Tropical scaling of a Lagrange-type linearization for matrix polynomial eigenvalue problems

Marc Van Barel
Department of Computer Science
KU Leuven, Belgium

Françoise Tisseur
School of Mathematics
The University of Manchester, UK

Matrix polynomial eigenvalue problem

Given a matrix polynomial $P(z) \in \mathbb{C}[z]^{s \times s}$, the matrix polynomial eigenvalue problem looks for scalars λ and corresponding nonzero vectors v:

$$P(\lambda)v = 0$$

where $\lambda \in \mathbb{C}$ (eigenvalue) and $v \in \mathbb{C}^m$ (eigenvector), $v \neq 0$.

Solution methods:

- linearization $\longrightarrow$ generalized eigenvalue problem
- contour integration
  [Asakura, Sakurai, Tadano, Ikegami, Kimura 2010]
  [Beyn 2012] [VB, Kravanja 2016] [VB 2016]
- Ehrlich-Aberth iteration, ...
  [Bini, Noferini 2013]

Large difference in magnitude of eigenvalues

- Problem: large difference in the order of magnitude of the eigenvalues.
  Using block companion linearization $\longrightarrow$ inaccurate results.

- Solution: tropical scaling of block companion linearization [Gaubert, Sharify 2009] [Noferini, Sharify, Tisseur 2015]

  $$P(z) \text{ has degree } d \Rightarrow d \text{ tropical roots } \tau_i$$

  For each of the different tropical roots τ_i: scaled block companion linearization:

  $$\gamma_i P(\tau_i z).$$

  Disadvantage: For a degree d polynomial matrix $P(z)$, it can be necessary to use d different scaled linearizations.

- Question: can we obtain the same accuracy using only one linearization? Hence, by solving only one GEVP?

Solution: Tropical scaling of a Lagrange-type linearization

- For certain structured matrix pencils, (modified) QZ algorithm computes the (well-conditioned) eigenvalues with a high relative precision even when these eigenvalues have a very large difference in magnitude.
- Linearize the polynomial eigenvalue problem into a matrix pencil having this structure:
  - Lagrange-type linearization parametrized by the choice of the interpolation points σ_i: $P(\sigma_i)$
  - Choosing these interpolation points σ_i as well-separated tropical roots $\Rightarrow$ desired matrix pencil

Structured matrix pencil - QZ algorithm

Consider the matrix pencil $A - zB$ with

- order of the magnitude of the nonzero elements of A is 10^0
- the matrix B a block diagonal matrix
  - the first block can be the zero matrix (implying trivial eigenvalues at infinity for the matrix pencil)
  - the other diagonal blocks are themselves diagonal matrices or dense $s \times s$ matrices
  - the nonzero elements from different diagonal blocks in B can be very different in magnitude.

The large difference in magnitude in the elements of B leads generically to a large difference in magnitude for the generalized eigenvalues of the matrix pencil.

- If the first block of matrix B is zero, deflate the trivial infinite eigenvalues based on the first block column.
- Reduce the matrix pencil to generalized Hessenberg form.
- Compute the generalized Schur form using `zhgeqz.f`.

Numerical example 1: matrix pencil $A - zB$, 2×2 blocks

- $a_{i,j} = x + \mathrm{i}y$ with x and y pseudorandomly chosen from the standard normal distribution and $\mathrm{i} = \sqrt{-1}$
- a first zero block on the diagonal of matrix B
- the other 16 diagonal blocks are dense where the elements are chosen as for matrix A but scaled with factors $10^{-5}, 10^{-4}, \ldots, 10^{10}$ and permuted
- 100 samples

