Pole swapping methods for the eigenvalue problem

Rational QR algorithms

Daan Camps

ICIAM 2019, Valencia

July 16, 2019

KU Leuven - University of Leuven - Department of Computer Science - NUMA Section

Acknowledgements

Collaborators:

- Raf Vandebril
- Karl Meerbergen
- Paul Van Dooren
- Thomas Mach
- David Watkins
- Nicola Mastronardi

Overview

Introduction

Generalized eigenvalue problems

- Let $A, B \in \mathbb{F}^{n \times n}$ define a *matrix pair* (A, B) or *matrix pencil* $A - \lambda B$.
- Regular: $\det(A - \lambda B) \neq 0$.
- Generalized eigenvalue problem:

$$A\boldsymbol{x} = \lambda B\boldsymbol{x}, \quad \lambda \in \bar{\mathbb{C}} := \mathbb{C} \cup \{\infty\}.$$

Generalized eigenvalue problems

- For a regular matrix pencil $A - \lambda B$ there exists unitary matrices Q and Z such that

$$S - \lambda T = Q^*(A - \lambda B)Z$$

  with $S - \lambda T$ upper triangular and $\Lambda(A, B) = \{s_{11}/t_{11}, s_{22}/t_{22}, \ldots\}$.
- The QZ method (Moler-Stewart, 1973) is the default algorithm to compute the generalized Schur decomposition.

Bulge chasing

- The QZ method generalizes the QR method of (Francis, 1961-62) for the standard eigenvalue problem
- It is a bulge chasing algorithm which consists out of two phases:
  1. Initial (direct) reduction to equivalent Hessenberg, upper triangular form

$$H - \lambda R = Q^*(A - \lambda B)Z$$

  2. Iterative bulge chasing phase to compute (real) generalized Schur decomposition

$$S - \lambda T = Q^*(A - \lambda B)Z$$

Bulge chasing

Bulge chasing =

- Motivated by implicit Q theorems
  $\Rightarrow$ iterates are uniquely determined by $\boldsymbol{q}_1 = p(AB^{-1})\boldsymbol{e}_1$
- Nested subspace iteration with a change of basis accelerated by polynomials (shifts) (Elsner-Watkins, 1991; Watkins, 1993)

$\rightarrow$ These results are based on a connection with Krylov subspaces.

Pole swapping

Rational Krylov

Polynomial	Rational
Krylov subspace $\mathcal{K}_{m+1}(A, \boldsymbol{v}) := \mathcal{R}(\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^m \boldsymbol{v})$	**rational Krylov subspace** $\mathcal{K}^{\text{rat}}_{m+1}(A, \boldsymbol{v}, \Xi) := q(A)^{-1}\mathcal{K}_{m+1}(A, \boldsymbol{v})$ • $\Xi = (\xi_1, \ldots, \xi_m) \subset \bar{\mathbb{C}} \setminus \Lambda$ • $q(z) = \prod_{\xi_i \neq \infty}(z - \xi_i)$
Arnoldi Decomposition (Arnoldi, 1951) $AV_m = V_{m+1}\underline{H}_m$	**rational Arnoldi** (Ruhe, 1998) $AV_{m+1}\underline{K}_m = V_{m+1}\underline{L}_m$ • $\ell_{i+1,i}/k_{i+1,i} = \xi_i$
Uniqueness $\boldsymbol{v}$ fixed $\Leftrightarrow (V_{m+1}, \underline{H}_m)$ unique	**Uniqueness** (Berljafa-Güttel, 2015) $\boldsymbol{v}$ and Ξ fixed $\Leftrightarrow (V_{m+1}, \underline{K}_m, \underline{L}_m)$ unique

Rational Krylov

Berljafa-Güttel (2015)

Two methods to change the poles in a rational Arnoldi decomposition

$$A V_{m+1} \underline{K}_m = V_{m+1} \underline{L}_m$$

- Implicitly by changing $\boldsymbol{v}_1$
- Explicitly by pole swapping

$\Rightarrow$ Equivalent if $\mathcal{R}(V_{m+1})$ is fixed.

C.-Meerbergen-Vandebril (2019a)

The pole swapping technique can be used as a direct method for the eigenvalue problem: rational QZ algorithm.

Rational QZ

Hessenberg pencils

$$
\underbrace{\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times
\end{matrix}}_{A}
\quad , \quad
\underbrace{\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times
\end{matrix}}_{B}
$$

Hessenberg pencils

$$\begin{matrix} \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times & \times & \times \\ & & \times & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times & \times \\ & & & & \times & \times & \times & \times \\ & & & & & \times & \times & \times \\ & & & & & & \times & \times \end{matrix} \quad , \quad \begin{matrix} \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times & \times & \times \\ & & \times & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times & \times \\ & & & & \times & \times & \times & \times \\ & & & & & \times & \times & \times \\ & & & & & & \times & \times \end{matrix}$$

