C}^{k \times l}$ with $k \geq l + n - 1$

An example of a two-parameter eigenvalue problem is $(\boldsymbol{A}_{00} + \lambda \boldsymbol{A}_{10} + \mu \boldsymbol{A}_{01})\boldsymbol{z} = \boldsymbol{0}$

Can be solved using the block Macaulay Matrix framework

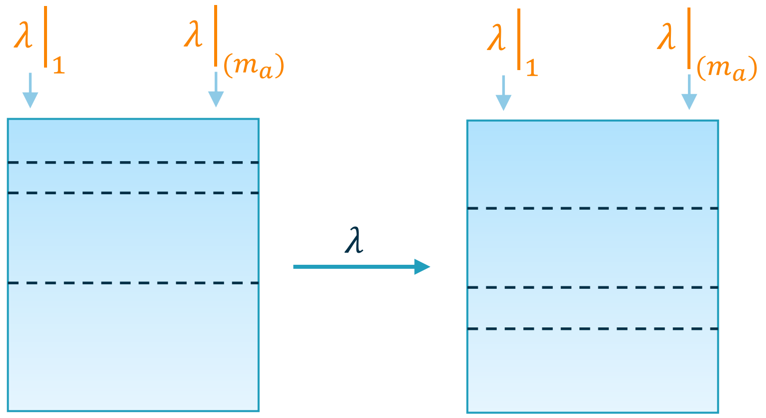
Block Macaulay Matrix Framework

Consider a linear two-parameter eigenvalue of degree $d_s = 1$

$$(\boldsymbol{A}_{00} + \lambda \boldsymbol{A}_{10} + \mu \boldsymbol{A}_{01})\boldsymbol{z} = \boldsymbol{0}$$

$$\begin{matrix} \\ \lambda \\ \mu \end{matrix} \underbrace{\begin{bmatrix} \boldsymbol{A}_{00} & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} & 0 & 0 & 0 \\ 0 & \boldsymbol{A}_{00} & 0 & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} & 0 \\ 0 & 0 & \boldsymbol{A}_{00} & 0 & \boldsymbol{A}_{10} & \boldsymbol{A}_{01} \end{bmatrix}}_{M} \begin{bmatrix} \boldsymbol{z} \\ \lambda \boldsymbol{z} \\ \mu \boldsymbol{z} \\ \lambda^2 \boldsymbol{z} \\ \lambda \mu \boldsymbol{z} \\ \mu^2 \boldsymbol{z} \end{bmatrix} = \boldsymbol{0}$$

null(M)

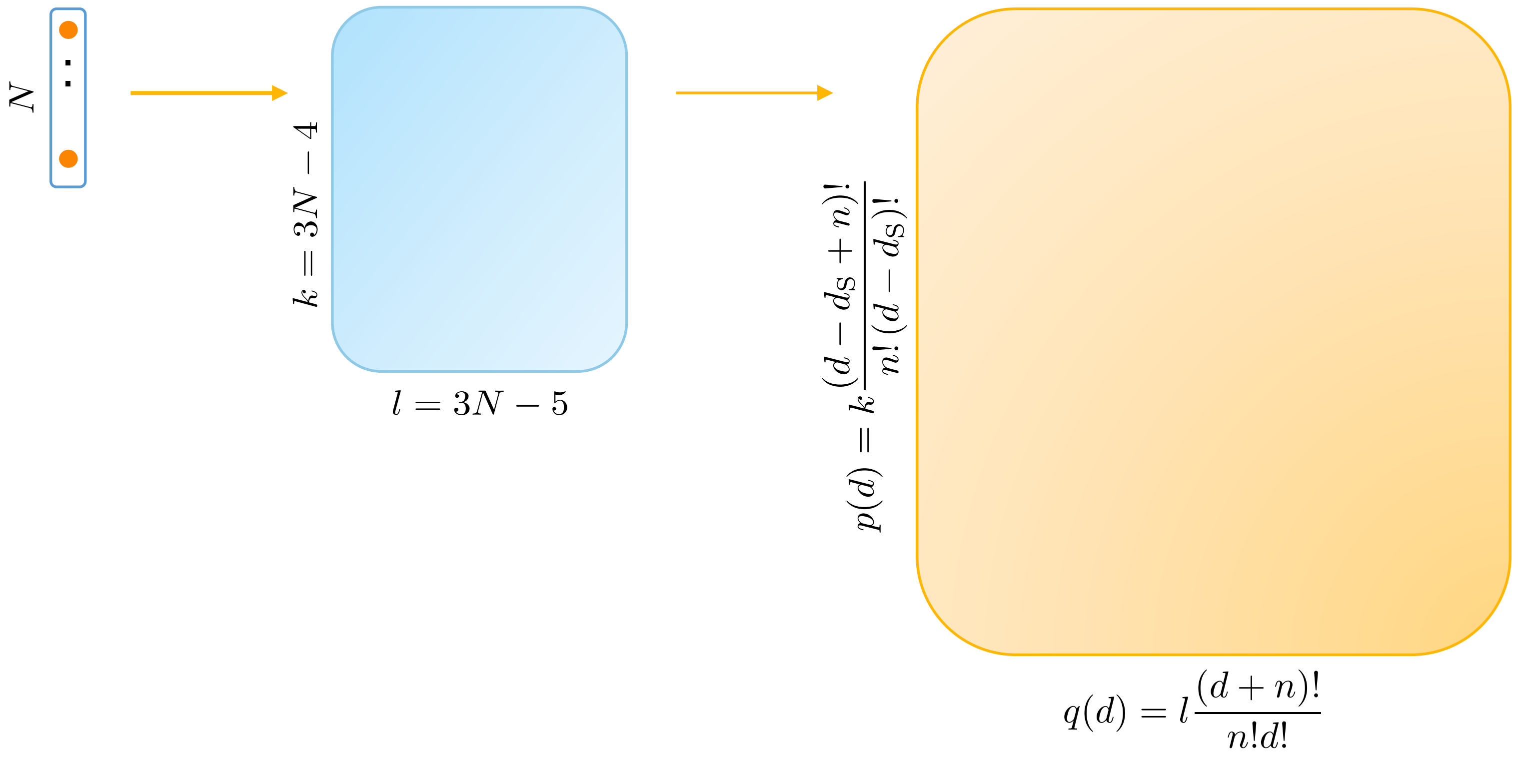


Generate a block Macaulay matrix of degree d using forward block multi shift recursions

Using multi shift invariance property of the null space construct a solvable single parameter eigenvalue problem.