Structured matrix pencil - QZ algorithm

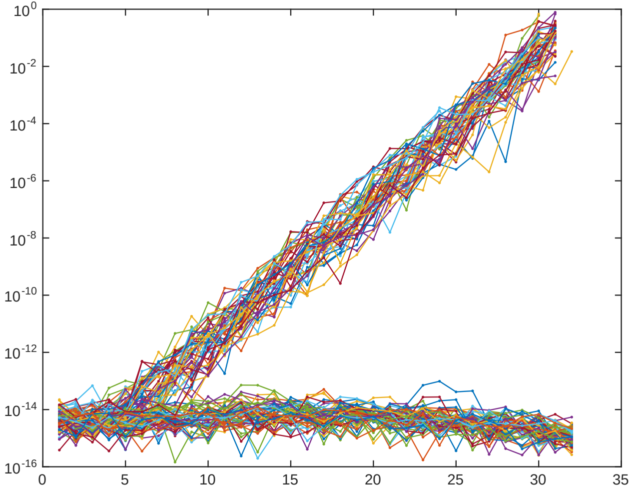


Figure: Relative error of computed eigenvalues of a structured matrix pencil using the unaltered QZ LAPACK implementation

Structured matrix pencil - QZ algorithm

- the eigenvalues have magnitudes of the order 10^{-10} up to 10^5
- for a lot of samples accurate eigenvalues over the whole scale
- there are samples such that only the smallest eigenvalues are relatively accurate while the accuracy decreases in the same degree as the magnitude increases
- at least one of the larger eigenvalues has been approximated by an eigenvalue at infinity
- unadapted version decides too fast that the corresponding computed eigenvalues are infinite

Modification of the LAPACK routine `zhgeqz.f`

- except for explicit infinite eigenvalues only finite eigenvalues are generated
- at two places in the fortran code, value of `BTOL` is replaced by the smallest positive nonzero floating point number

Structured matrix pencil - QZ algorithm

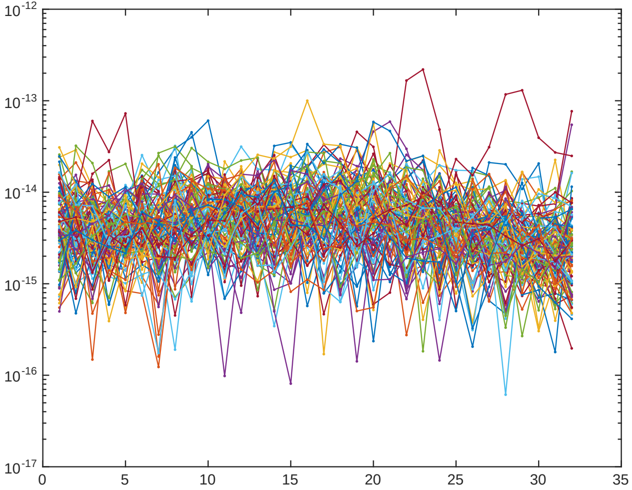


Figure: Relative error of computed eigenvalues of a structured matrix pencil using our adaptation of the QZ LAPACK implementation

[Amiraslani, Corless, Lancaster 2009] Let P be of degree d.

Take d points $\sigma_i \in \mathbb{C}$, $i = 1, 2, \ldots, d$ with corresponding barycentric weights $\beta_i = (\prod_{j \neq i}(\sigma_i - \sigma_j))^{-1}$.

Let the highest degree coefficient of P be denoted as P_d.

$$\left[\begin{array}{c|cccc} P_d & \beta_1 P(\sigma_1) & \beta_2 P(\sigma_2) & \cdots & \beta_d P(\sigma_d) \\ \hline -I_s & (z-\sigma_1) I_s & & & \\ -I_s & & (z-\sigma_2) I_s & & \\ \vdots & & & \ddots & \\ -I_s & & & & (z-\sigma_d) I_s \end{array}\right]$$

Proof:

$$\left[\begin{array}{c|cccc} P_d & \beta_1 P(\sigma_1) & \beta_2 P(\sigma_2) & \cdots & \beta_d P(\sigma_d) \\ \hline -I_s & (z-\sigma_1)I_s & & & \\ -I_s & & (z-\sigma_2)I_s & & \\ \vdots & & & \ddots & \\ -I_s & & & & (z-\sigma_d)I_s \end{array}\right] \left[\begin{array}{c} \ell(z)I_s \\ \hline \ell_1(z)I_s \\ \ell_2(z)I_s \\ \vdots \\ \ell_d(z)I_s \end{array}\right] = \left[\begin{array}{c} P(z) \\ \hline 0 \\ 0 \\ \vdots \\ 0 \end{array}\right]$$

with $\ell(z) = \prod_{i=1}^{d}(z-\sigma_i)$ and $\ell_i(z) = \ell(z)/(z-\sigma_i)$.

