Control of the search space size and solving linear systems in the FEAST algorithm

Lukas Krämer

Bergische Universität Wuppertal
—Applied Computer Science—

July 4th, 2014

Joint work with

Martin Galgon, Bruno Lang (Wuppertal),
Achim Basermann, Jonas Thies, (DLR Köln)
Andreas Alvermann, Ernst-Moritz-Arndt-Universität Greifswald

This work was partially supported by DFG within the priority programme 1648—Software for Exascale Computing (SPPEXA) within the project ESSEX (Equipping Sparse Solvers for Exascale).

http://www.sppexa.de
http://blogs.fau.de/essex/

Outline

Outline

Symmetric eigenvalue problem

We aim at solving the problem:

Given: $\mathsf{A} \in \mathbb{R}^{n\times n},\ \mathsf{A} = \mathsf{A}^\star$,
an interval

$$I_\lambda = [\,\underline{\lambda},\, \overline{\lambda}\,]\,.$$

Sought: $\mathsf{X} \in \mathbb{R}^{n\times m}$, $m \leq n$, $\mathsf{X}^\star\mathsf{X} = \mathsf{I}$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ such that

$$\mathsf{AX} = \mathsf{X}\Lambda \text{ and all } \lambda_j \in [\underline{\lambda},\, \overline{\lambda}]\,.$$

✎ Solve the (partial) symmetric eigenproblem. Find all eigenvalues in a given interval.

(Everything can be generalized to complex matrices and to a pair (A, B) with B hpd.)

Framework: Rayleigh-Ritz approach

- Choose a subspace $\mathcal{U} = \mathrm{span}(\mathsf{U}), \mathsf{U} \in \mathbb{C}^{n \times m}$.
- Compute the Rayleigh quotients $\mathsf{A}_\mathsf{U} := \mathsf{U}^\star \mathsf{A} \mathsf{U},\ \mathsf{B}_\mathsf{U} := \mathsf{U}^\star \mathsf{B} \mathsf{U}$.

Framework: Rayleigh-Ritz approach

- Choose a subspace $\mathcal{U} = \mathrm{span}(\mathsf{U}), \mathsf{U} \in \mathbb{C}^{n \times m}$.
- Compute the Rayleigh quotients $\mathsf{A}_\mathsf{U} := \mathsf{U}^\star \mathsf{A} \mathsf{U},\ \mathsf{B}_\mathsf{U} := \mathsf{U}^\star \mathsf{B} \mathsf{U}$.
- Compute eigenpairs $(\widetilde{\Lambda}, \widetilde{\mathsf{W}})$ of $\mathsf{A}_\mathsf{U} \mathsf{W} = \mathsf{B}_\mathsf{U} \mathsf{W} \Lambda$.
- Form Ritz pairs $(\widetilde{\Lambda}, \mathsf{U}\widetilde{\mathsf{W}})$ of $\mathsf{A}\mathsf{X} = \mathsf{B}\mathsf{X}\Lambda$.

- Choose a subspace $\mathcal{U} = \mathrm{span}(\mathsf{U}), \mathsf{U} \in \mathbb{C}^{n\times m}$.
- Compute the Rayleigh quotients $\mathsf{A}_\mathsf{U} := \mathsf{U}^\star \mathsf{A}\mathsf{U},\ \mathsf{B}_\mathsf{U} := \mathsf{U}^\star \mathsf{B}\mathsf{U}$.
- Compute eigenpairs $(\widetilde{\Lambda}, \widetilde{\mathsf{W}})$ of $\mathsf{A}_\mathsf{U}\mathsf{W} = \mathsf{B}_\mathsf{U}\mathsf{W}\Lambda$.
- Form Ritz pairs $(\widetilde{\Lambda}, \mathsf{U}\widetilde{\mathsf{W}})$ of $\mathsf{A}\mathsf{X} = \mathsf{B}\mathsf{X}\Lambda$.

