Absolute value preconditioning for symmetric linear systems and eigenvalue problems

Eugene Vecharynski
(joint work with Andrew Knyazev)

Georgia State University
Department of Mathematics and Statistics

Scientific Computing Group Meeting
Emory University, Atlanta, GA
October 17, 2012

Two problems

1. Symmetric indefinite linear system

$$Ax = b, \qquad A = A^*$$

- Mixed FE discretizations of PDEs in fluid and solid mechanics
- Discretizations of PDEs in acoustics
- Inner steps of interior point methods in optimization
- Inner steps of symmetric eigensolvers (e.g., Jacobi-Davidson)

2. Symmetric interior eigenvalue problem

$$Kv = \lambda Mv, \qquad K = K^*, \; M = M^* > 0$$

- Vibration modes analysis
- Electronic structure calculations

$Ax = b$ and $Kv = \lambda Mv$

General setting

- Large size, ill-conditioned
- Sparse matrices or matrix-free environment
- Direct methods are inapplicable. Iterate!
- Use **preconditioners** to improve convergence

$Ax = b$ and $Kv = \lambda Mv$

General setting

- Large size, ill-conditioned
- Sparse matrices or matrix-free environment
- Direct methods are inapplicable. Iterate!
- Use **preconditioners** to improve convergence

$\Rightarrow$ This talk concerns the construction of **symmetric positive definite** (SPD) preconditioners for both problems

Preconditioning

Traditionally,

- For $Ax = b$, define $T \approx A^{-1}$
- For $Kv = \lambda Mv$, define $T \approx (K - c^2 M)^{-1}$

$\Rightarrow$ T is not SPD!

Outline

In this talk ...

- Advocate SPD preconditioning
- Motivate and introduce **absolute value** (AV) preconditioning
- Construct an AV preconditioner for a model linear system and eigenvalue problem

Symmetric indefinite linear systems

How to solve $Ax = b, \ A = A^*$?

Preconditioned system

$$TAx = Tb$$

How to solve $Ax = b, \ A = A^*$?

Preconditioned system

$$TAx = Tb$$

- T is **symmetric indefinite** or **nonsymmetric** (e.g., $T \approx A^{-1}$)
  $\Rightarrow$ TA is *nonsymmetric* in any inner product.
  The symmetry is lost.
  Methods: GMRES, BiCG, BiCGstab, QMR, etc.

How to solve $Ax = b,\ A = A^*$?

Preconditioned system

$$TAx = Tb$$

- T is **symmetric indefinite** or **nonsymmetric** (e.g., $T \approx A^{-1}$)
  $\Rightarrow$ TA is *nonsymmetric* in any inner product.
  The symmetry is lost.
  Methods: GMRES, BiCG, BiCGstab, QMR, etc.
- T is **SPD**.
  $\Rightarrow$ TA is *symmetric* in the T^{-1}-based inner product $(\cdot,\cdot)_{T^{-1}} = (\cdot, T^{-1}\cdot)$
  The symmetry is preserved.
  Method: PMINRES (**optimal, short-term recurrent**).

How to solve $Ax = b,\ A = A^*$?

Preconditioned system

$$TAx = Tb$$

- T is **symmetric indefinite** or **nonsymmetric** (e.g., $T \approx A^{-1}$)
  $\Rightarrow$ TA is *nonsymmetric* in any inner product.
  The symmetry is lost.
  Methods: GMRES, BiCG, BiCGstab, QMR, etc.
- T is **SPD**.
  $\Rightarrow$ TA is *symmetric* in the T^{-1}-based inner product $(\cdot,\cdot)_{T^{-1}} = (\cdot, T^{-1}\cdot)$
  The symmetry is preserved.
  Method: PMINRES (**optimal, short-term recurrent**).

$\Rightarrow$ How to correctly define an SPD preconditioner T?

Matrix absolute value

The eigenvalue decomposition

$$A = V \Lambda V^*$$

- The matrix absolute value of A is defined as

$$|A| = V\,|\Lambda|\,V^*, \; |\Lambda| = \mathrm{diag}\{\left|\lambda_j\right|\}.$$

- The matrix sign of A is defined as

$$\mathrm{sign}(A) = V \mathrm{sign}(\Lambda)\, V^*, \; \mathrm{sign}(\Lambda) = \mathrm{diag}\left\{\mathrm{sign}(\lambda_j)\right\}.$$

The **polar decomposition**

$$A = |A|\,\mathrm{sign}(A) = \mathrm{sign}(A)\,|A|\,.$$

An ideal SPD preconditioner

Let $T = |A|^{-1}$. Then $TAx = Tb$ is

$$\operatorname{sign}(A)x = |A|^{-1} b.$$

The matrix $TA = \operatorname{sign}(A)$ has two distinct eigenvalues: -1 and 1.