(Vermeersch, C. and De Moor, B., 2022.)

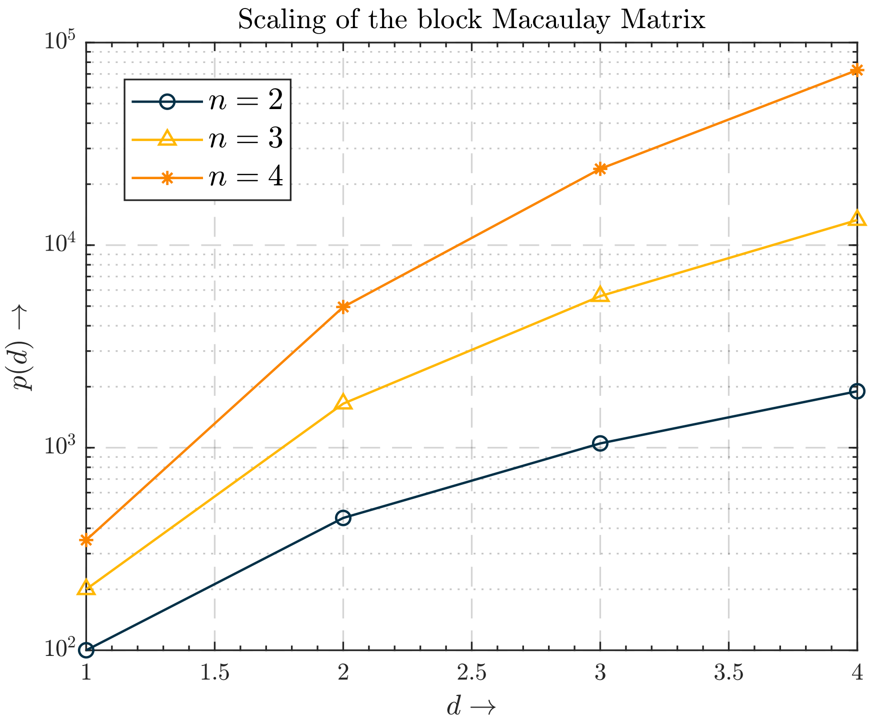
Block Macaulay bottleneck



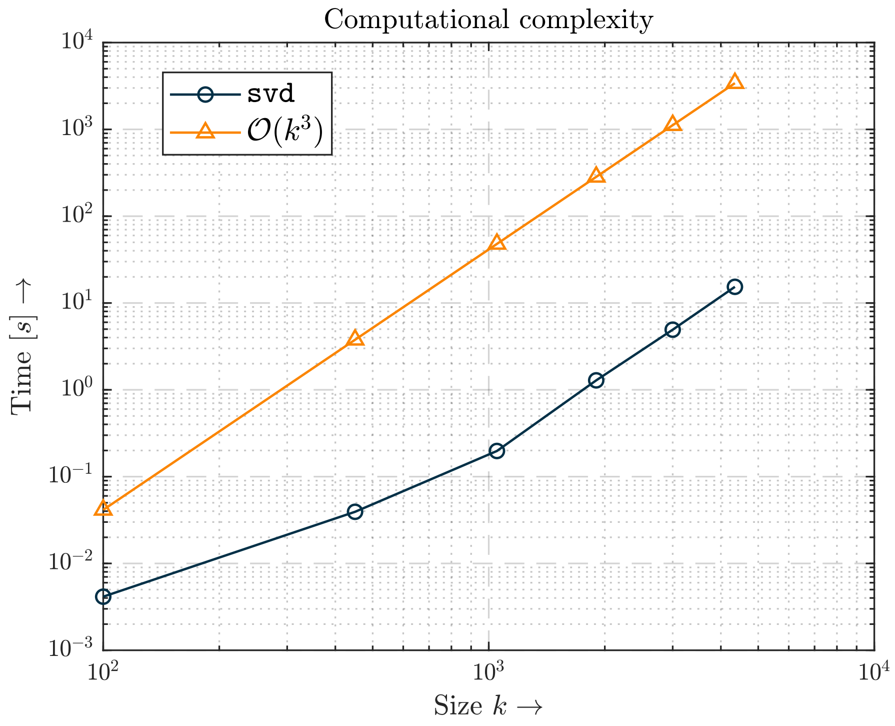
null(M)
$\mathcal{O}(pq \times \min(p, q))$

$\boldsymbol{y} \in \mathbb{R}^N$ $\boldsymbol{A} \in \mathbb{R}^{k \times l}$ $\boldsymbol{M} \in \mathbb{R}^{p(d) \times q(d)}$