$$A \qquad\qquad\qquad B$$

Rational QZ

Hessenberg pencils

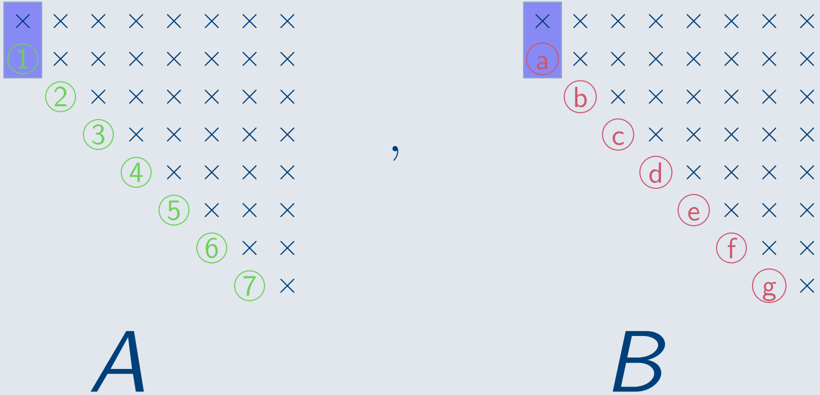
$$
\begin{array}{cccccccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\textcircled{1} & \times & \times & \times & \times & \times & \times & \times \\
 & \textcircled{2} & \times & \times & \times & \times & \times & \times \\
 & & \textcircled{3} & \times & \times & \times & \times & \times \\
 & & & \textcircled{4} & \times & \times & \times & \times \\
 & & & & \textcircled{5} & \times & \times & \times \\
 & & & & & \textcircled{6} & \times & \times \\
 & & & & & & \textcircled{7} & \times
\end{array}
\quad , \quad
\begin{array}{cccccccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\textcircled{a} & \times & \times & \times & \times & \times & \times & \times \\
 & \textcircled{b} & \times & \times & \times & \times & \times & \times \\
 & & \textcircled{c} & \times & \times & \times & \times & \times \\
 & & & \textcircled{d} & \times & \times & \times & \times \\
 & & & & \textcircled{e} & \times & \times & \times \\
 & & & & & \textcircled{f} & \times & \times \\
 & & & & & & \textcircled{g} & \times
\end{array}
$$

$$A \qquad\qquad B$$

$$\text{pole tuple } \Xi = \left(\frac{\textcircled{1}}{\textcircled{a}}, \frac{\textcircled{2}}{\textcircled{b}}, \ldots\right) \subset \bar{\mathbb{C}}$$

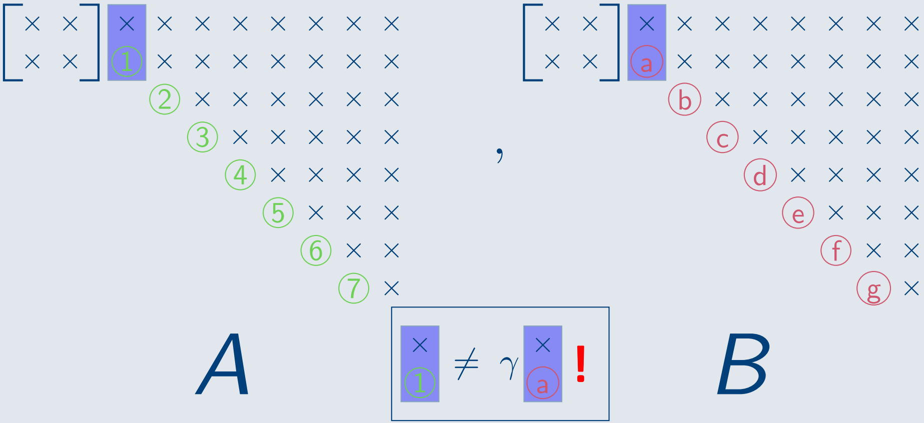
Rational QZ: an example

Introducing a shift



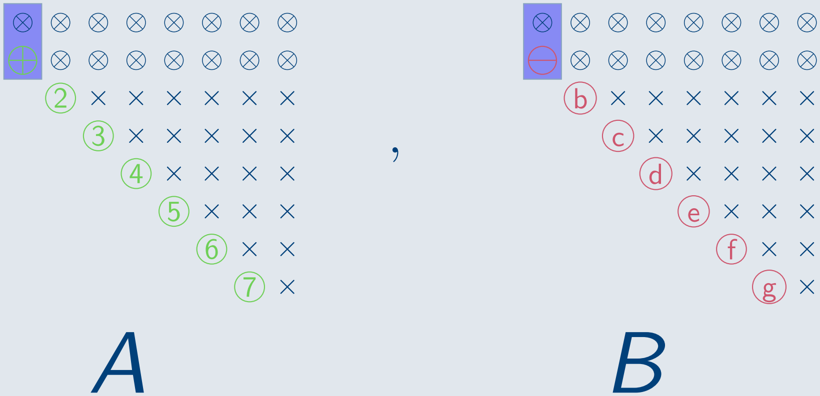
Rational QZ: an example

Introducing a shift



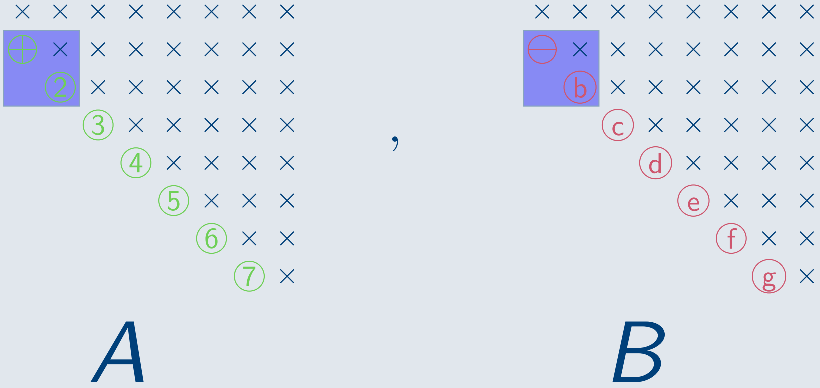
Rational QZ: an example

Introducing a shift



Rational QZ: an example

Swapping poles



Rational QZ: an example

Swapping poles

Classical problem in NLA:

Reordering generalized Schur form

(Van Dooren, 1981), (Kågström, 1993), (C.-Mach-Vandebril-Watkins, 2019)

$$\frac{\oplus}{\ominus} \neq \frac{②}{ⓑ}\,! \qquad \frac{\oplus}{\ominus} \neq \frac{0}{0}\,! \qquad \frac{②}{ⓑ} \neq \frac{0}{0}\,!$$

$$\Rightarrow Q^* = \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix}, Z = \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix}$$

