Optimizing Least-Squares Rational Filters for Solving Interior Eigenvalue Problems

Jan Winkelmann, Edoardo Di Napoli
{winkelmann,dinapoli}@aices.rwth-aachen.de

Aachen Institute for Advanced Studies in Computational Engineering Science

July 6, 2016

Filtered Subspace Iteration

Consider the interior hermitian eigenvalue problem

$$Ax = \lambda x, \text{ with } A = A^H \in \mathbb{C}^{n \times n} \text{ sparse and } \lambda \in \mathcal{I} = [\lambda_{min}, \lambda_{max}]$$

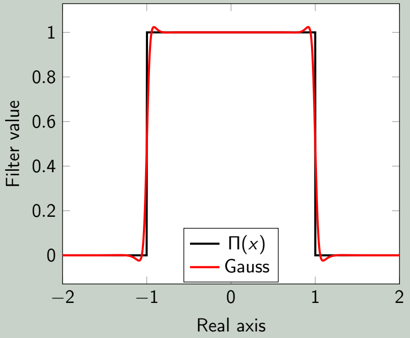
Can be extended to generalized interior (non-hermitian) eigenproblems

Algorithm Sketch

1. Filter: $\hat{X}_{n+1} := \rho(A) X_n$
2. Rayleigh-Ritz:
   - $[Q, \Lambda] := \text{eig}(\hat{X}_{n+1}^H \cdot A \cdot \hat{X}_{n+1})$
   - $X_{n+1} := \hat{X}_{n+1} \cdot Q$

- X_0 initial guess of eigenvectors
- $\rho(A)$ spectral projector

Example (Gauss filter)



Rational filter methods

Classic rational filter methods use numerical integration of a contour

$$\rho(A)X = \frac{1}{2\pi\imath}\oint_{\mathcal{C}}(zI - A)^{-1}X\mathrm{d}z \approx \sum_{i=1}^{N}\omega_i(z_iI - A)^{-1}X$$

We can use any function that yields cheap system solves. For Example:

$$\rho(A) = \sum_{i=1}^{N}\alpha_i(z_iI - A)^{-1}$$

Rational filter methods

Classic rational filter methods use numerical integration of a contour

$$\rho(A)X = \frac{1}{2\pi\imath}\oint_{\mathcal{C}}(zI - A)^{-1}X\mathrm{d}z \approx \sum_{i=1}^{N}\omega_i(z_iI - A)^{-1}X$$

We can use any function that yields cheap system solves. For Example:

$$\rho(A) = \sum_{i=1}^{N}\alpha_i(z_iI - A)^{-1}$$

Previous Work

- SS-Method [1]
- FEAST [2]
- Zolotarev Filters [3]
- Least-Squares [4,5]

---

[1] Sakurai, Sugiura. Journal of Computational and Applied Mathematics; 2003

[2] Polizzi. Physical Review B; 2009

[3] Güttel, et.al. SIAM Journal on Scientific Computing; 2015

[4] Barel. Linear Algebra and its Applications; 2015

[5] Xi, Saad, Y. Preprint; 2015

Least-Squares based filters

The idea:

$$\min_{\alpha_i, z_i \; 1 \leq i \leq N} \int_{-\infty}^{\infty} \left| \prod(x) - \sum_{i=1}^{N} \frac{\alpha_i}{z_i - x} \right|^2 \mathrm{d}x$$

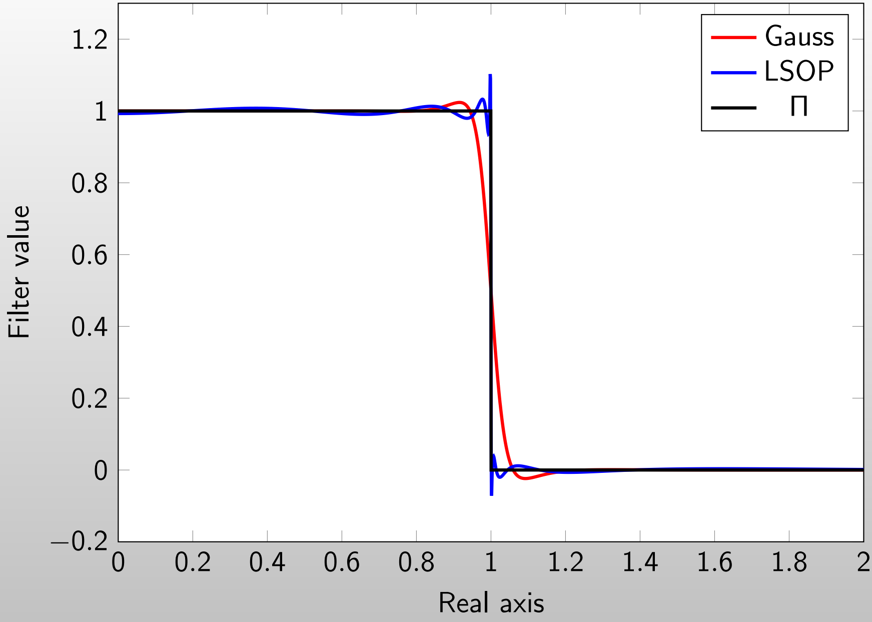
Note that:

- $\prod(x)$ is the the indicator function on $[-1, 1]$.
- Coefficients α_i, and poles z_i are complex-valued.
- We optimize coefficients and poles of the rational function.