Block Macaulay bottleneck



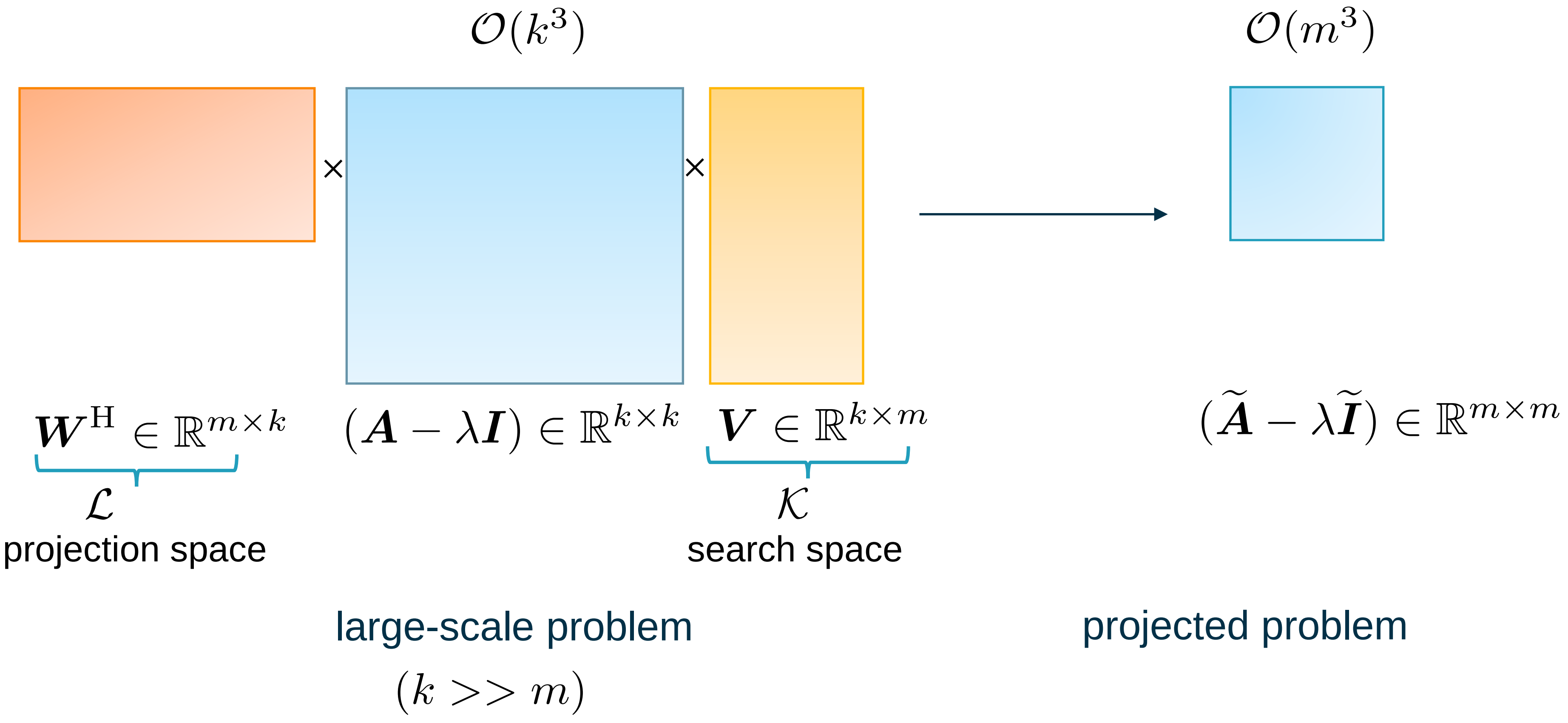
The rows of the block Macaulay matrix for a $k \times l$ matrix with $k = 100$



The computational complexity of the SVD

Subspace methods

For a one-parameter problem $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{z} = \boldsymbol{0}$



Subspace methods

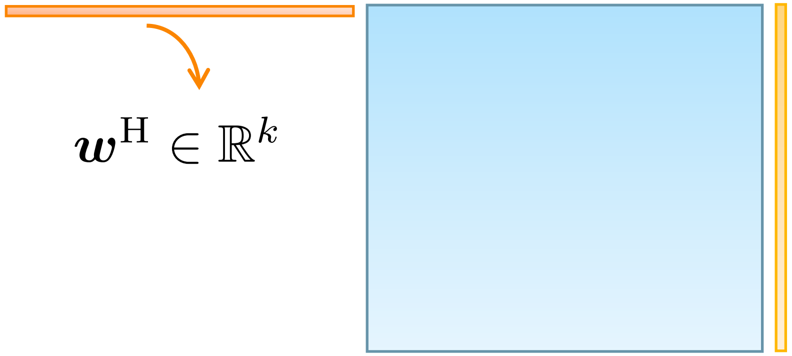
Two natural questions arise

1. Extraction: Does the smaller problem have approximations to the large-scale problem ?
2. Expansion: How to manipulate the large-scale problem to assist in convergence ?

Extraction: Rayleigh quotient

Does the smaller problem contain approximations to the large scale problem ?

Consider $m = 1$ it can be shown that the Rayleigh Quotient yields the eigenvalue when $\boldsymbol{w}, \boldsymbol{v}$ are the left and right eigenvectors.



$\boldsymbol{w}^{\mathrm{H}} \in \mathbb{R}^k$

$\boldsymbol{z}^{\mathrm{H}} \in \mathbb{R}^k$

$= 0$

$(\boldsymbol{A} - \lambda \boldsymbol{I}) \in \mathbb{R}^{k \times k}$

$$\lambda = \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{z}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{z}}$$

Rayleigh quotient

(Parlett 1974)

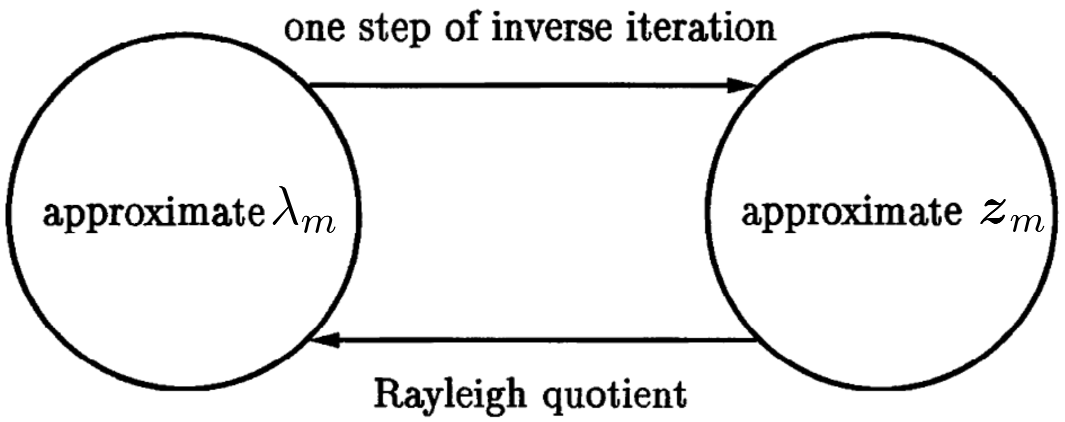
Extraction: Rayleigh quotient iteration

We do not know the eigenvectors before hand

Using Inverse iteration estimates of the eigenvector can be refined

$$\boldsymbol{z}_{m+1} = (\boldsymbol{A} - \lambda_m \boldsymbol{I})^{-1} \boldsymbol{z}_m$$

$$\boldsymbol{w}_{m+1} = (\boldsymbol{A}^{\mathrm{H}} - \lambda_m \boldsymbol{I})^{-1} \boldsymbol{w}_m$$