...and hope that $(\widetilde{\Lambda}, \mathsf{U}\widetilde{\mathsf{W}})$ are good approximations to some eigenpairs.
Otherwise, set $\mathsf{U} := \mathsf{U}\widetilde{\mathsf{W}}$ and iterate.

The subspace in use

E. Polizzi, Phys. Rev. B, 2009, 79, 115112

In FEAST: The subspace

$$U := \frac{1}{2\pi\imath}\int_{\mathcal{C}} (zB - A)^{-1} \mathrm{d}z BY$$

is used.

- $\mathcal{C}$ = curve in complex plane surrounding I_λ.
- Y = starting base of (possibly) random vectors.

It follows

$$\frac{1}{2\pi\imath}\int_{\mathcal{C}} (zB - A)^{-1} B \mathrm{d}z = \text{orth. projector onto} \quad \mathrm{span}(X)$$

The FEAST Algorithm: Skeleton

Input: $I_\lambda := [\underline{\lambda}, \overline{\lambda}]$, an estimate $\widetilde{m}$ of the number of eigenvalues in I_λ.

Output $\hat{m} \leq \widetilde{m}$ eigenpairs with eigenvalue in I_λ.

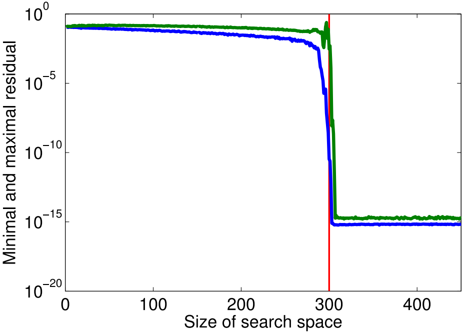
Perform:

1. Choose $\mathsf{Y} \in \mathbb{C}^{n \times \widetilde{m}}$ of full rank and compute $\mathsf{U} := \frac{1}{2\pi \mathbf{i}} \int_{\mathcal{C}} (z\mathsf{B} - \mathsf{A})^{-1} \mathsf{B} \, \mathrm{d}z \, \mathsf{Y}$,
2. Form $\mathsf{A}_\mathsf{U} := \mathsf{U}^\star \mathsf{A} \mathsf{U}, \ \mathsf{B}_\mathsf{U} := \mathsf{U}^\star \mathsf{B} \mathsf{U}$,
3. Solve size-$\widetilde{m}$ eigenproblem $\mathsf{A}_\mathsf{U} \widetilde{\mathsf{W}} = \mathsf{B}_\mathsf{U} \widetilde{\mathsf{W}} \widetilde{\Lambda}$,
4. Compute $(\widetilde{\Lambda}, \widetilde{\mathsf{X}} := \mathsf{U} \cdot \widetilde{\mathsf{W}})$,
5. If no convergence: go to Step 1 with $\mathsf{Y} := \widetilde{\mathsf{X}}$.

Problem: How to choose search space size $\widetilde{m}$?

Experiment: Choose $I_\lambda = [\underline{\lambda}, \overline{\lambda}]$ such that it contains the $m = 300$ lowest eigenvalues of size 1059-matrix. Let $\widetilde{m}$ vary, $\widetilde{m} = 1, \ldots, 450$.

Residuals vs. search space size $\widetilde{m}$:



K., Di Napoli, Galgon, Lang, Bientinesi: *Dissecting the FEAST algorithm for generalized eigenproblems*, 2013

Problem: How to choose search space size $\widetilde{m}$?

Do not choose too small!

$$\text{Determine a good choice for } \widetilde{m} \Leftrightarrow \text{Count eigenvalues in } I_\lambda$$

Leads to stopping criterion:

$$\text{“all” eigenpairs found} \Rightarrow \text{Stop Algorithm}$$

Counting eigenvalues

- $\mathrm{rank}(U) = m =$ number of eigenvalues in I_λ
- In exact arithmetic: U has singular values 0 or 1
- numerical rank?