Rational QZ: an example

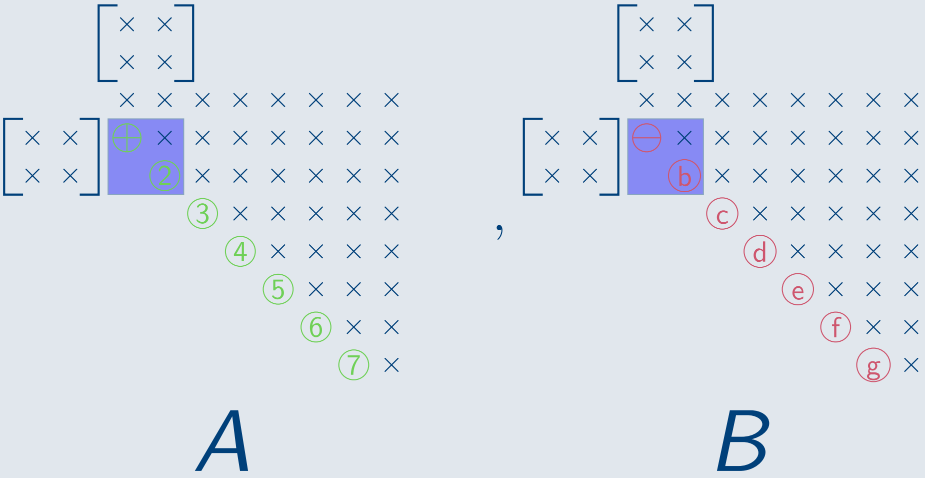
Swapping poles

Table 1: Distribution of errors $\hat{a}_{21}/\|A\|$ and $\hat{b}_{21}/\|B\|$ for our method, Van Dooren's method, and the Sylvester method.

$\hat{x}_{21}/\|X\|$		$[0, 10^{-16}]$	$(10^{-16}, 10^{-15}]$	$(10^{-15}, 10^{-10}]$	$(10^{-10}, 10^{-5}]$	$(10^{-5}, 10^{0}]$
Our method	A	99.71%	0.29%	0%	0%	0%
	B	99.85%	0.15%	0%	0%	0%
Van Dooren	A	98.19%	0.55%	0.93%	0.27%	0.06%
	B	98.19%	0.55%	0.93%	0.27%	0.06%
Sylvester	A	93.34%	5.88%	0.57%	0.17%	0.04%
	B	93.34%	5.88%	0.57%	0.17%	0.04%