Rayleigh Quotient Iteration a diagram from *Trefethen and Bau*

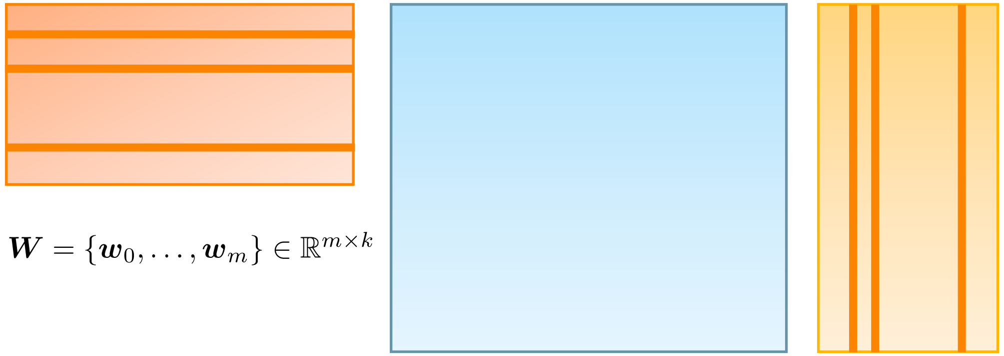
Expansion: Rayleigh-Ritz

Convergence can be accelerated by expanding the search and projection spaces as

There are added benefits to expanding the spaces

- Numerically stable for $\lambda \in \mathbb{C}$
- Can handle larger matrices
- Can help compute a subset of λ

$$(\boldsymbol{W}\hat{w})^{\mathrm{H}}(\boldsymbol{A} - \lambda \boldsymbol{I})(\boldsymbol{V}\hat{\boldsymbol{z}}) = 0$$



$= 0$

$\boldsymbol{W} = \{\boldsymbol{w}_0, \ldots, \boldsymbol{w}_m\} \in \mathbb{R}^{m \times k}$

$(\boldsymbol{A} - \lambda \boldsymbol{I}) \in \mathbb{R}^{k \times k}$

$\boldsymbol{V} = \{\boldsymbol{z}_0, \ldots, \boldsymbol{z}_m\} \in \mathbb{R}^{k \times m}$

Rectangular multiparameter eigenvalue problems

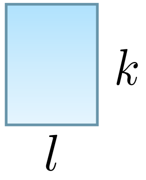
Type 1

$$(\boldsymbol{A} - \lambda \boldsymbol{B})\boldsymbol{z} = 0$$

(k+1) equations	*(m+1)* equations
(1+k) variables	*(1+m)* variables

Type 3

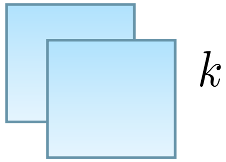
$$(\boldsymbol{A} - \lambda \boldsymbol{B})\boldsymbol{z} = 0$$



(k+1) equations	*(m+1)* equations
(l+1) variables	*(1+m)* variables

Type 2

$$(\boldsymbol{A}_1 - \lambda \boldsymbol{B}_1 - \mu \boldsymbol{C}_1)\boldsymbol{z}_1 = \boldsymbol{0}$$
$$(\boldsymbol{A}_2 - \lambda \boldsymbol{B}_2 - \mu \boldsymbol{C}_2)\boldsymbol{z}_2 = \boldsymbol{0}$$

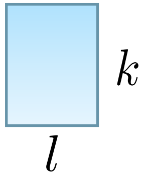


(nk+n) equations	*(mk+m)* equations
(n+nk) variables	*(m+mk)* variables

(Hochstenbach *et al* 2004)

Type 4

$$(\boldsymbol{A} - \lambda \boldsymbol{B} - \mu \boldsymbol{C})\boldsymbol{z} = 0$$



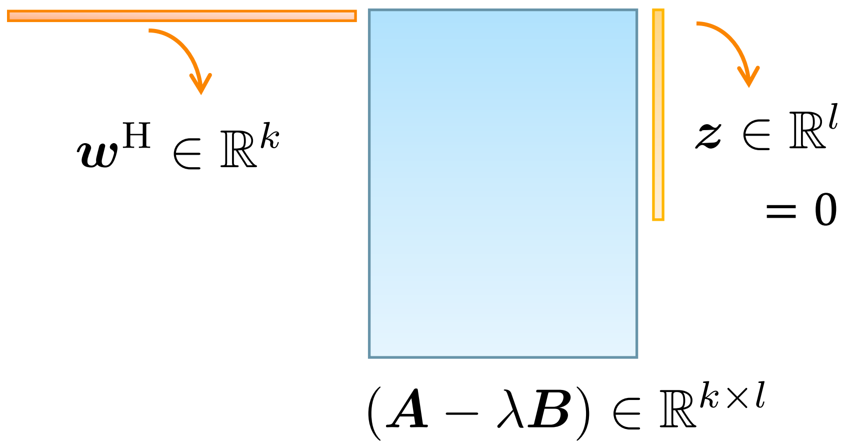
((l+n-1)+1) equations	*(m+1)* equations
(l+n) variables	*(n+m)* variables

underdetermined ?