Counting eigenvalues

- $\mathrm{rank}(U) = m =$ number of eigenvalues in I_λ
- In exact arithmetic: U has singular values 0 or 1
- numerical rank?
- It can be shown: If $m \leq \widetilde{m}$ we have $m = \#$ of singular values of U that are $\geq 1/2$ (Tang & Polizzi, 2013).
- Number $1/2$ depends on implementation (integration scheme). E.g., valid for Gauß–Legendre integration, midpoint rule of even order.

Counting eigenvalues: SVD of U

SVD $\mathsf{U} = \mathsf{W}\Sigma\mathsf{V}^\star$ with $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_m, 0, \ldots, 0)$

$\Rightarrow$ $\mathrm{span}(\mathsf{W}(:, 1:m)) = \mathrm{span}(\mathsf{U})$

$\Rightarrow$ Found the correct subspace, but cost $= \mathcal{O}(n^2 \cdot \widetilde{m})$

SVD of $\mathsf{B}_\mathsf{U} = \mathsf{U}^\star\mathsf{B}\mathsf{U}$:

$\Rightarrow$ $\mathsf{B}_\mathsf{U} = \mathsf{V}\Sigma^\star\Sigma\mathsf{V}^\star$ with $\mathsf{V} \in \mathbb{R}^{\widetilde{m}\times\widetilde{m}}$, orthogonal. No order n cost.

$\Rightarrow$ $\mathsf{U}' = \mathsf{U} \cdot \mathsf{V}(:, 1:m)$ spans the wanted space $\rightsquigarrow \mathcal{O}(n \cdot m)$.

Counting eigenvalues: SVD of U

SVD $\mathsf{U} = \mathsf{W}\Sigma\mathsf{V}^\star$ with $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_m, 0, \ldots, 0)$

$\Rightarrow$ $\mathrm{span}(\mathsf{W}(:, 1:m)) = \mathrm{span}(\mathsf{U})$

$\Rightarrow$ Found the correct subspace, but cost $= \mathcal{O}(n^2 \cdot \widetilde{m})$

SVD of $\mathsf{B}_\mathsf{U} = \mathsf{U}^\star\mathsf{B}\mathsf{U}$:

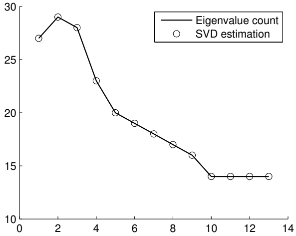
$\Rightarrow$ $\mathsf{B}_\mathsf{U} = \mathsf{V}\Sigma^\star\Sigma\mathsf{V}^\star$ with $\mathsf{V} \in \mathbb{R}^{\widetilde{m}\times\widetilde{m}}$, orthogonal. No order n cost.

$\Rightarrow$ $\mathsf{U}' = \mathsf{U} \cdot \mathsf{V}(:, 1:m)$ spans the wanted space $\rightsquigarrow \mathcal{O}(n \cdot m)$.

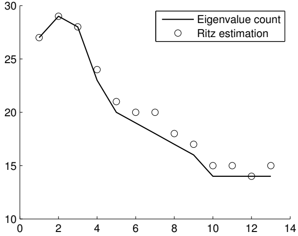
This means the dimension and a basis of the search space can be computed.

Proceed with a subspace slightly larger than m.

SVD count



Number of Ritz values in I_λ



1 Point on abscissa $\equiv$ 1 interval I_λ

$\circ =$ count from respective method

line $=$ exact count

✎ Ritz count is almost for free and not too bad. It does not deliver the subspace. Typically overestimates.