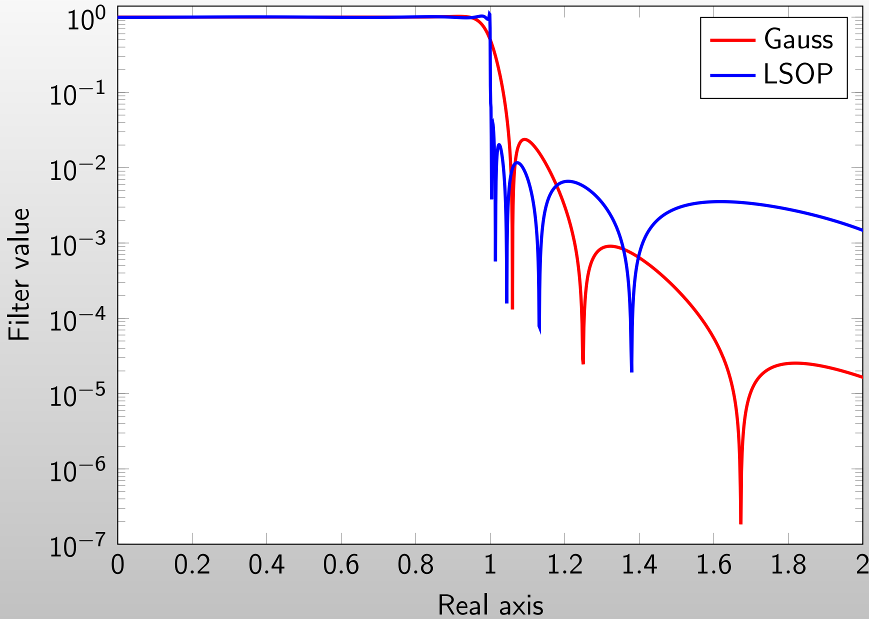
Potential Benefits:

- Faster convergence
- Increased flexibility
- Problem specific filters

First Results



First Results



Our Goals

- Filtered Subspace Iteration is sensitive to the spectrum
- Appropriate choice of subspace size and degree required
- Different filter types have different strengths
- Even with some prior knowledge:
  Optimal choice of parameters is difficult.
- This is a difficult problem $\Rightarrow$ We will not change this

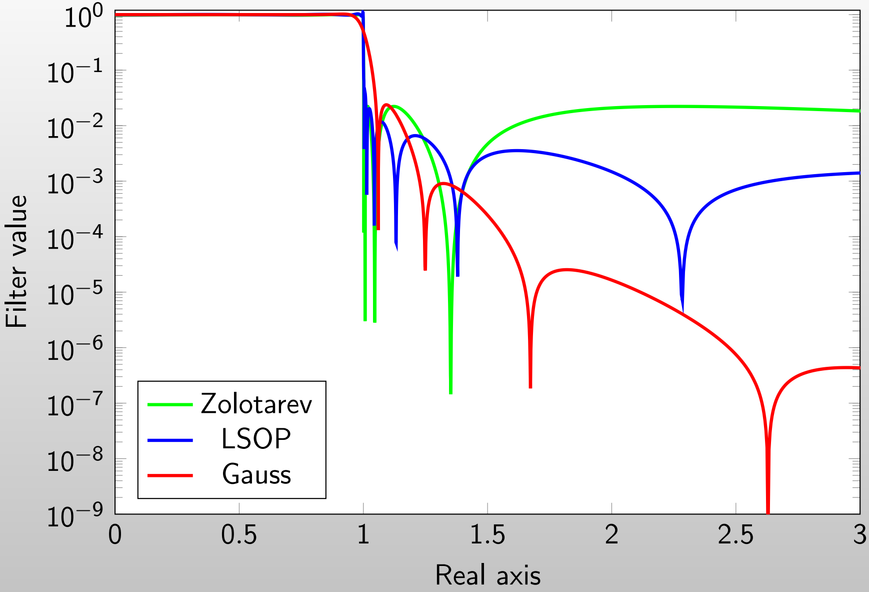
- The filter should be "smart"
- Take advantage of the spectrum
- Ideally, this would work automatically