Lagrange-type linearization

Dividing the second block column by σ_1, the third block column by σ_2, ...: matrix pencil $zB - A$

$$\left[\begin{array}{c|cccc} P_d & \beta_1 P(\sigma_1)/\sigma_1 & \beta_2 P(\sigma_2)/\sigma_2 & \cdots & \beta_d P(\sigma_d)/\sigma_d \\ \hline -I_s & (z/\sigma_1 - 1) I_s & & & \\ -I_s & & (z/\sigma_2 - 1) I_s & & \\ \vdots & & & \ddots & \\ -I_s & & & & (z/\sigma_d - 1) I_s \end{array}\right]$$

- Except for the first block row A has nonzero elements equal to one.
- The matrix B is a (block-)diagonal matrix with diagonal nonzero blocks: $\sigma_i^{-1} I_s$.
- Idea: choose σ_i such that also the first block row of A contains elements whose magnitude is of order 1.

Suppose we have information on the magnitude of the modulus of the eigenvalues of the PEP via tropical roots.
Note that this information comes in "blocks" of s eigenvalues

tropical root τ_l with multiplicity μ_l $\longleftrightarrow$ $\mu_l s$ eigenvalues

Choice of the points σ_i based on the knowledge of the tropical roots τ_l and their "multiplicity" μ_l:

$$\{\sigma_i\}_{i=i',\dots,i'+(\mu_l-1)} = \{\mu_l \text{ roots of unity multiplied by } \tau_l\}$$

$$\begin{array}{ccccccc} \tau_1 & < & \tau_2 & < & \cdots & < & \tau_t \\ \mu_1 & & \mu_2 & & \cdots & & \mu_t \\ \sigma_1, \sigma_2, \dots \sigma_{\mu_1} & & \sigma_{\mu_1+1}, \dots, \sigma_{\mu_1+\mu_2} & & \cdots & & \sigma_{d-\mu_t+1}, \dots, \sigma_d \end{array}$$

Using this choice of the interpolation points σ_i gives us a matrix pencil $zB - A$ with top block row of A:

$$\begin{bmatrix} -P_d & -\frac{\beta_1 P(\sigma_1)}{\sigma_1} & -\frac{\beta_2 P(\sigma_2)}{\sigma_2} & \cdots & -\frac{\beta_d P(\sigma_d)}{\sigma_d} \end{bmatrix}$$

where the magnitude of the ith block is determined by the corresponding tropical root τ_l

$$\frac{\|P(\sigma_i)\||\beta_i|}{|\sigma_i|} \leq \frac{(d+1)}{\mu_l \Gamma_l}\|P_d\|$$

with separation parameter $\gamma < 1$

$$\gamma = \min\{ \quad \tilde{\gamma} \quad | \quad \tau_l \leq \tilde{\gamma}\tau_{l+1}, \quad l = 1, 2, \dots, t-1\}$$

$$\Gamma_l = \prod_{j=1, j\neq l}^{t} (1 - \gamma^{|l-j|})^{\mu_j}$$

Well-separated tropical roots

- Put an upper bound on the separation parameter $\gamma < 1$: $\tau_l \leq \gamma\tau_{l+1}$, e.g., γ should be less than 1/5.
- Hence, Γ_l becomes closer to one.

$$\Gamma_l = \prod_{j=1, j\neq l}^{t} (1 - \gamma^{|l-j|})^{\mu_j}$$

- The smaller value for γ has to be compensated by a relaxation in the condition for the tropical roots measured by a relaxation parameter ρ:

$$\frac{\|P(\sigma_i)\||\beta_i|}{|\sigma_i|} \leq \frac{\rho(d+1)}{\mu_l\Gamma_l}\|P_d\|$$