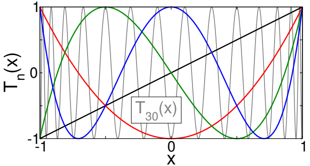
Counting eigenvalues: alternatives

Chebyshev polynomials

two-term recurrence

$$T_n(x) = \cos(n \arccos x)$$

$$T_0(x) = 1\ , \quad T_1(x) = x\ ,$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$



Expansion of functions

for example: density of states (DOS) $\hat{=}$ eigenvalue count

$$\rho(\omega) \propto \mu_0 + 2\sum_{n=1} \mu_n T_n(\omega)$$

with coefficients (moments) of the matrix A

$$\mu_n = \int_{-1}^{1} \rho(\omega)\, T_n(\omega)\, \mathrm{d}\omega$$
$$= \mathsf{x}^\star T_n(\mathsf{A})\mathsf{x}$$

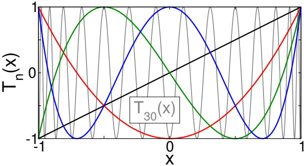
Counting eigenvalues: alternatives

Chebyshev polynomials

two-term recurrence

$$T_n(x) = \cos(n \arccos x)$$

$$T_0(x) = 1\,, \quad T_1(x) = x\,,$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$



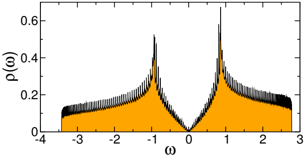
Expansion of functions

for example: density of states (DOS) $\hat{=}$ eigenvalue count

idea: use as input for FEAST

further technical details: Weiße, Wellein, Alvermann, Fehske, Rev. Mod. Phys. 78, 275 (2006).

cf. recent report (Di Napoli *et al.*, 2013)



Outline

Linear systems

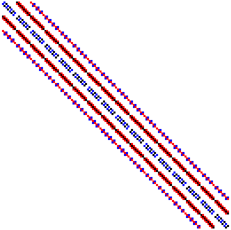
Central task: Solution of

$$(z\mathsf{B} - \mathsf{A})\mathsf{V} = \mathsf{B}\mathsf{Y}.$$

Matrix $(z\mathsf{B} - \mathsf{A})$ is

- large and sparse ~~Cholesky~~ ~~LU~~
- non-Hermitian (z complex) ~~Cholesky~~ ~~CG~~
- complex valued
- often ill-conditioned ~~GMRES~~ GMRES??

Problems from Graphene nanoribbons



some lead to banded systems ⇒ special banded direct solver.
See (Hockney & Jesshope, 1988) and (Lawrie & Sameh, 1984).

Results with banded solver

- Matrix size $n \approx 1.2M$, ratio $\frac{\text{nnz}}{n^2} \approx 0.001\%$ (12 nonzero entries per row).
- Graphene II with gap around zero.

Matrix	est.	found	iters.
Graphene I	535	535	7
Graphene II (ℓ)	492	492	4
Graphene II (r)	546	546	8
Graphene III	574	572	25

- Eigenpair residuals: $10^{-9} \dots 10^{-16}$
- Linear system residuals: $10^{-10} \dots 10^{-16}$

General (non-banded) problems

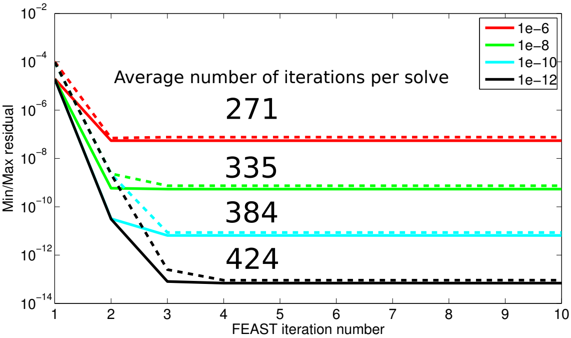
Problem large & sparse: Iterative method needed.
Candidate 1: Special tailored method for systems

$$(\mathsf{R} + \mathbf{i}\mathsf{S})(\mathsf{x} + \mathbf{i}\mathsf{y}) = \varphi + \mathbf{i}\psi$$
$$\mathsf{R} = \mathsf{R}^T, \mathsf{S} = \mathsf{S}^T, \text{ real}$$

(Axelsson & Kucherov, 2000)