Extraction: BMM-RQI

Consider a one-parameter rectangular eigenvalue problem

What is the Rayleigh-like quotient?



$$\lambda = \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{z}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{B} \boldsymbol{z}}$$

block Macaulay Rayleigh quotient

Extraction: BMM-RQI

Equations of the Inverse iteration do not hold up for rectangular matrices

We can write for an approximation of the left and right eigenvectors at step m

$$(\boldsymbol{A} - \lambda_m \boldsymbol{B})\boldsymbol{z}_{m+1} = \sigma_l|_m \boldsymbol{w}_{m+1}$$

$$\lambda_{m+1} - \lambda_m = \frac{\sigma_l|_m}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{B} \boldsymbol{v}}$$

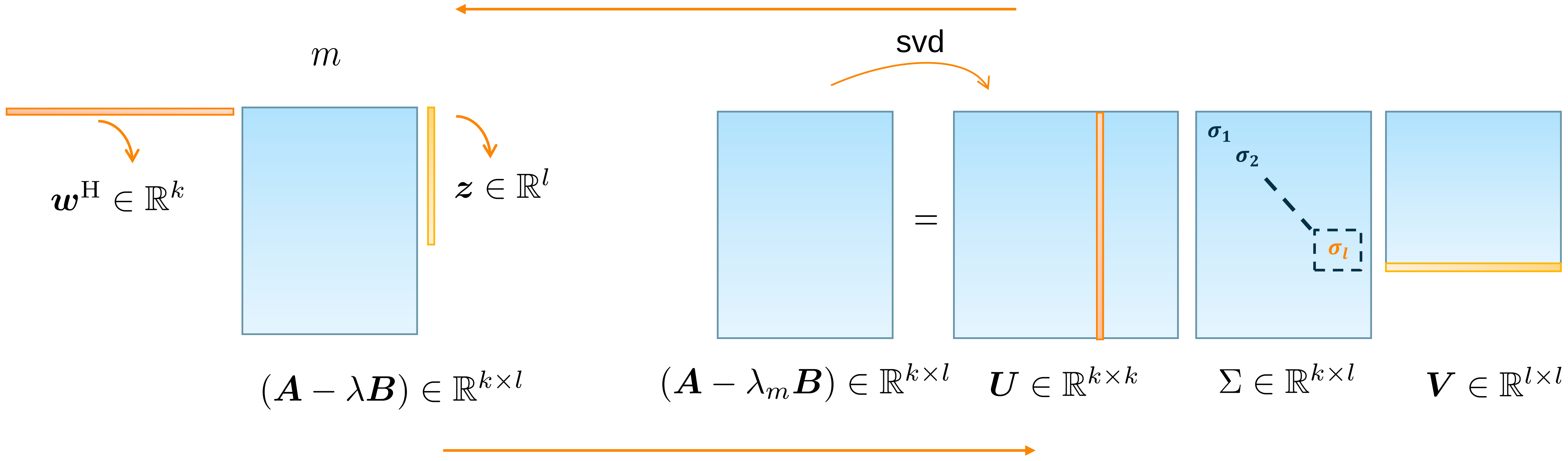
By choosing the vectors which make the pencil the most rank deficient we refine the λ estimate

This can be visualized as traversing to a minima in the Pseudospectra

$\sigma_{min}(A - \lambda B)$ for $\lambda \in \mathbb{C}$

Extraction: BMM-RQI

Adapting the Trefethen and Bau flow to the Rectangular case



Approximate the eigenvalue

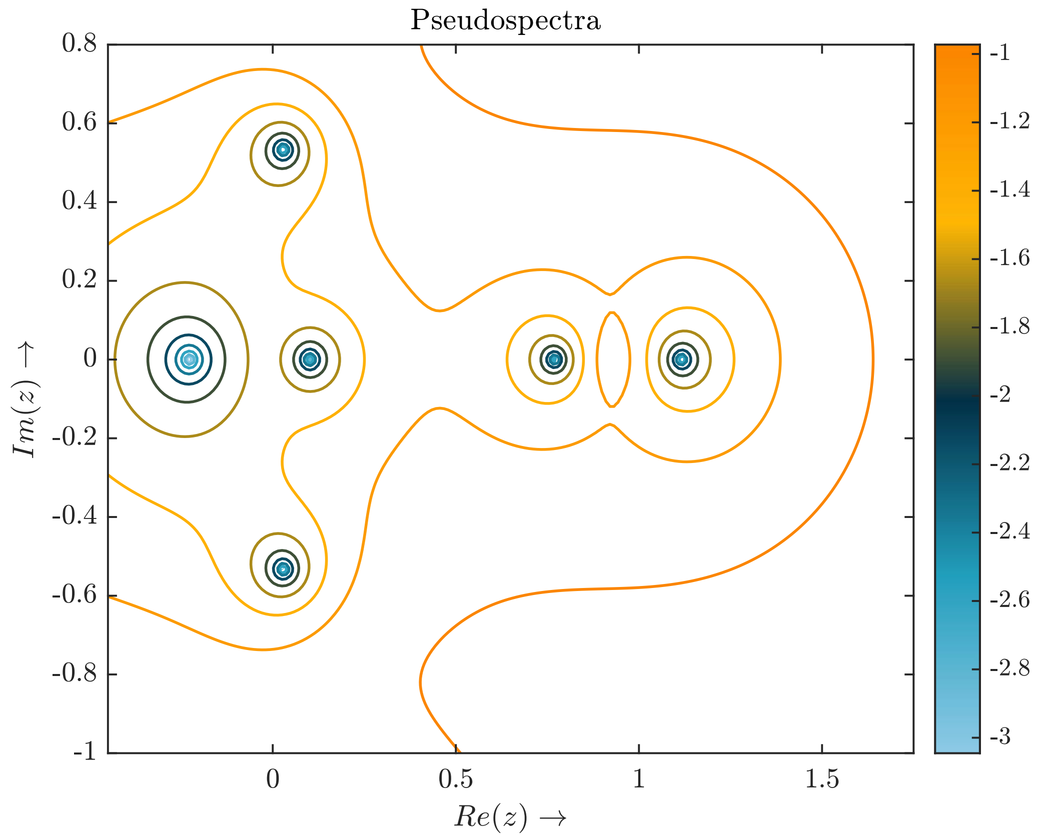
Approximate the eigenvector

Convergence analysis

Consider a randomly generated Rectangular problem

λ
1.027429454379017
1.099012545384454
1.115994481273991
1.116571414400400
1.116572007766462
1.116572007767087