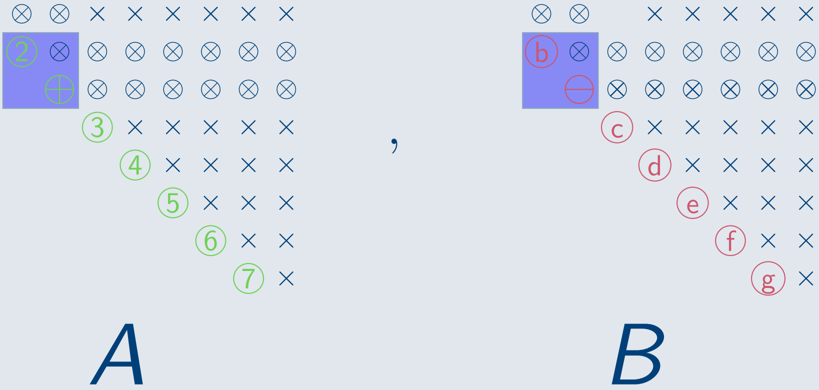
Rational QZ: an example

Swapping poles



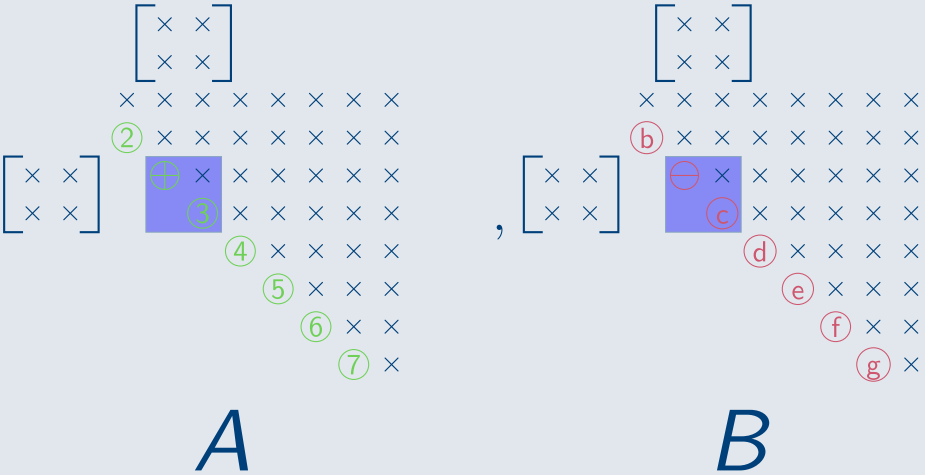
Rational QZ: an example

Swapping poles



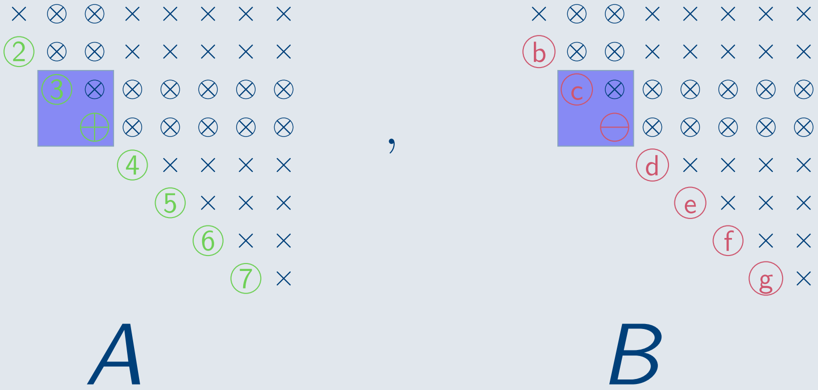
Rational QZ: an example

Swapping poles



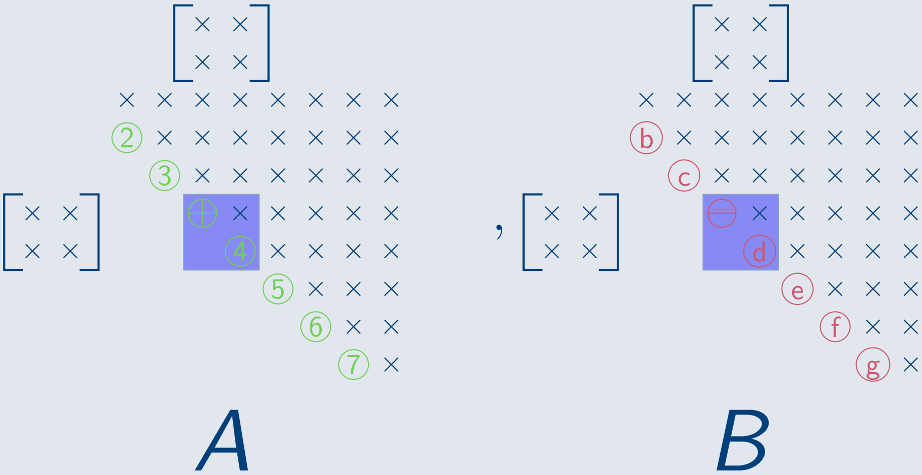
Rational QZ: an example

Swapping poles



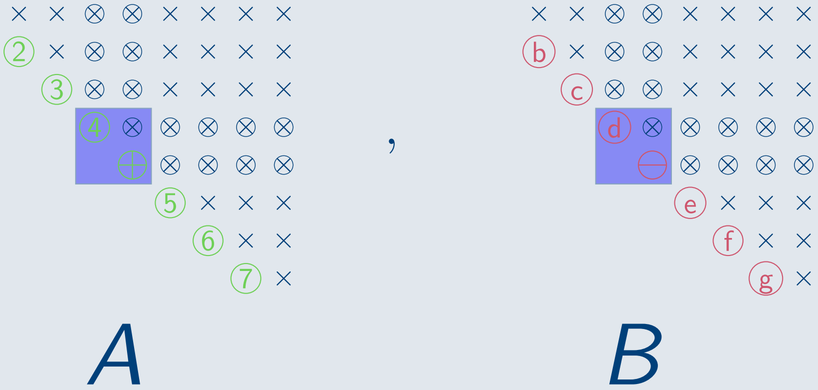
Rational QZ: an example

Swapping poles



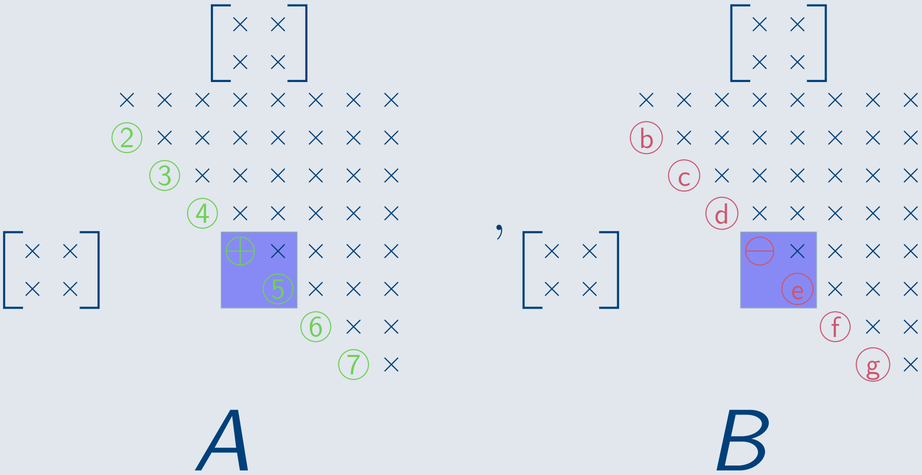
Rational QZ: an example

Swapping poles



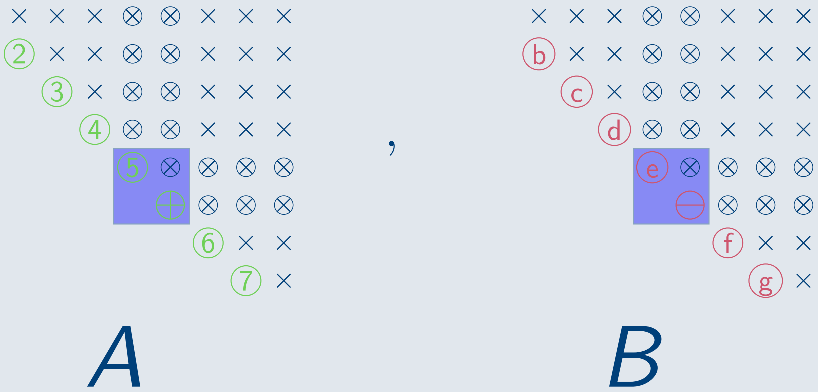
Rational QZ: an example

Swapping poles



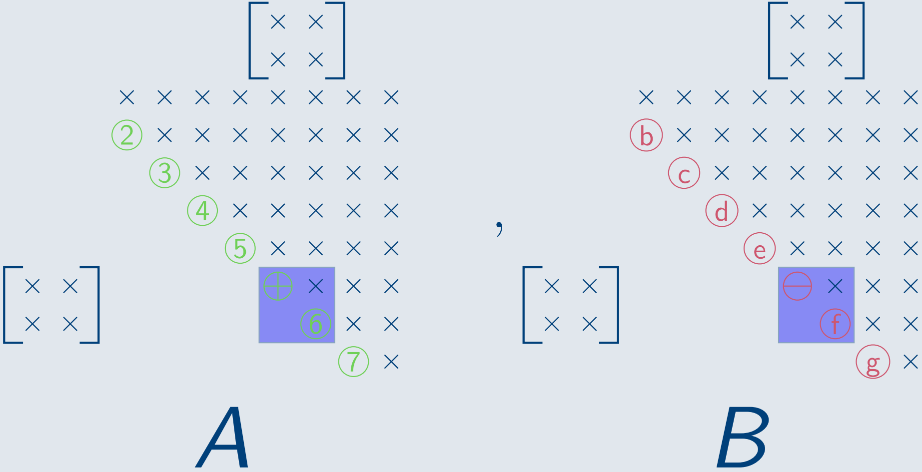
Rational QZ: an example

Swapping poles



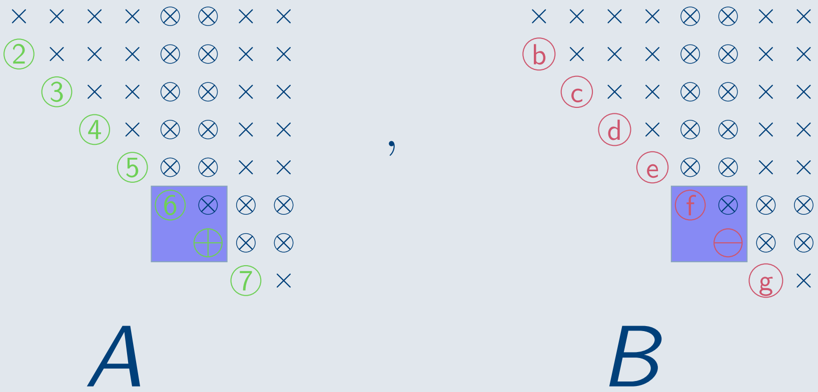
Rational QZ: an example

Swapping poles



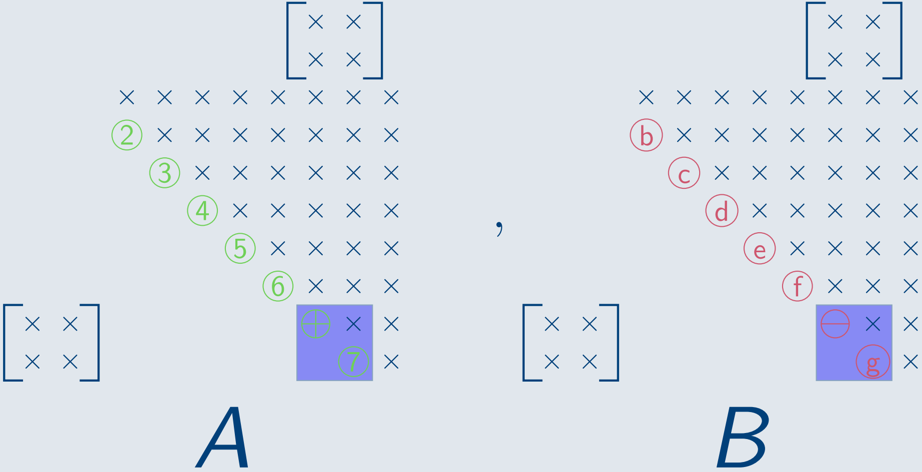
Rational QZ: an example

Swapping poles



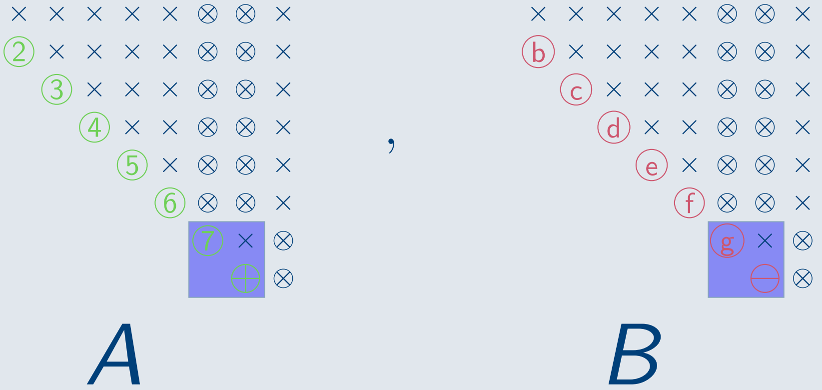
Rational QZ: an example

Swapping poles



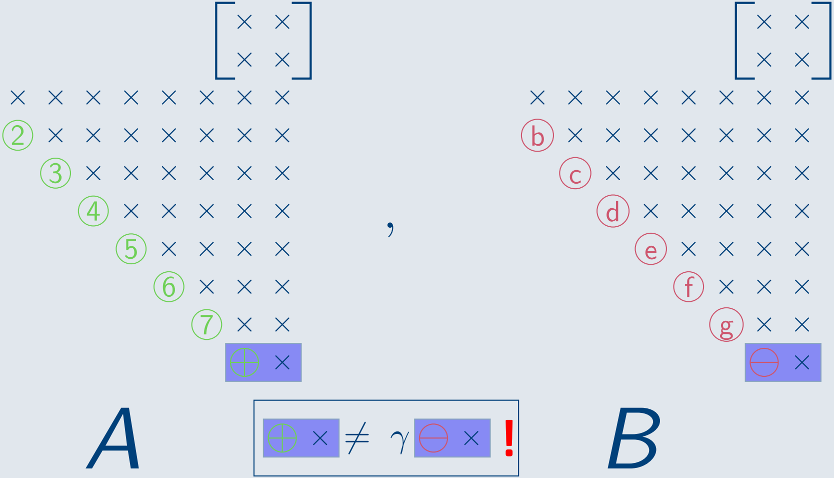
Rational QZ: an example

Swapping poles



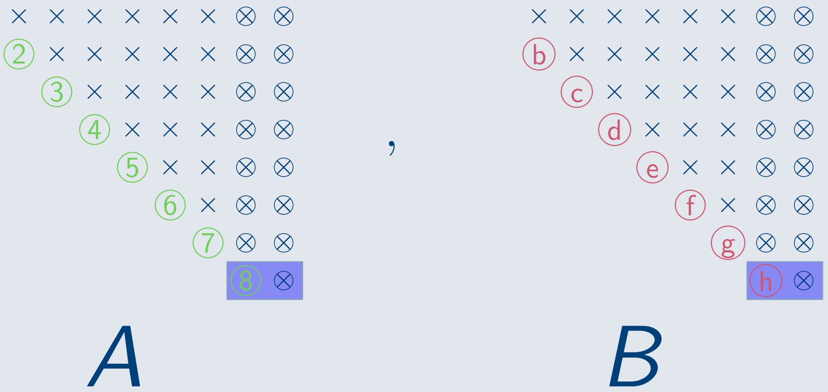
Rational QZ: an example