- $(z\mathsf{B} - \mathsf{A})\mathsf{V} = \mathsf{BY}$ can be brought to this form.
- Leads to real systems of size $2n$.
- Not efficient enough in practice

Problem large & sparse: Iterative method needed.
Candidate 2: CARP-CG
(cf. Talk of Jonas Thies, Code also by Jonas)
Experiment: Graphene, $n = 4{,}200$, $m = 17$ interior eigenpairs
Achievable residuals $\|\mathsf{A}\mathsf{x} - \mathsf{x}\lambda\|$:



$$\max_{z_j} \kappa(z_j \mathsf{I} - \mathsf{A}) = 1097.$$

Larger example

Example 1: Graphene, $n = 84{,}000$, sought $m = 102$ eigenpairs from interior.

Average: 567 iterations/solve.

- $\mathrm{tol}_{\mathrm{lin}} = 10^{-10}$
- $\|\mathsf{A}\mathsf{x} - \mathsf{x}\lambda\| = 10^{-7} \ldots 10^{-10}$ (convergence criterion: 10^{-7}).
- Orthogonality $\max_{i \neq j} |\mathsf{x}_i^\star \mathsf{x}_j| = 10^{-15}$.

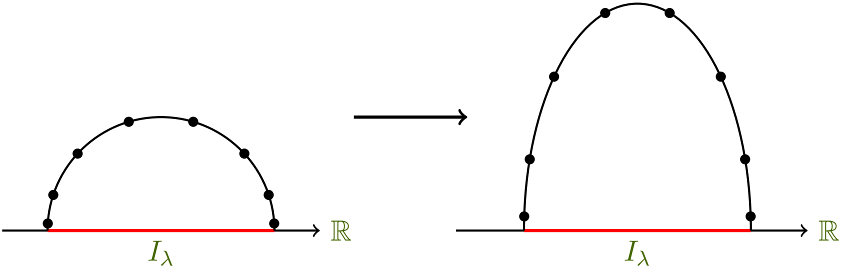
Example 2: Graphene, $n = 176{,}000$, sought $m = 201$ eigenpairs from interior.

Average: 655 iterations/solve.

- $\mathrm{tol}_{\mathrm{lin}} = 10^{-10}$
- $\|\mathsf{A}\mathsf{x} - \mathsf{x}\lambda\| = 10^{-8} \ldots 10^{-10}$ (convergence criterion: 10^{-7}).
- Orthogonality $\max_{i \neq j} |\mathsf{x}_i^\star \mathsf{x}_j| = 10^{-15}$.

Other contours

Idea: Bring integration points away from real axis.

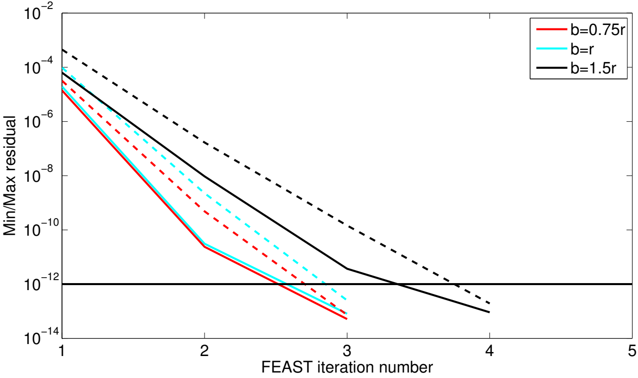


Use $\mathsf{U} = \frac{1}{2\pi\mathbf{i}}\mathsf{A}\int_{\mathcal{C}} \frac{1}{z}(z\mathsf{I} - \mathsf{A})^{-1}\mathrm{d}z$