State of the Art Filters

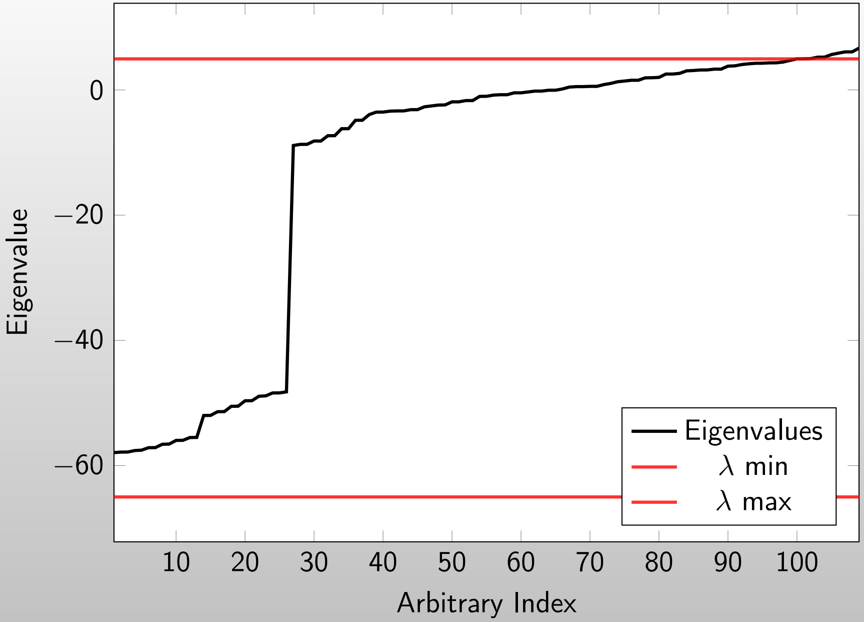
- Gauss Filter
  - From Gauss(-Legendre) Quadrature
  - Benefits from larger search subspace (≈ 1.5)
- Zolotarev Filter
  - From Cauer, Elliptical, or Zolotarev Filters
  - Does not benefit from larger subspaces
  - May have better convergence or load-balancing

- Choosing optimal filter and parameters is difficult
- FEAST 3.0 as Benchmark tool
- All Timings are for single threaded execution
  - Maximum of 40 iterations
  - Termination criterion: Residual $\leq 10^{-14}$
  - Filter–type, –degree and search subspace as indicated

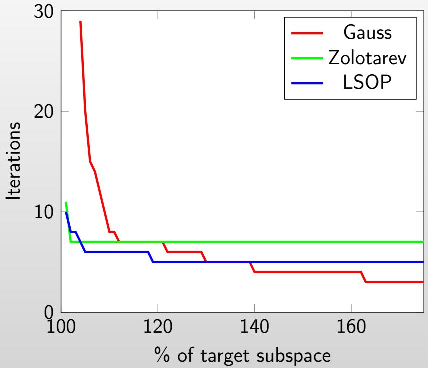
Filter "Comparison", 4 Poles per Quadrant

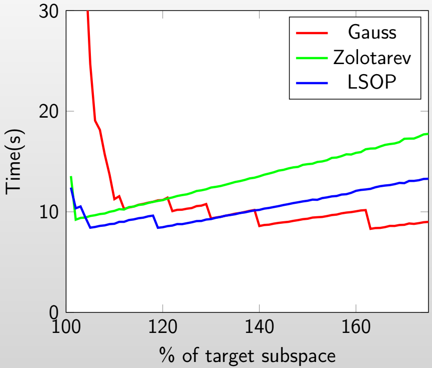


2D FEM: CNT Spectrum Visualization



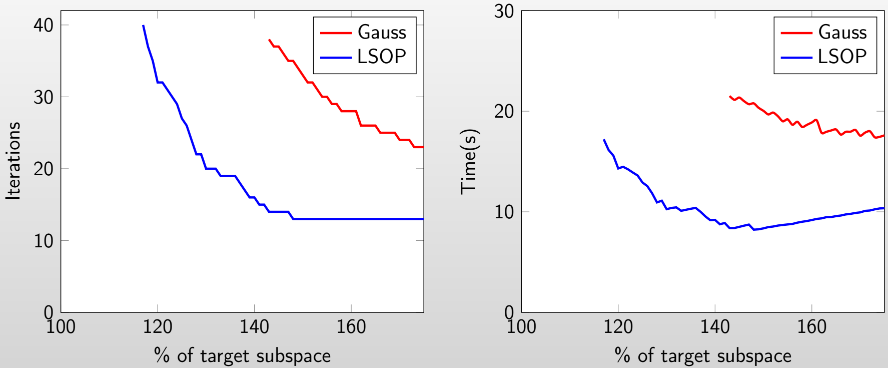
2D FEM: CNT, 4 Poles per Quadrant





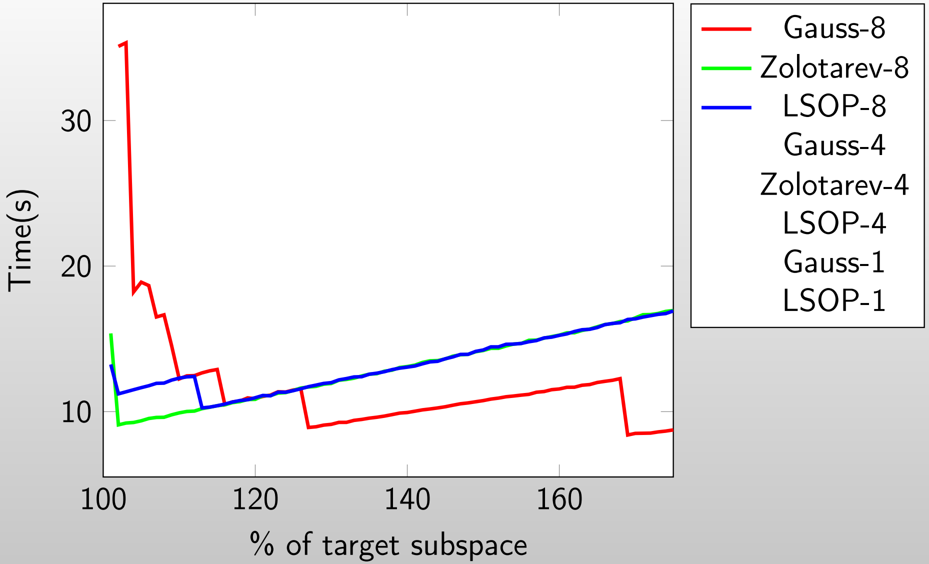
- Gauss filter benefits from larger subspaces
- Least-Squares filter requires less iterations than Zolotarev
  Somewhat benefits from larger subspace