- notation $\mu_{i,j} = \sum_{l=i}^{j} \mu_l$.
- $\tau_l^{\mu_l} = \frac{\|P_{\mu_{1,l-1}}\|}{\|P_{\mu_{1,l}}\|}, \qquad l = 1, 2, \ldots, t$
- For a chosen value of the separation parameter $\gamma < 1$

$$\tau_l \leq \gamma \tau_{l+1}, \qquad l = 1, 2, \ldots, t-1$$

- With $\rho \geq 1$ the relaxation parameter

$$\begin{aligned} \tau_l^{\mu_{1,l-1}} \|P_{\mu_{1,l-1}}\| &= \tau_l^{\mu_{1,l}} \|P_{\mu_{1,l}}\| \\ &\geq \rho^{-1} \tau_l^{j} \|P_j\|, \qquad j \notin \{\mu_{1,l-1}, \mu_{1,l}\} \end{aligned}$$

- The well-known algorithm to compute the classical tropical roots based on the convex hull of the points $(k, \log(\|P_k\|))$, $k = 0, 1, \ldots, d$ can be extended to compute well-separated tropical roots τ_l.

- Compute well-separated tropical roots τ_l such that

$$\frac{\|P(\sigma_i)\||\beta_i|}{|\sigma_i|} \le \frac{\rho(d+1)}{\mu_l \Gamma_l}\|P_d\| = \mathcal{O}(1)\|P_d\|.$$

- Scale the polynomial matrix $\beta P(z)$ such that the norm of the highest degree coefficient is equal to one.
- Construct the Lagrange-type linearization $zB - A$ with the top block row of A equal to

$$\begin{bmatrix} -P_d & -\frac{\beta_1 P(\sigma_1)}{\sigma_1} & -\frac{\beta_2 P(\sigma_2)}{\sigma_2} & \dots & -\frac{\beta_d P(\sigma_d)}{\sigma_d} \end{bmatrix}$$

- (Refine the balancing/scaling, e.g., using the balancing strategy of [Lemonnier, Van Dooren 2006].)
- Apply the modified QZ algorithm excluding nontrivial roots at infinity.

After scaling/balancing, the norms of the blocks in the two matrices A and B of the matrix pencil $A - \lambda B$ have orders of magnitude:

$$A = \left[\begin{array}{c|cccc} 10^0 & 10^0 & 10^0 & \cdots & 10^0 \\ \hline 10^0 & 10^0 & & & \\ 10^0 & & 10^0 & & \\ \vdots & & & \ddots & \\ 10^0 & & & & 10^0 \end{array}\right]$$

$$B = \left[\begin{array}{c|cccc} & & & & \\ \hline & \sigma_1^{-1} & & & \\ & & \sigma_2^{-1} & & \\ & & & \ddots & \\ & & & & \sigma_d^{-1} \end{array}\right]$$

Backward error

- We know that the QZ algorithm for solving the generalized eigenvalue problem is backward stable, that is, it solves a nearby problem $\lambda(B+\Delta B)-(A+\Delta A)$ with

$$\|\Delta A\|_2 \leq p_A(n)\epsilon_{\text{mach}}\|A\|_2, \qquad \|\Delta B\|_2 \leq p_B(n)\epsilon_{\text{mach}}\|B\|_2,$$

  where ϵ_{mach} is the machine precision and $p_A(n)$, $p_B(n)$ are polynomial in n.
- Because $\|B\|_2$ is of the order σ_1^{-1}, when the backward error $\|\Delta B\|_2$ would not be structured, this error would change the diagonal elements σ_i^{-1} significantly when $\sigma_i^{-1} \ll \sigma_1^{-1}$.
- However, the backward error ΔB on B is structured and follows the graded structure of the inverse of the σ_i.
- This allows to compute the (well-conditioned) eigenvalues (eigenvectors) with high relative precision, even when they differ several orders of magnitude.