Use of ellipse: Results

Experiment: Graphene, $n = 4{,}200$, $m = 17$ interior eigenpairs

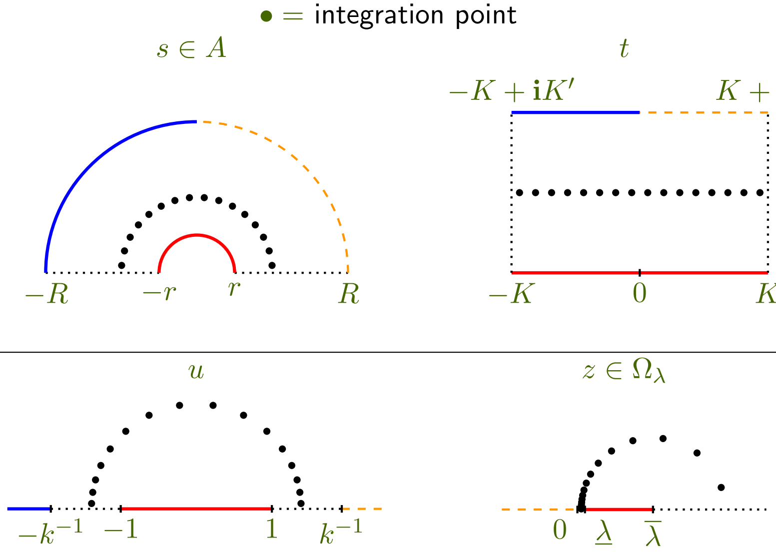
$r =$ main semi axis, second semi axis $b = \alpha r$, vary α



Required $\|Ax - x\lambda\| \leq 10^{-12}$.

$\Downarrow$ CARP-CG iterations $\leftrightarrow$ FEAST iterations $\Uparrow$

Outline



● = integration point
$s \in A$
$-R$
$-r$
r
R
t
$-K + \mathbf{i}K'$
$K + \mathbf{i}K'$
$-K$
0
K
u
$-k^{-1}$
-1
1
k^{-1}
$z \in \Omega_\lambda$
0
$\underline{\lambda}$
$\overline{\lambda}$

Outlook: Transforming the integration region

... for computing upper eigenpairs

Adapted from (Hale *et al.*, 2008): results in

$$\mathsf{U} = -\frac{\sqrt{\underline{\lambda}\overline{\lambda}}}{\pi \mathbf{i} k} \int_{-K+\mathbf{i}K'/2}^{3K+\mathbf{i}K'/2} \frac{\mathrm{cn}(t)\mathrm{dn}(t)}{(k^{-1} - \mathrm{sn}(t))^2} (z(t)\mathsf{B} - \mathsf{A})^{-1} \mathsf{B} \mathrm{d}t \mathsf{Y},$$

with $k = \frac{\sqrt{\overline{\lambda}/\underline{\lambda}}-1}{\sqrt{\overline{\lambda}/\underline{\lambda}}+1} < 1$. Then use midpoint rule.

Outlook: Transforming the integration region

... for computing upper eigenpairs

Adapted from (Hale *et al.*, 2008): results in

$$\mathsf{U} = -\frac{\sqrt{\underline{\lambda}\overline{\lambda}}}{\pi \mathbf{i} k} \int_{-K+\mathbf{i}K'/2}^{3K+\mathbf{i}K'/2} \frac{\mathrm{cn}(t)\mathrm{dn}(t)}{(k^{-1} - \mathrm{sn}(t))^2} (z(t)\mathsf{B} - \mathsf{A})^{-1} \mathsf{B} \mathrm{d}t \mathsf{Y},$$

with $k = \frac{\sqrt{\overline{\lambda}/\underline{\lambda}}-1}{\sqrt{\overline{\lambda}/\underline{\lambda}}+1} < 1$. Then use midpoint rule.