2D FEM: CNT, 1 Pole per Quadrant

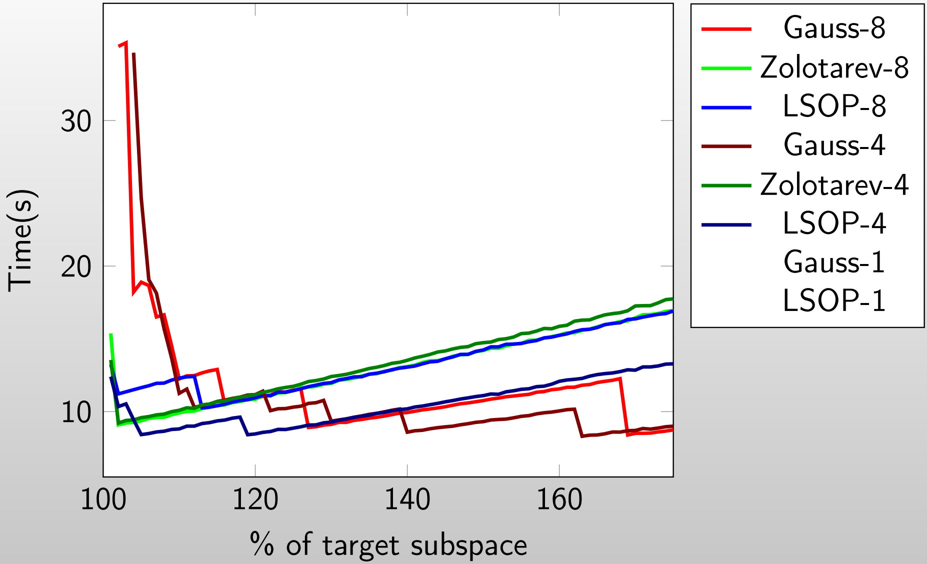


- Zolotarev does not converge within 40 iterations

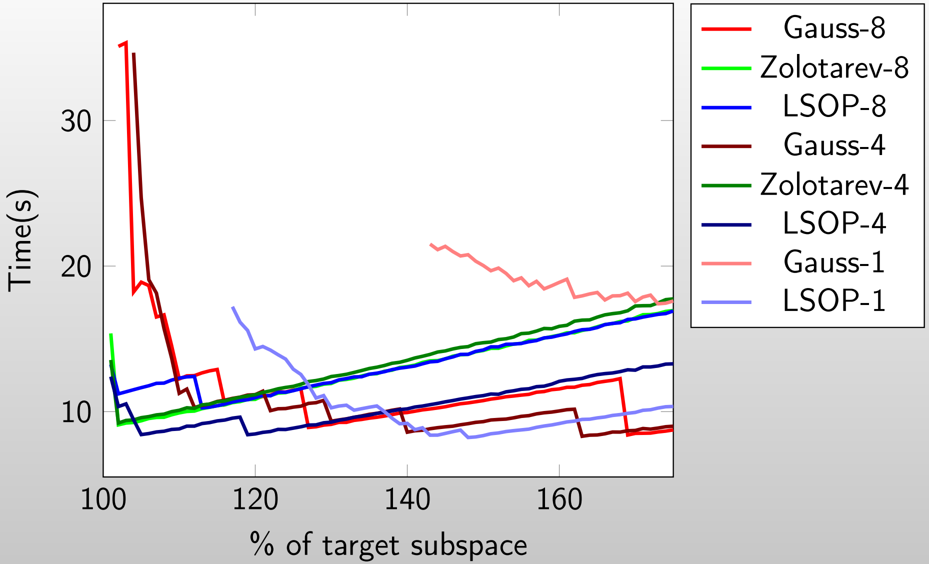
2D FEM: CNT, Complete Timings



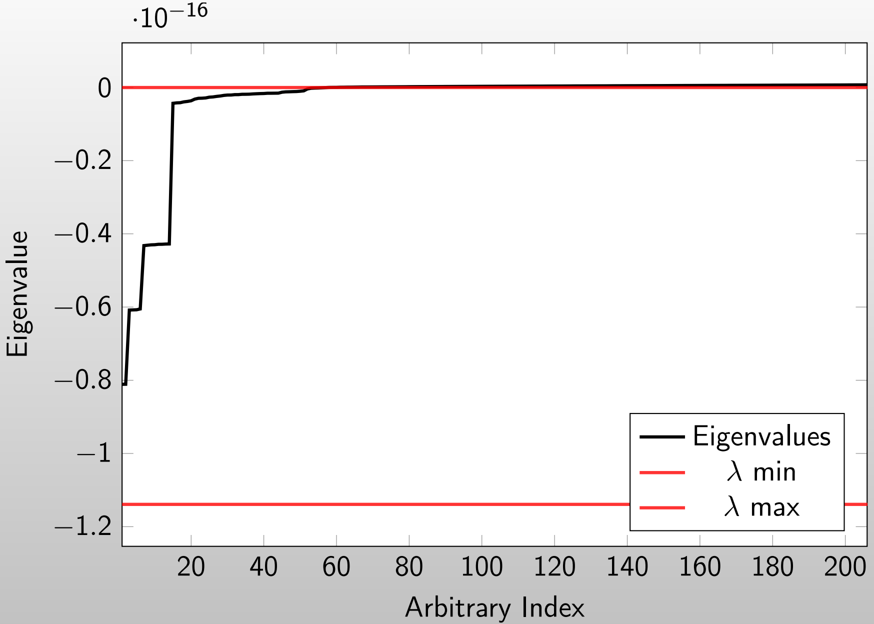
2D FEM: CNT, Complete Timings



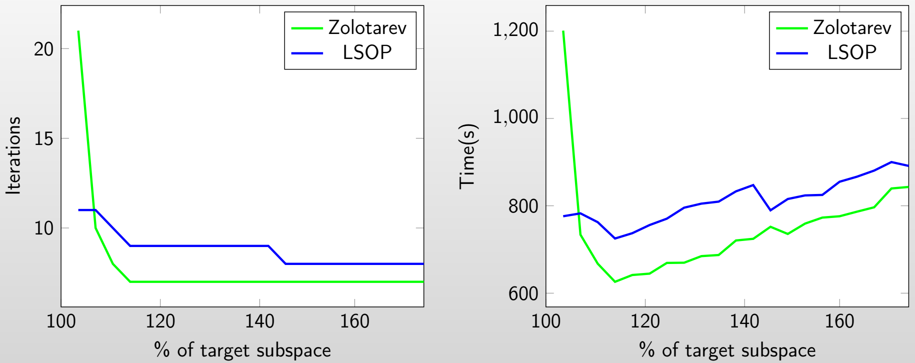
2D FEM: CNT, Complete Timings



3D FEM: Caffeinep2 Spectrum Visualization

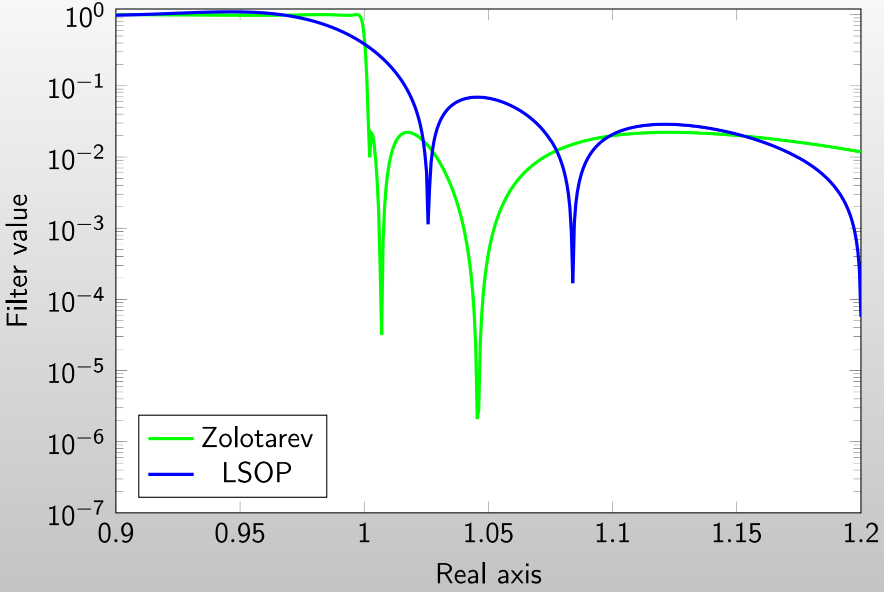


3D FEM: Caffeinep2, 4 Poles per Quadrant



- Gauss does not converge
  (recall the large cluster at the edge of the interval)
- Zolotarev outperforms Least-Squares filters for any number of poles

Recall: Zolotarev vs LSOP Filters



Wrap-up: Non Problem-Specific Filters

- Results for unconstrained Least-Squares filters
- Works well even with low degrees
- This is not the take away today

The take-away should be:

1. Gather spectrum information (separate topic)
2. Use appropriate filter and parameters
3. Facilitated via custom Least-Squares filters

- Least-Squares filters are very flexible
  Constraints for asymmetry, steepness, decay possible
- Fast generation of filters possible

Approximating $\Pi(x)$

$$\min_{\gamma_i, z_i \; 1 \le i \le N/4} \int_{-\infty}^{\infty} \left| \prod(x) - f(x,z,\gamma) \right|^2 \mathrm{d}x, \text{ with}$$

$$f(x,z,\gamma) = \sum_{i=1}^{N/4} \frac{\gamma_i}{z_i - x} + \frac{\overline{\gamma_i}}{\overline{z_i} - x} - \frac{\gamma_i}{z_i + x} - \frac{\overline{\gamma_i}}{\overline{z_i} + x}$$