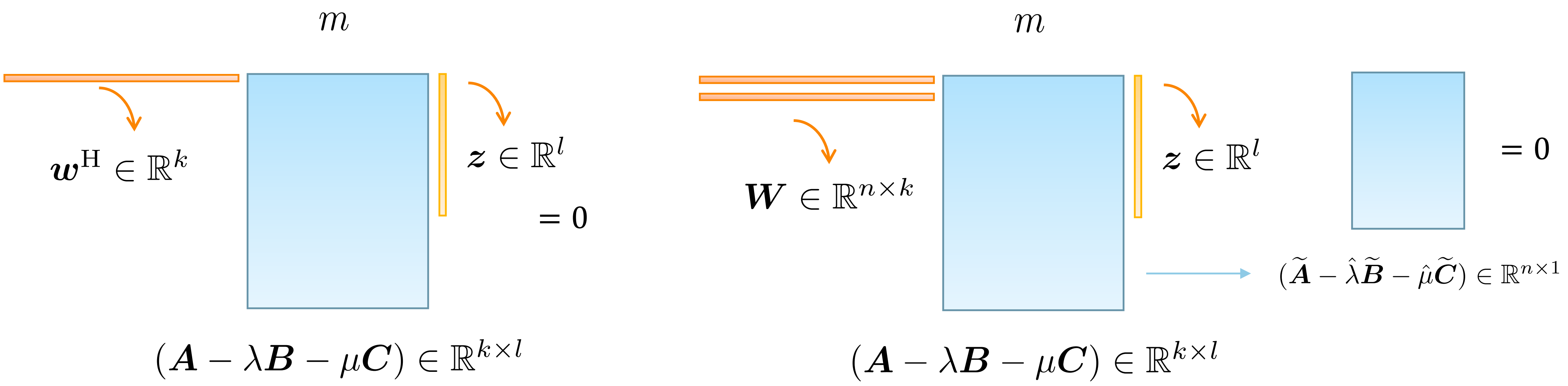
Quadratic convergence of the BMM-RQI



Pseudospectra of the rectangular problem

Extraction: BMM-RQI

Consider a multiparameter eigenvalue problem



For a multiparameter problem the above
Construction of the BMM-RQI is underdetermined

Introduce extra constraints !
- Random vectors
- Or orthogonal to $\boldsymbol{w}$

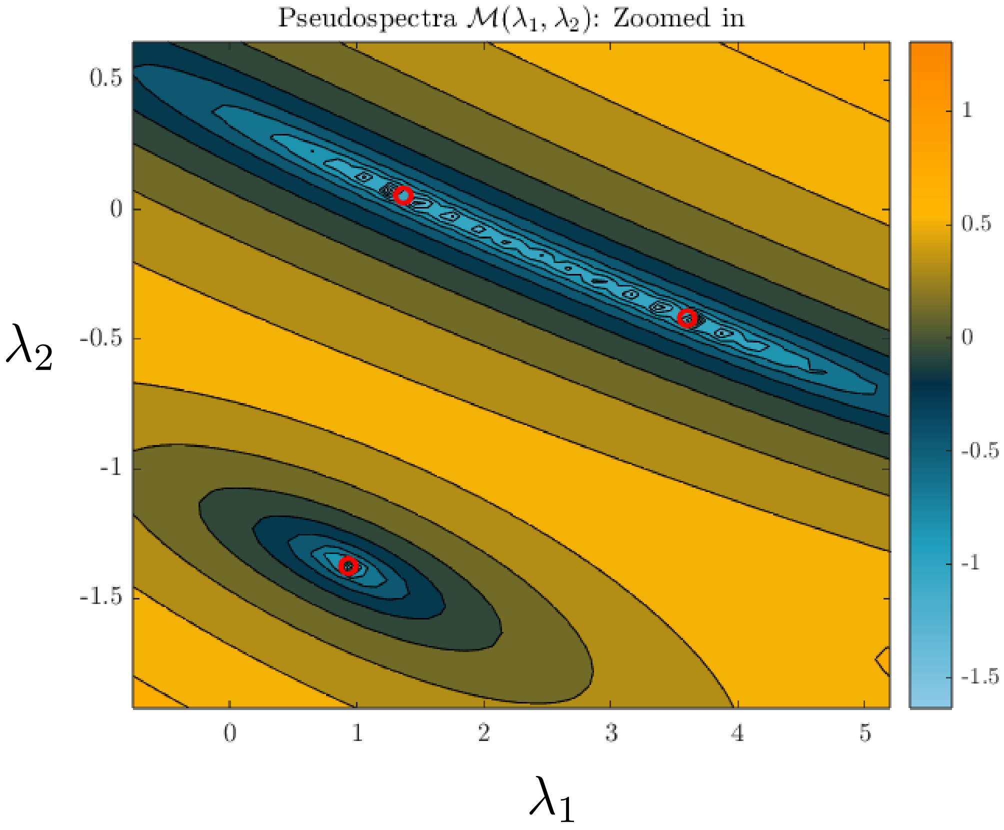
Example: A linear 2-parameter eigenvalue problem

Consider a two-parameter problem

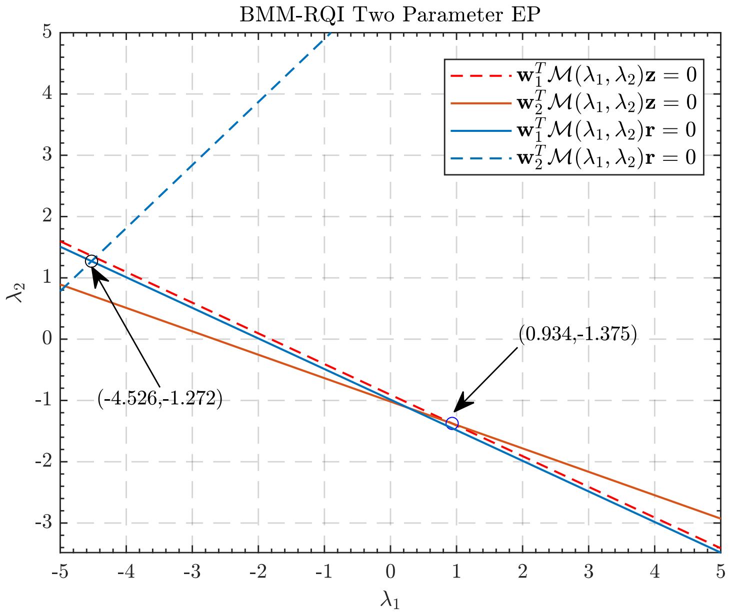
$$\underbrace{\left(\begin{bmatrix} 2 & 6 \\ 4 & 5 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} + \mu \begin{bmatrix} 4 & 2 \\ 0 & 8 \\ 1 & 1 \end{bmatrix}\right)}_{\boldsymbol{\mathcal{M}}(\lambda,\mu)} \boldsymbol{z} = \mathbf{0}$$