Exp. 01: Classical versus well-separated tropical roots

Upper-bound for $\|P(\sigma_i)\||\beta_i|/|\sigma_i|$

$$\frac{\|P(\sigma_i)\||\beta_i|}{|\sigma_i|} \leq \frac{\rho(d+1)}{\mu_l \Gamma_l}\|P_d\|.$$

For well-separated roots $\|P(\sigma_i)\||\beta_i|/|\sigma_i| \lesssim 2$.

- Take as tropical roots $\tau_l = 10^{2(l-1)}$, $l = 1, 2, \ldots, 10$ having multiplicity $\mu_l = 1$.
- The polynomial matrix coefficients are chosen as scaled random unitary matrices having such a norm that we obtain the corresponding tropical roots.
- The values $\|P(\sigma_i)\||\beta_i|/|\sigma_i|$ for 100 samples for $P(z)$.

Exp. 01: Classical versus well-separated tropical roots

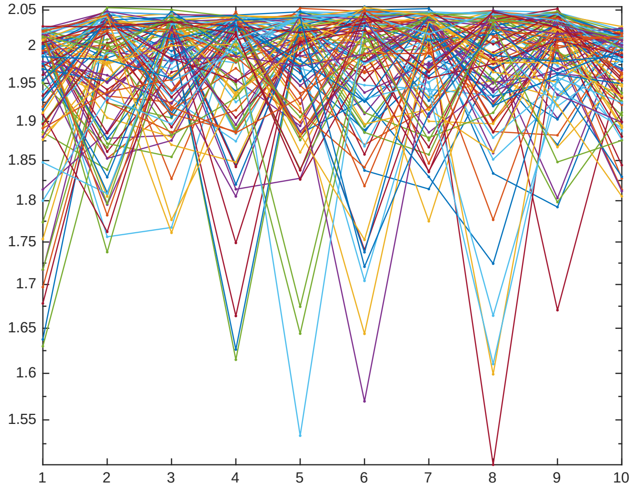
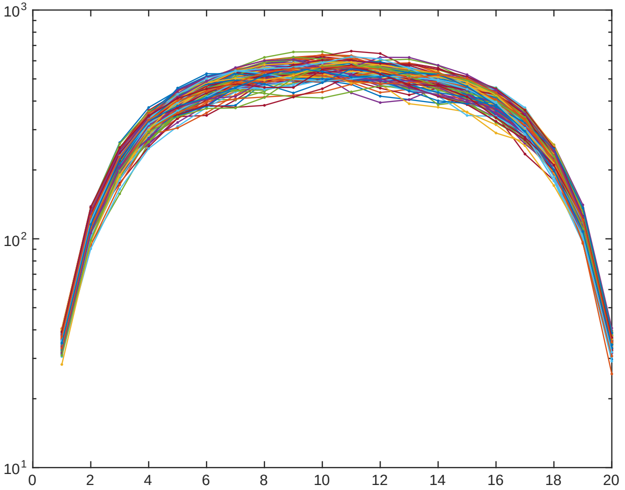
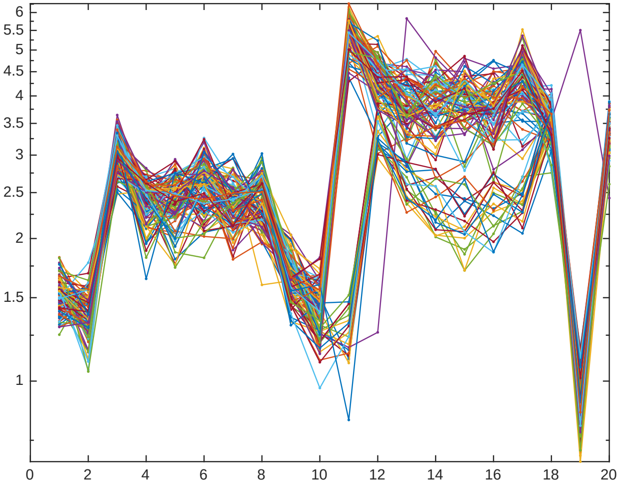


Figure: The values $\|P(\sigma_i)\||\beta_i|/|\sigma_i|$, $i = 1, 2, \ldots, d$ for well-separated tropical roots

In this case, we choose the tropical roots not well-separated, e.g., $\tau_l = (1.5)^{l-1}$, $l = 1, 2, \ldots, 20$. Note that $1.5^{19} \approx 2.2 \cdot 10^3$.