Theorem: (Hale *et al.*, 2008)

Let A, B *be real symmetric and let* $I_\lambda = [\underline{\lambda}, \overline{\lambda}] \subset (0, +\infty)$, *let* $\widetilde{\mathsf{U}}_p$ *formed by midpoint rule, then for all* $\underline{\lambda}, \overline{\lambda} > 0$

$$\|\mathsf{U} - \widetilde{\mathsf{U}}_p\| = \mathcal{O}\left(\exp\left(\frac{-\pi^2 p}{\log(\overline{\lambda}/\underline{\lambda}) + 3}\right)\right).$$

Adapted from (Hale *et al.*, 2008): results in

$$\mathsf{U} = -\frac{\sqrt{\underline{\lambda}\overline{\lambda}}}{\pi \mathbf{i} k} \int_{-K+\mathbf{i}K'/2}^{3K+\mathbf{i}K'/2} \frac{\mathrm{cn}(t)\mathrm{dn}(t)}{(k^{-1} - \mathrm{sn}(t))^2} (z(t)\mathsf{B} - \mathsf{A})^{-1} \mathsf{B} \mathrm{d}t \mathsf{Y},$$

with $k = \frac{\sqrt{\overline{\lambda}/\underline{\lambda}}-1}{\sqrt{\overline{\lambda}/\underline{\lambda}}+1} < 1$. Then use midpoint rule.

Theorem: (Hale *et al.*, 2008)

Let A, B *be real symmetric and let* $I_\lambda = [\underline{\lambda}, \overline{\lambda}] \subset (0, +\infty)$, *let* $\widetilde{\mathsf{U}}_p$ *formed by midpoint rule, then for all* $\underline{\lambda}, \overline{\lambda} > 0$

$$\|\mathsf{U} - \widetilde{\mathsf{U}}_p\| = \mathcal{O}\left(\exp\left(\frac{-\pi^2 p}{\log(\overline{\lambda}/\underline{\lambda}) + 3}\right)\right).$$

$\mathcal{O}\left(\log\left(\frac{\overline{\lambda}}{\underline{\lambda}}\right)\right)$ instead of $\mathcal{O}\left(\frac{\overline{\lambda}}{\underline{\lambda}}\right)$ solves!

Transforming the integration region: results

(A, B) from structural engineering, $n = 1473$, seek 300 highest eigenpairs.

Order (p)	Transformation method		Gauß–Legendre		Trapezoidal rule	
4	300	eigenpairs	0	eigenpairs	65	eigenpairs
	9	iterations	10	iterations	10	iterations
	13k	solves	18k	solves	17k	solves
8	300	eigenpairs	0	eigenpairs	68	eigenpairs
	5	iterations	10	iterations	10	iterations
	14k	solves	36k	solves	35k	solves
16	300	eigenpairs	294	eigenpairs	82	eigenpairs
	2	iterations	**10**	iterations	**10**	iterations
	14k	solves	58k	solves	69k	solves

!

References I

Axelsson, Owe, & Kucherov, Andrey. 2000.
Real valued iterative methods for solving complex symmetric linear systems.
Numer. Linear Algebra Appl., **7**, 197–218.

Di Napoli, E., Polizzi, E., & Saad, Y. 2013.
Efficient estimation of eigenvalue counts in an interval.
http://arxiv.org/abs/1308.4275.

Hale, Nicholas, Higham, Nicholas J., & Trefethen, Lloyd N. 2008.
Computing A^α, $\log(A)$ and Related Matrix Functions by Contour Integrals.
SIAM J. Numer. Anal., **46**(5), 2505–2523.

References II

Hockney, R.W., & Jesshope, C.R. 1988.
Parallel Computers 2.
Hilger.

Krämer, Lukas, Di Napoli, Edoardo, Galgon, Martin, Lang, Bruno, & Bientinesi, Paolo. 2013.
Dissecting the FEAST algorithm for generalized eigenproblems.
J. Comput. Appl. Math., **244**, 1–9.

Lawrie, D. H., & Sameh, A. H. 1984.
The Computation and Communication Complexity of a Parallel Banded System Solver.
ACM Trans. Math. Softw., **10**(2), 185–195.

References III

Tang, Ping Tak Peter, & Polizzi, Eric. 2013.
Subspace Iteration with Approximate Spectral Projection.
http://arxiv.org/abs/1302.0432 [math.NA], version 3.