- Reflection Symmetry: $f(-x,z,\gamma) = f(x,z,\gamma)$
- Complex Conjugation: $\overline{f(x,z,\gamma)} = f(x,z,\gamma)$

- Symmetries force evenness and real-ness
- Analogous to existing integration rules
- For simplicity we disallow poles on the axes
- Asymmetric filters lack reflection symmetry

Approximating $\Pi(x)$

$$\min_{\gamma_i, z_i \; 1 \le i \le N/4} \int_{-\infty}^{\infty} \left| \prod(x) - f(x,z,\gamma) \right|^2 \mathrm{d}x, \text{ with}$$

$$f(x,z,\gamma) = \sum_{i=1}^{N/4} \frac{\gamma_i}{z_i - x} + \frac{\overline{\gamma_i}}{\overline{z_i} - x} - \frac{\gamma_i}{z_i + x} - \frac{\overline{\gamma_i}}{\overline{z_i} + x}$$

- Reflection Symmetry: $f(-x,z,\gamma) = f(x,z,\gamma)$
- Complex Conjugation: $\overline{f(x,z,\gamma)} = f(x,z,\gamma)$

- Symmetries force evenness and real-ness
- Analogous to existing integration rules
- For simplicity we disallow poles on the axes
- Asymmetric filters lack reflection symmetry

We require cheap computation of:

- Least-Squares Residual
- Gradients w.r.t. α_i, z_i

Calculating Least-Squares Residuals

$$\int_{-\infty}^{\infty} w(x) \left| \prod(x) - \sum_{i=1}^{N/4} \frac{\gamma_i}{z_i - x} + \frac{\overline{\gamma_i}}{\overline{z_i} - x} - \frac{\gamma_i}{z_i + x} - \frac{\overline{\gamma_i}}{\overline{z_i} + x} \right|^2 \mathrm{d}x$$

- An appropriate $w(x)$ simplifies the calculation:

$$w(x) = \begin{cases} 0, & \text{if } |x| > a \\ \beta, & \text{if } |x| \leq 1 \\ 1, & \text{else} \end{cases}$$

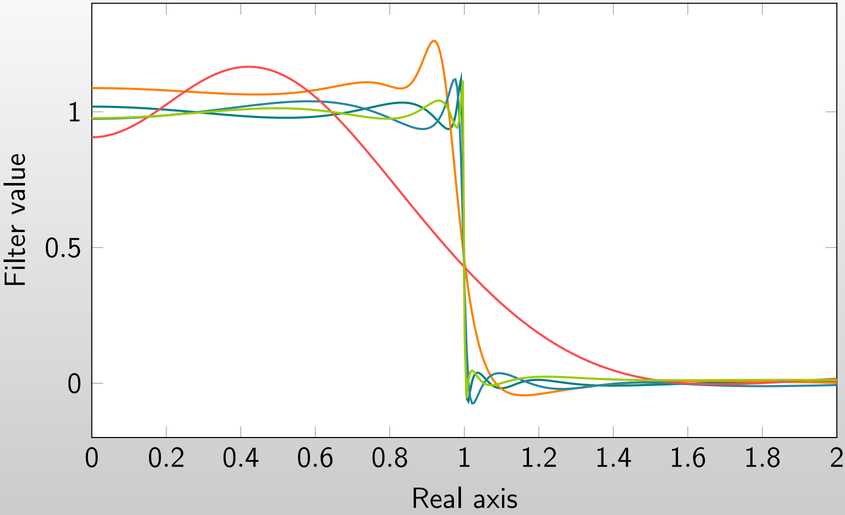
- With $\beta \approx 0.4$, and $a \approx 5$.
- Exact 1.0 in $[-1, 1]$ not required
  $\Rightarrow$ Small β increases sharpness at -1,1
- 'Cut-off' after $a \Rightarrow$ Definite integrals
- We can provide a (matrix) form for the residual

Optimization

- We can derive Gradients w.r.t. z_i, and α_i
- Use of symmetries $\Rightarrow$ Cheaper Gradients
- Fast filter generation via Levenberg-Marquardt
- Details in forthcoming publication

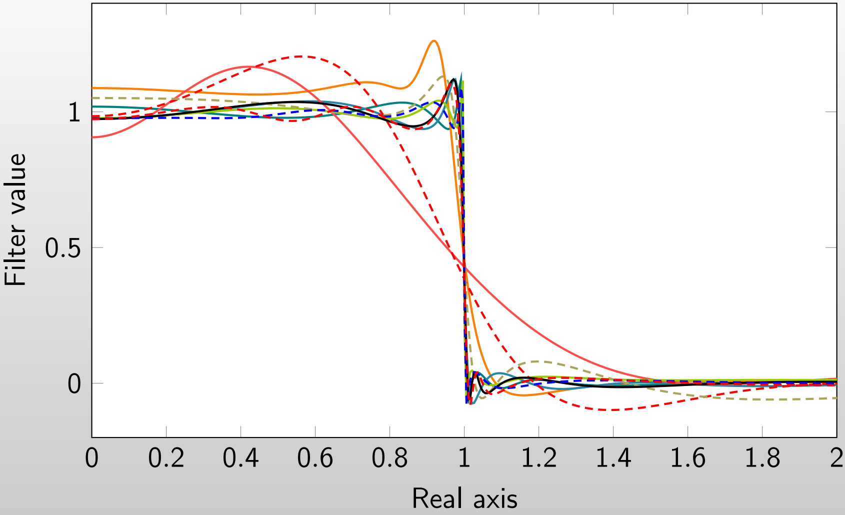
- Descent methods require a starting position
- Very relevant for non-convex optimization