Using well-separated tropical values with a separation parameter $\gamma = 5^{-1}$.



Example 4 of [Gaubert, Sharify 2009]: Matlab-code:

```
s = 8; d = 10;
scaling = [-5,-2,-3,-4,2,0,3,-3,4,2,5];
for i = 1:d+1
    P{i} = randn(s) * 10^scaling(i);
end
```

The backward error is computed as [Tisseur 2000]:

$$\frac{\|P(\tilde{\lambda})^{-1}\|_2^{-1}}{\sum_{i=0}^{d} |\tilde{\lambda}|^i \|P_i\|_2} = \frac{\sigma_{\min}(P(\tilde{\lambda}))}{\sum_{i=0}^{d} |\tilde{\lambda}|^i \|P_i\|_2}.$$

The maximum backward error is plotted for 100 samples using the two types of tropical roots.

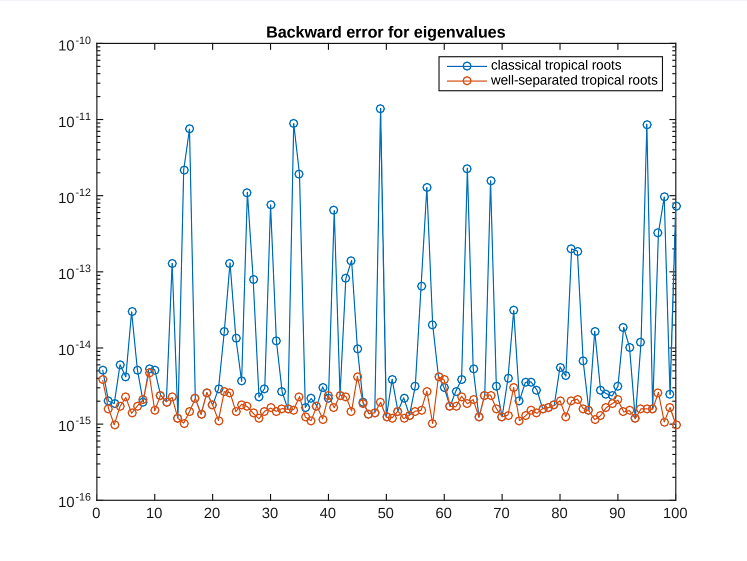


Figure: Backward error for 100 samples of Experiment 3 using classical tropical roots and well-separated tropical roots

Exp. 03: Classical versus well-separated tropical roots

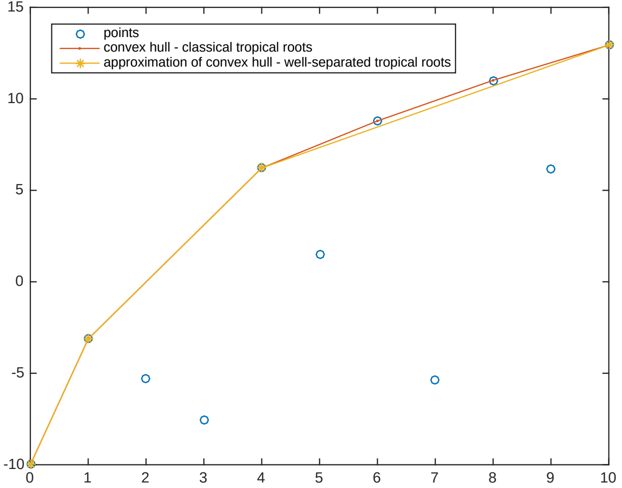


Figure: Convex hull of the points $(i, \log \|P_i\|)$, $i = 0, 1, \ldots, d$ and approximation of this convex hull

Exp. 04: Polynomial eigenvalue problems from NLEVP

For all square problems from NLEVP having size s less than or equal to 300, apply the following five algorithms:

1. tropical scaling for the Lagrange linearization using the classical tropical roots;
2. tropical scaling for the Lagrange linearization using well-separated tropical roots;
3. `polyeig` of MATLAB;
4. `quadeig` (if the degree is equal to 2);
5. tropical scaling for the block companion linearization using the classical tropical roots.