Introducing a pole



Rational QZ: an example

Introducing a pole

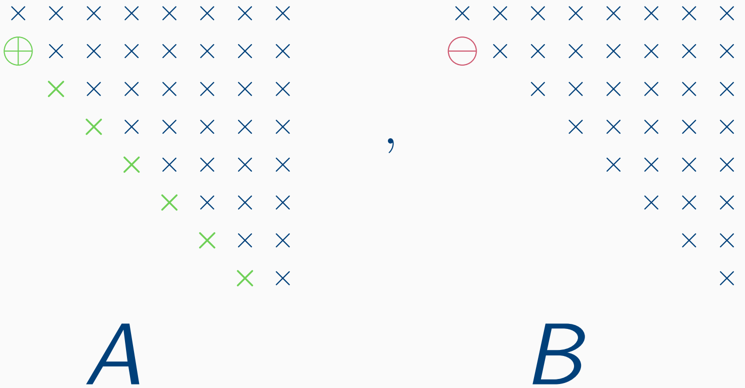


Classical QZ as a special case

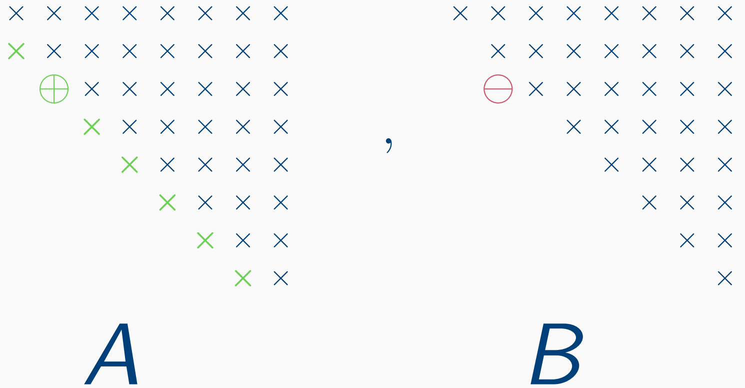
$$
\begin{array}{cccccccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times
\end{array}
\quad , \quad
\begin{array}{cccccccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times \\
 & & & & & & & \times
\end{array}
$$

$$A \qquad\qquad\qquad B$$

Classical QZ as a special case



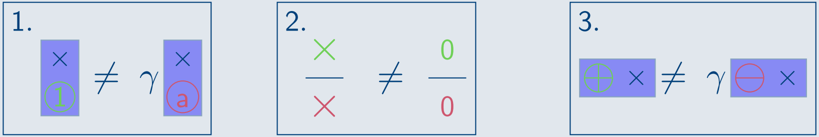
Classical QZ as a special case



Theoretical results

Definition: Properness.

The Hessenberg pair (A, B) is called *proper* if:



Theoretical results

Theorem. (C.-Meerbergen-Vandebril, 2019a)

If (A, B) is a proper Hessenberg pair with poles $(\xi_1, \ldots, \xi_{n-1})$ distinct from the eigenvalues. Then for $i = 1, \ldots, n$:

$$\mathcal{K}_i^{\text{rat}}(AB^{-1}, \boldsymbol{e}_1, (\xi_1, \ldots, \xi_{i-1})) = \mathcal{E}_i := \mathcal{R}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i),$$

while for $i = 1, \ldots, n-1$:

$$\mathcal{K}_i^{\text{rat}}(B^{-1}A, \boldsymbol{e}_1, (\xi_2, \ldots, \xi_i)) = \mathcal{E}_i.$$

Theoretical results

Implicit Q Theorem. (C.-Meerbergen-Vandebril, 2019a)

Given a regular matrix pair (A, B). The matrices Q and Z that transform it to proper Hessenberg form,

$$(\hat{A}, \hat{B}) = Q^* \, (A, B) \, Z,$$

are determined *essentially unique* if $Q\boldsymbol{e}_1$ and the (order of the) poles are fixed.

Theoretical results

Rational accelerated subspace iteration. (C.-Meerbergen-Vandebril, 2019a)

A rational QZ step with shift $\varrho \notin \{\Lambda, \Xi\}$ on a proper Hessenberg pencil with poles $(\xi_1, \ldots, \xi_{n-1})$ and new pole ξ_n, all distinct from Λ, performs nested subspace iteration for $i = 1, \ldots, n{-}1$ accelerated by

$$Q\mathcal{E}_i = \mathcal{R}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_i) = (A - \varrho B)(A - \xi_i B)^{-1}\mathcal{E}_i$$
$$Z\mathcal{E}_i = \mathcal{R}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_i) = (A - \xi_{i+1} B)^{-1}(A - \varrho B)\mathcal{E}_i,$$

followed by a change of basis.