λ_1	λ_2	z_1	z_2	w_1	w_2	w_3
0.9338	-1.375	0.7848	0.6197	-0.8381	-0.5432	-0.0500
1.3683	0.0552	0.8623	-0.5065	0.6064	-0.2791	-0.7446
3.6026	-0.4183	0.7958	-0.6056	0.7056	-0.7084	0.0192

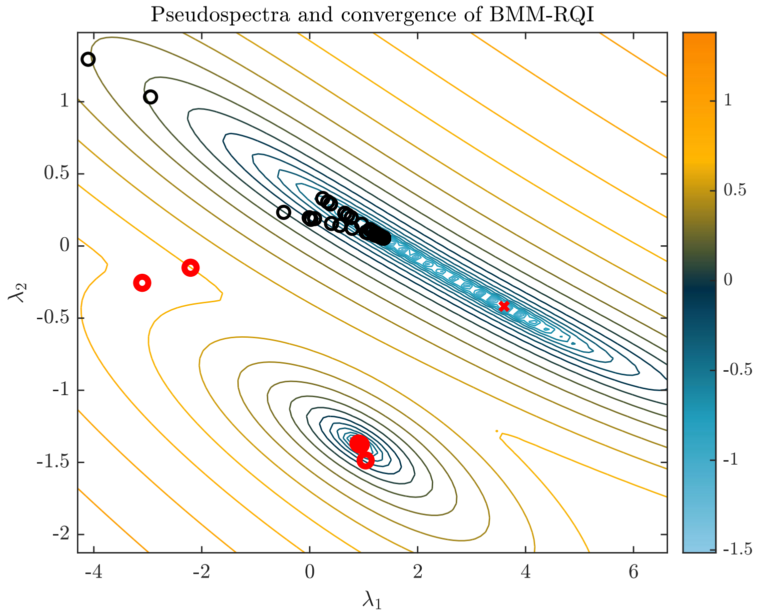
Example: A linear 2-parameter eigenvalue problem



BMM-RQI for the two-parameter problem



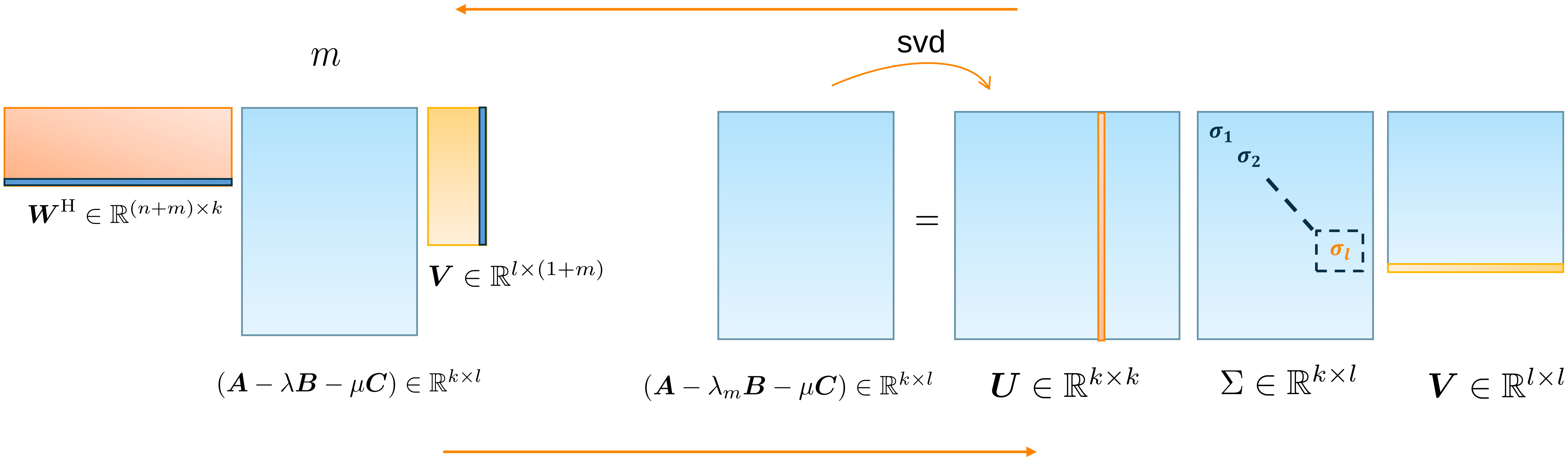
The BMM-RQI equations evaluated at eigenvectors and the random vectors



Convergence starting with different initial conditions

Expansion: BMM-RR

Expanding the search and projections spaces similar to the RR approach



Approximate the eigenvalue

Approximate the eigenvector

BMM-RR Speed Up !

Solving the MEP associated with a randomly generated order two Least-squares realization problem

N	Problem size	Computation time(s)	Residual error
50	146×145	0.8604	4.5576×10^{-9}
100	296×295	3.9671	9.1061×10^{-9}
150	446×445	5.3200	8.0039×10^{-9}
200	596×595	11.0338	5.5277×10^{-9}
250	746×745	16.9015	8.5270×10^{-9}

Using the existing (block) Macaulay approaches, it can take up-to 100s to solve a 14 x 13 problem !

(Vermeersch and De Moor 2022b)

Conclusion and Future work

- A first-of-its-kind subspace algorithm for rectangular multiparameter eigenvalue problems
- Can be leveraged to find solutions to large-scale system identification problem
- Linearizing nonlinear MEPs arising from system identification can help locate real roots to the problem
- Rigorous proof of convergence for the multiparameter problems need to be carried out

Questions

References

- De Moor, B., 2019, June. Least squares realization of LTI models is an eigenvalue problem. In *2019 18th European Control Conference (ECC)* (pp. 2270-2275). IEEE.
- Vermeersch, C. and De Moor, B., 2022. Two complementary block Macaulay matrix algorithms to solve multiparameter eigenvalue problems. Linear Algebra and its Applications, 654, pp.177-209.
- Trefethen, L.N. and Bau, D., 2022. Numerical linear algebra (Vol. 181). SIAM.
- Hochstenbach, M.E., Kosir, T. and Plestenjak, B., 2004. A Jacobi--Davidson type method for the two-parameter eigenvalue problem. SIAM Journal on Matrix Analysis and Applications, 26(2), pp.477-497.
- Parlett, B.N., 1974. The Rayleigh quotient iteration and some generalizations for nonnormal matrices. Mathematics of Computation, 28(127), pp.679-693.
- Vermeersch, C. and De Moor, B., 2022b. Two double recursive block macaulay matrix algorithms to solve multiparameter eigenvalue problems. IEEE Control Systems Letters, 7, pp.319-324.