Exp. 04: Polynomial eigenvalue problems from NLEVP

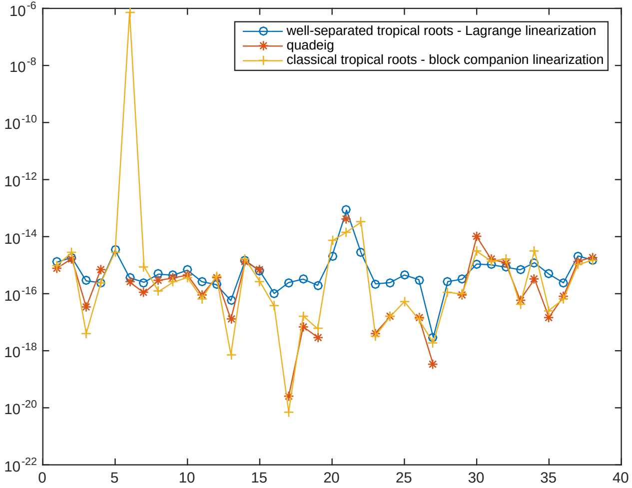


Figure: Backward error for algorithms 1, 4 and 5

Exp. 04: Polynomial eigenvalue problems from NLEVP

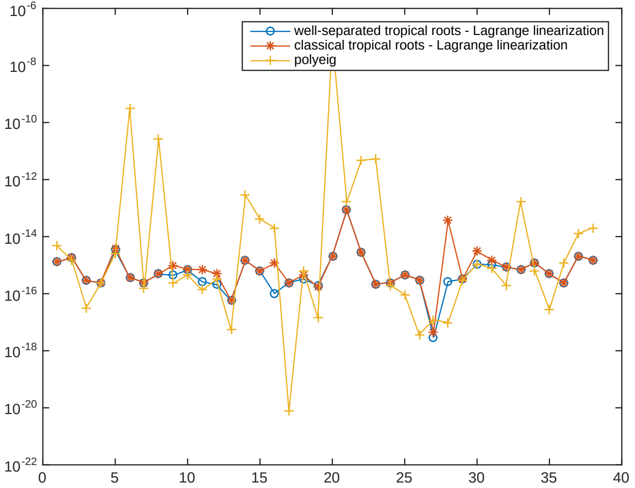


Figure: Backward error for algorithms 1, 2 and 3

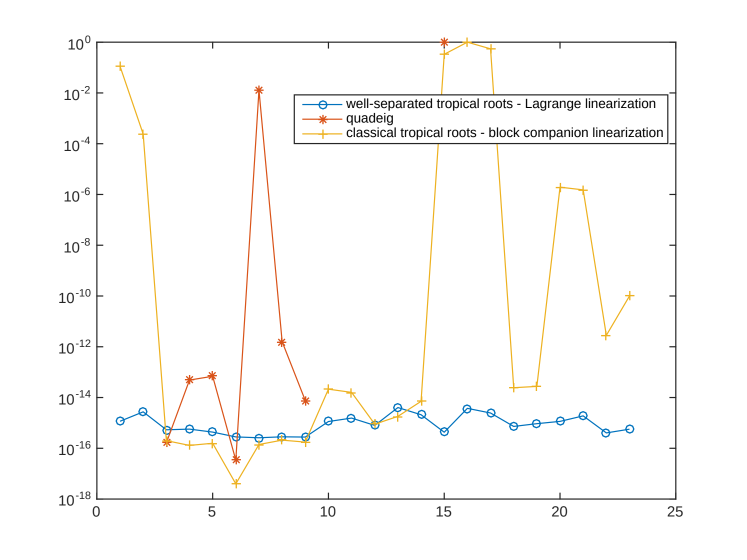


Figure: Backward error for algorithms 1, 4 and 5

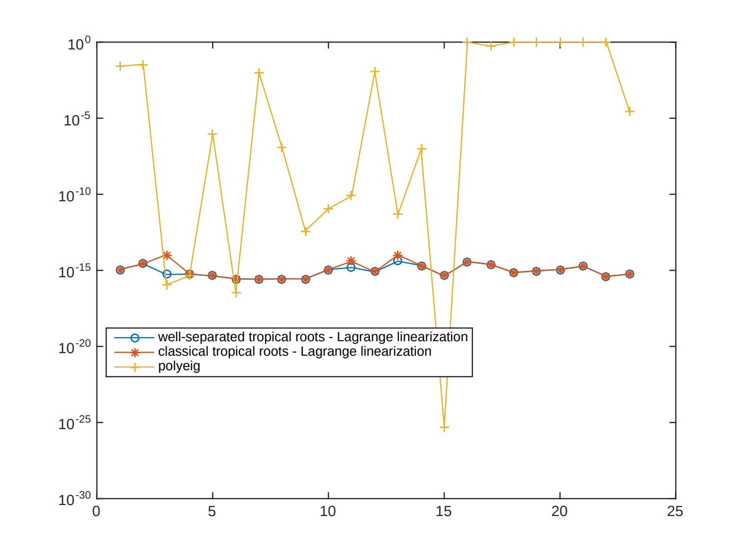


Figure: Backward error for algorithms 1, 2 and 3

A. Amiraslani, R. M. Corless, and P. Lancaster.
Linearization of matrix polynomials expressed in polynomial bases.
IMA Journal of Numerical Analysis, 29:141–157, 2009.

J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, and K. Kimura.
A numerical method for polynomial eigenvalue problems using contour integrals.
Japan J. Indust. Appl. Math., 27:73–90, 2010.

W.-J. Beyn.
An integral method for solving nonlinear eigenvalue problems.
Linear Algebra and its Applications, 436:3839–3863, 2012.

D. A. Bini and V. Noferini.
Solving polynomial eigenvalue problems by means of the ehrlich-aberth method.
Linear Algebra and its Applications, 439:1130–1149, 2013.

D. A. Bini, V. Noferini, and M. Sharify.
Locating the eigenvalues of matrix polynomials.
SIAM Journal on Matrix Analysis and Applications, 34(4):1708–1727, 2013.

H.-Y. Fan, W.-W. Lin, and P. Van Dooren.