→ Subspace iteration with rational filter

→ More modular (single swap) convergence theory: (C.-Mach-Vandebril-Watkins, 2019).

Theoretical results

Exactness result (C., 2019)

Let (A, B) be a proper Hessenberg pencil with poles Ξ. Furthermore, let ϱ be an eigenvalue of (A, B) which is distinct from Ξ. A rational QZ step, $Q^*(A, B)Z$, with shift ϱ leads to a deflation in the last rows of $Q^*(A, B)Z$.

Rational QZ

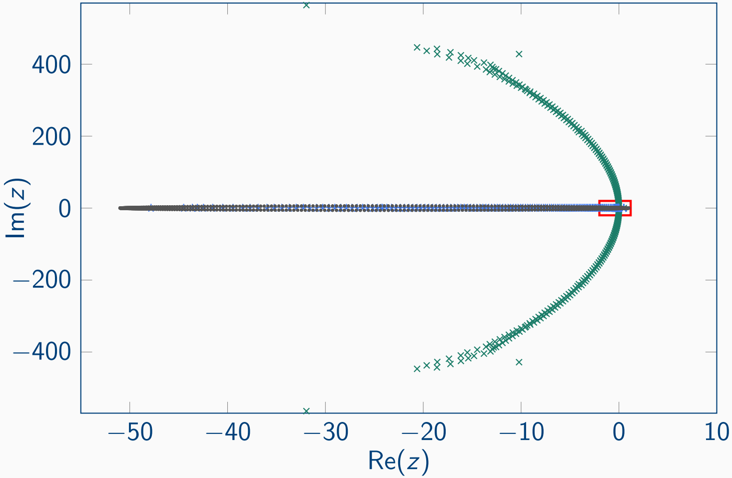
Pole swapping =

- Motivated by implicit Q theorems
  $\Rightarrow$ iterates are uniquely determined by $\boldsymbol{q}_1 = q(AB^{-1})\boldsymbol{e}_1$ and poles in pencil
- Nested subspace iteration with a change of basis accelerated by rational functions (shifts and poles)

$\rightarrow$ These results are based on a connection with rational Krylov subspaces

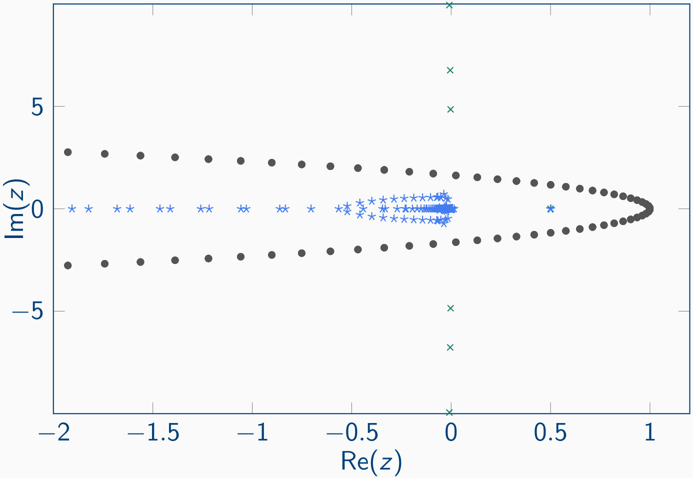
Numerical example: Reduction to Hessenberg form

Data: MHD matrix pair from MatrixMarket, $n = 1280$



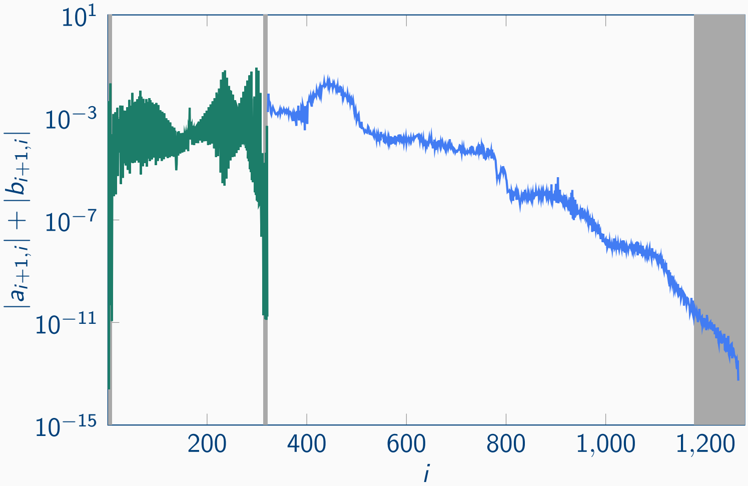
Numerical example: Reduction to Hessenberg form

Data: MHD matrix pair from MatrixMarket, $n = 1280$



Numerical example: Reduction to Hessenberg form

Data: MHD matrix pair from MatrixMarket, $n = 1280$



Multishift, multipole rational QZ

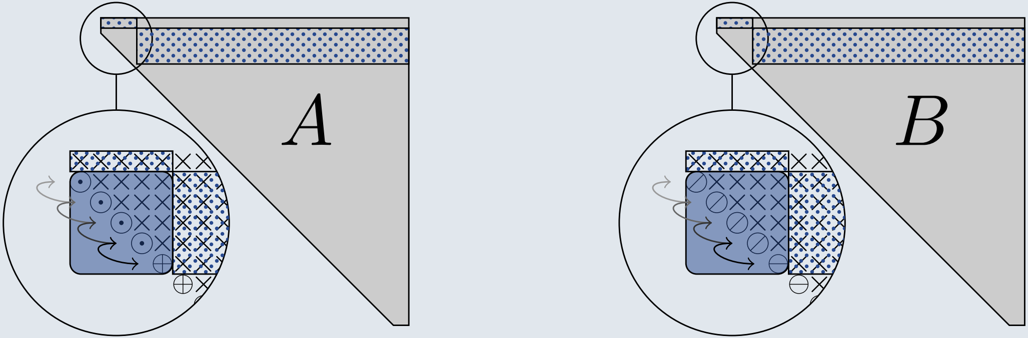
Multishift, multipole rational QZ

Motivation: make the rational QZ method competitive with state-of-the-art.