Extraction: BMM-RQI

Equations of the Inverse iteration do not hold up for rectangular matrices

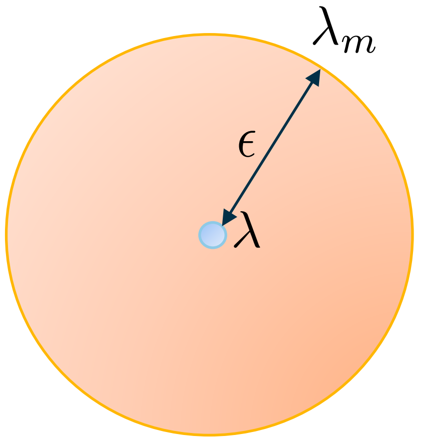
However, we can write at Iteration m

Smallest singular value of the matrix pencil

$$(\boldsymbol{A} - \lambda_m \boldsymbol{B})\boldsymbol{z}_{m+1} = \sigma_l|_m \boldsymbol{w}_{m+1}$$

$$\lambda_{m+1} - \lambda_m = \frac{\sigma_l|_m}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{B} \boldsymbol{v}}$$

$$|\lambda - \lambda_m| = \mathcal{O}(\epsilon^2)$$



Convergence analysis

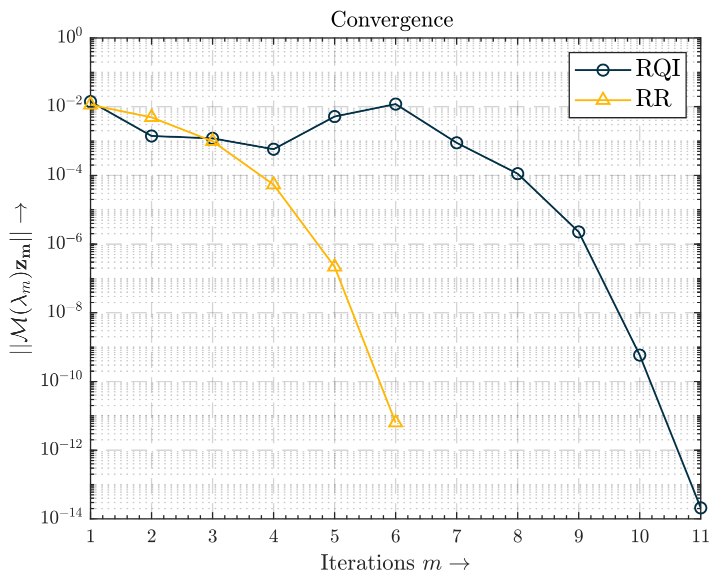
Rayleigh quotient iteration

λ
5.000000000000001
5.213114754098361
5.214319743184031
5.214319743377537

Symmetric $\boldsymbol{A} \in \mathbb{R}^{3\times 3}$

$|\lambda_{m+1} - \lambda_m| \propto \mathcal{O}(\epsilon^3)$

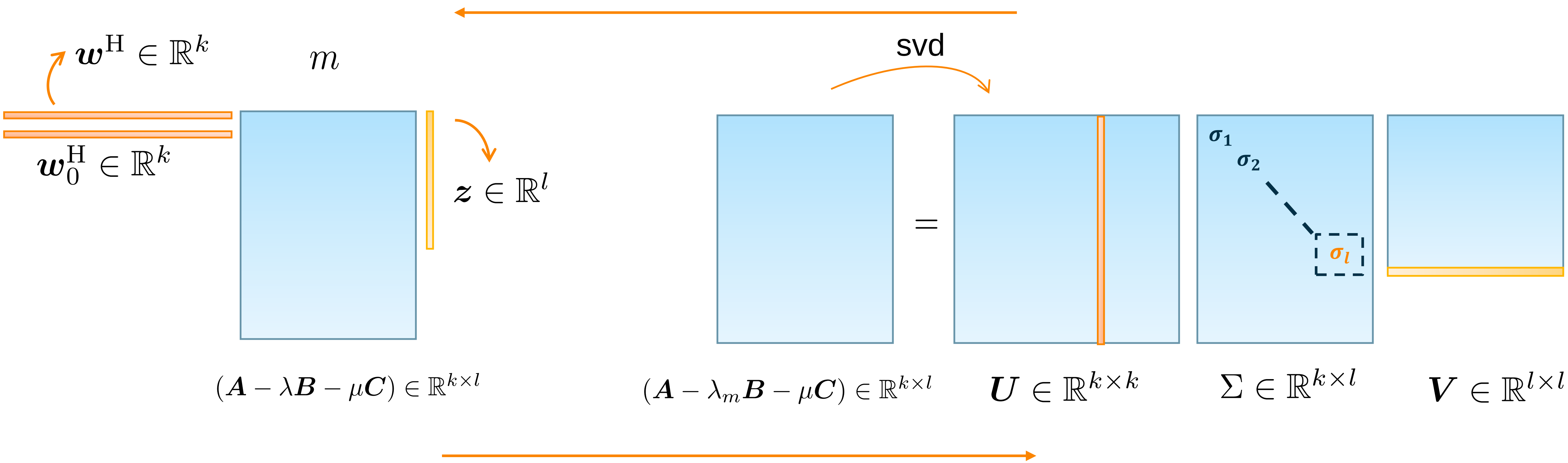
Rayleigh-Ritz method



Random $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{500\times 500}$

RR takes lesser # of iterations

Extraction: BMM-RQI



Approximate the eigenvalue

Approximate the eigenvector