Local Minima for 4 Poles per Quadrant



- The filter is non-convex in the poles z_i
- Resulting filter depends on the starting position
- Many approaches do not yield good results

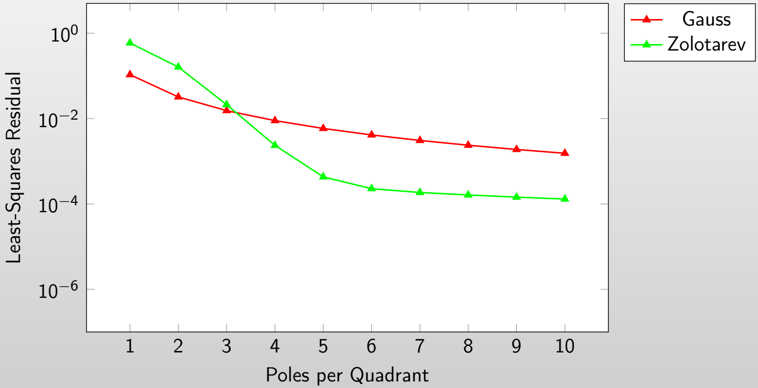
Local Minima for 4 Poles per Quadrant



- The filter is non-convex in the poles z_i
- Resulting filter depends on the starting position
- Many approaches do not yield good results

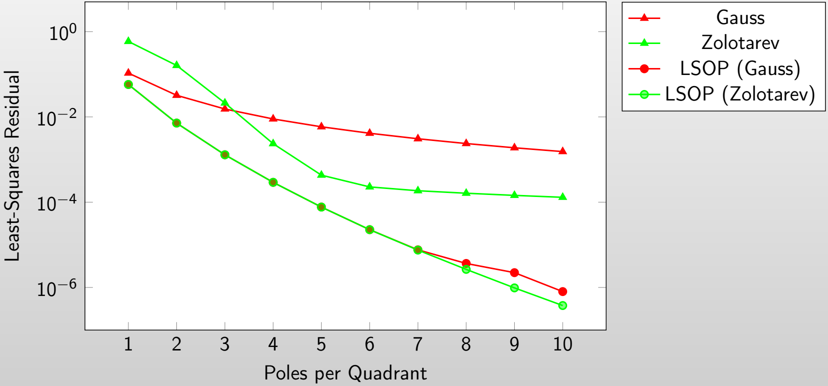
Use Existing Methods as Starting Points

- Existing filters already have good residuals



Use Existing Methods as Starting Points

- Existing filters already have good residuals



- $\Rightarrow$ Good results as starting positions
- Both produce the same results for ≤ 7 poles per quadrant
- So far, these are the best residuals we know of

Conclusion

- Filters are important for contour-based eigensolvers
- No single filter type is best

We present Least-Squares based filters:

- Rich framework for filter generation
- Potential for custom filters
- Competetive results for unconstrained Least-Squares filters
- Fast generation of filters

- No theoretical insights on generated filters
  (i.e. worst-case convergence ratio)
- Systematic comparisons of filters is tricky

Future Work

- Parallelism and convergence of linear system solves
- Further details in forthcoming publication:
- More constrained optimization of filters
- Systematic comparisons of filters

Thank you for your attention
Questions?

Deutsche Forschungsgemeinschaft
DFG

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

Constrained LSOP

- Poles near the real axis $\Rightarrow$ Almost real shift z_i of Matrix

$$\left(\sum_{i=1}^{N/4} \frac{\gamma_i}{I z_i - A} + \frac{\overline{\gamma_i}}{I \overline{z_i} - A} - \frac{\gamma_i}{I z_i + A} - \frac{\overline{\gamma_i}}{I \overline{z_i} + A} \right) X = B$$

- The shifted $(z_i I - A)$ may become (almost) singular

Constrained LSOP

- Poles near the real axis $\Rightarrow$ Almost real shift z_i of Matrix

$$\left(\sum_{i=1}^{N/4} \frac{\gamma_i}{Iz_i - A} + \frac{\overline{\gamma_i}}{I\overline{z_i} - A} - \frac{\gamma_i}{Iz_i + A} - \frac{\overline{\gamma_i}}{I\overline{z_i} + A}\right) X = B$$