- Extension of the rational QZ method from Hessenberg to block Hessenberg pencils
- Shifts and poles of larger multiplicity
- Real-valued generalized eigenproblems in real arithmetic
- Swapping 2×2 blocks: Iterative refinement via Newton steps (C.-Mastronardi-Vandebril-Van Dooren, 2019).

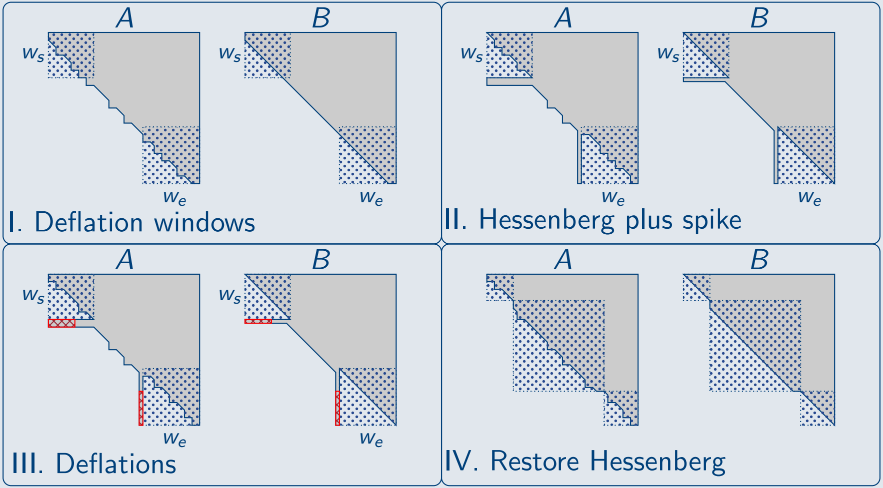
Multishift, multipole rational QZ

Batched operations



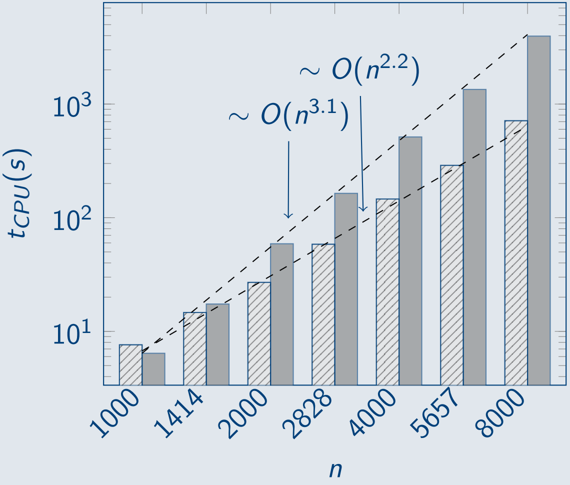
Multishift, multipole rational QZ

Aggressive early deflation (Braman-Byers-Mathias, 2002)



Multishift, multipole rational QZ

Numerical experiments with `libRQZ v0.1`



Conclusion

Conclusion and outlook

1. We have presented a novel interpretation of QR-type methods:

   $$\textit{bulge chasing} \leftrightarrow \textit{pole swapping}.$$

2. This results in a more general class of algorithms.
3. Convergence is determined by rational functions instead of polynomials.
4. Faster and more flexible eigensolvers.

References

Arnoldi, W. E. (1951). The principle of minimized iteration in the solution of the matrix eigenvalue problem. *Quart. Appl. Math.*, 9:17–29.

Berljafa, M., and Güttel, S. (2015). Generalized rational Krylov decompositions with an application to rational approximation. *SIMAX*, 36(2):894–916.

Braman, K., Byers R., and Mathias, R. (2002) The multishift QR algorithm. Part II: aggressive early deflation. *SIMAX*, 23(4):948–973.

Camps, D., Meerbergen, K., and Vandebril, R. (2019). A rational QZ method. *SIMAX* Accepted.

Camps, D., Meerbergen, K., and Vandebril, R. (2019). A multishift, multipole rational QZ method with aggressive early deflation. Submitted.

Camps, D., Mach, T., Vandebril, R., and Watkins, D. S. (2019). On pole-swapping algorithms for the eigenvalue problem. Submitted.

Camps, D. (2019). Pole swapping methods for the eigenvalue problem – rational QR algorithms. PhD thesis.

Camps, D., Mastronardi, N., Vandebril, R., and Van Dooren P. (2019). Swapping 2×2 blocks in the Schur and generalized Schur form. *JCAM* Available online.

References (cont.)

Elsner, L., and Watkins, D. S. Convergence of algorithms of decomposition type for the eigenvalue problem. *Lin. Alg. Appl.*, 143:19–47.

Francis, J. G. F. (1961-62) The QR transformation, a unitary analogue to the LR transformation — Part 1 and 2 *The Computer Journal*.

Kågström, B. (1993) A direct method for reordering eigenvalues in the generalized real Schur form of a regular matrix pair (A, B). In *Linear Algebra for Large Scale and Real-Time Applications*.

Moler, C. B., and Stewart, G. W. (1973) An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.*, 10(2):241–256.

Ruhe, A. (1998) Rational Krylov: A practical algorithm for large sparse nonsymmetric matrix pencils. *SISC*, 19(5):1535–1551

Van Dooren, P. (1981) A generalized eigenvalue approach for solving Riccati equations *SISC*, 2(2):121–135

Watkins, D. S. (1993) Some perspectives on the eigenvalue problem. *SIREV*, 35:430–471.