Normwise scaling of second order polynomial matrices.
SIAM Journal on Matrix Analysis and Applications, 26(1):252–256, 2004.

S. Gaubert and M. Sharify.
Tropical scaling of polynomial matrices.
In *Positive Systems*, volume 389 of *Lecture Notes in Control and Inf. Sci.*, pages 291–303. Springer-Verlag, Berlin, 2009.

D. Lemonnier and P. Van Dooren.
Balancing regular matrix pencils.
SIAM Journal on Matrix Analysis and Applications, 28(1):253–263, 2006.

V. Noferini, M. Sharify, and F. Tisseur.
Tropical roots as approximations to eigenvalues of matrix polynomials.
SIAM Journal on Matrix Analysis and Applications, 236(1):138–157, 2015.

M. Sharify.
Scaling Algorithms and Tropical Methods in Numerical Matrix Analysis: Applica- tion to the Optimal Assignment Problem and to the Accurate Computation of Eigenvalues.
PhD thesis, Ecole Polytechnique, Palaiseau, France, 2011.

F. Tisseur.
Backward error and condition of polynomial eigenvalue problems.
Linear Algebra and its Applications, 309(1-3):339–361, 2000.

M. Van Barel.
Designing rational filter functions for solving eigenvalue problems by contour integration.
Linear Algebra and its Applications, 502:346–365, 2016.

M. Van Barel and P. Kravanja.
Nonlinear eigenvalue problems and contour integrals.
Journal of Computational and Applied Mathematics, 292:526 – 540, 2016.