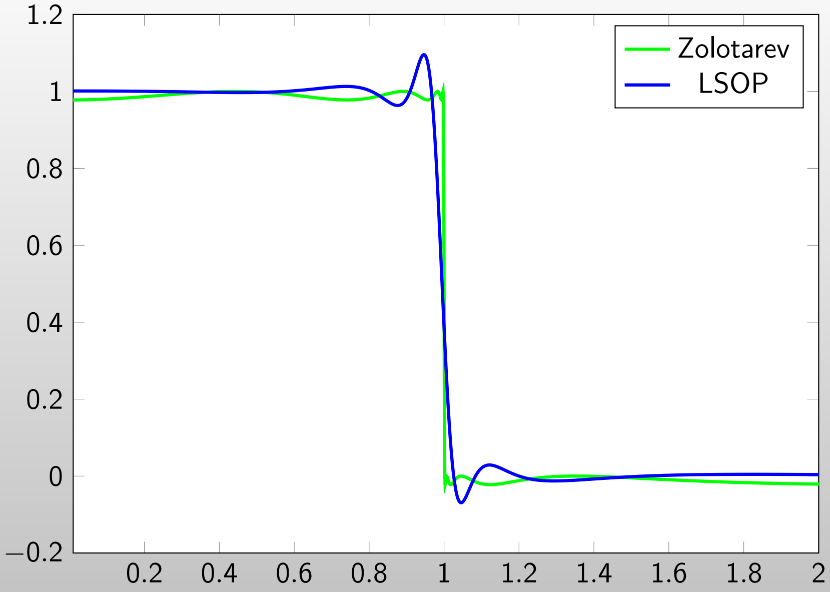
- The shifted $(z_i I - A)$ may become (almost) singular
- Larger imaginary part on the poles can prevent this "problem"

$$\min_{\gamma_i, z_i \; 1 \le i \le N/4} F(x), \text{ such that}$$

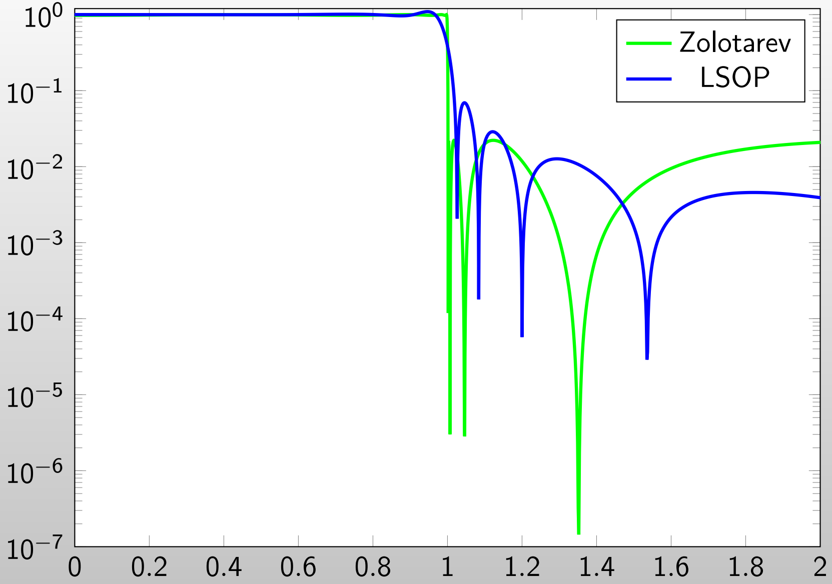
$$\Im(z_i) \ge 0.1, \;\; 1 \le i \le N/4$$

$$F(x) = \int_{-\infty}^{\infty} w(x) \left| \prod(x) - \sum_{i=1}^{N/4} \frac{\gamma_i}{z_i - x} + \frac{\overline{\gamma_i}}{\overline{z_i} - x} - \frac{\gamma_i}{z_i + x} - \frac{\overline{\gamma_i}}{\overline{z_i} + x} \right|^2 \mathrm{d}x$$

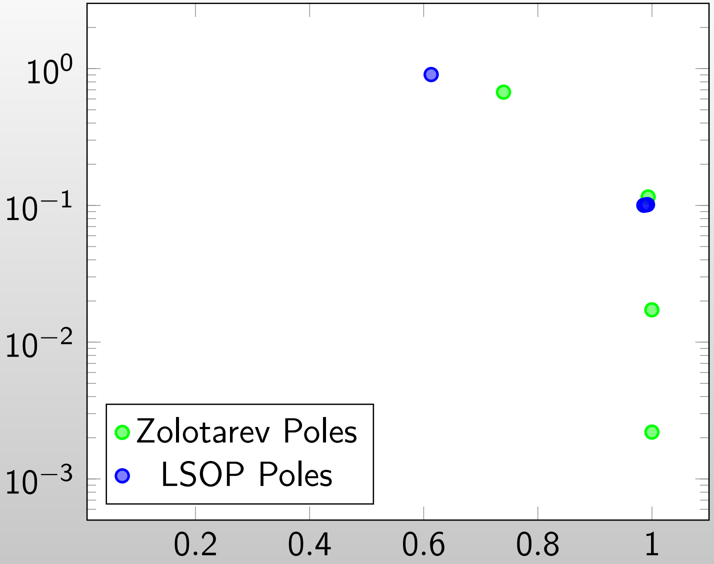
Zolotarev vs constrained LSOP, 4 Poles per Quadrant



Zolotarev vs constrained LSOP, 4 Poles per Quadrant



Zolotarev vs constrained LSOP, 4 Poles per Quadrant



CAFF timings