$\Rightarrow$ PMINRES converges in at most two steps

An ideal SPD preconditioner

Let $T = |A|^{-1}$. Then $TAx = Tb$ is

$$\text{sign}(A)x = |A|^{-1}\, b.$$

The matrix $TA = \text{sign}(A)$ has two distinct eigenvalues: -1 and 1.

$\Rightarrow$ PMINRES converges in at most two steps

- $T = |A|^{-1}$ is an **ideal SPD preconditioner**
- Construction of the *exact* $|A|^{-1}$ is prohibitively expensive
- **AV preconditioning:** Construct an SPD preconditioner T as an approximation to $|A|^{-1}$, i.e.,

$$T \approx |A|^{-1}\,.$$

Bounds on eigenvalues of TA

Theorem

Given a symmetric indefinite A, an SPD T, and constants $\delta_1 \geq \delta_0 > 0$, such that

$$\delta_0(v, T^{-1}v) \leq (v, |A|\, v) \leq \delta_1(v, T^{-1}v), \ \forall v.$$

Then

$$\Lambda(TA) \subset [-\delta_1, -\delta_0] \bigcup [\delta_0, \delta_1] \,.$$

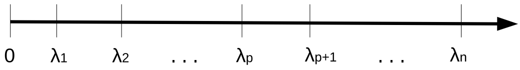
Challenges

For **practical applications**:

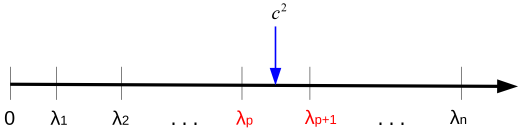
- Can we construct $T \approx |A|^{-1}$ using A (matvecs with A)?
- Can we construct $T \approx |A|^{-1}$ at an optimal cost $\mathcal{O}(\text{nnz}(A))$?

Eigenvalue problems

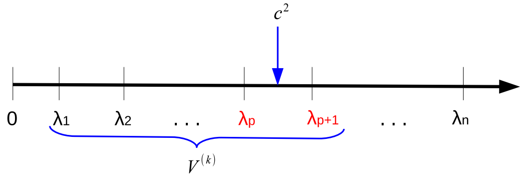
$Kv = \lambda Mv$



$Kv = \lambda Mv$



$Kv = \lambda Mv$



LOBPCG:

$$
\begin{aligned}
\Lambda^{(k)} &= (V^{(k)})^* K V^{(k)}, \; (V^{(k)})^* M V^{(k)} = I \\
V^{(k+1)} &= V^{(k)} C_V^{(k)} + T \left(K V^{(k)} - M V^{(k)} \Lambda^{(k)} \right) C_R^{(k)} + P^{(k)} C_P^{(k)} \\
P^{(k+1)} &= V^{(k+1)} - V^{(k)} C_V^{(k)}
\end{aligned}
$$

$Kv = \lambda Mv$

LOBPCG:

$$\begin{aligned}
\Lambda^{(k)} &= (V^{(k)})^* K V^{(k)},\ (V^{(k)})^* M V^{(k)} = I \\
V^{(k+1)} &= V^{(k)} C_V^{(k)} + T \left(K V^{(k)} - M V^{(k)} \Lambda^{(k)} \right) C_R^{(k)} + P^{(k)} C_P^{(k)} \\
P^{(k+1)} &= V^{(k+1)} - V^{(k)} C_V^{(k)}
\end{aligned}$$

Choice of the preconditioner

- $T \approx K^{-1} \Rightarrow$ emphasizes the extreme (smallest) eigenpairs
- $T \approx \left(K - c^2 M\right)^{-1} \Rightarrow$ indefinite preconditioning
- Our approach: $T \approx \left|K - c^2 M\right|^{-1}$

The same challenges!

For **practical applications**:

- Can we construct $T \approx |K - c^2 M|^{-1}$ using K and M?
- Can we construct $T \approx |K - c^2 M|^{-1}$ at an optimal cost $\mathcal{O}(\text{nnz}(K)+\text{nnz}(M))$?

Model problems

- **The "shifted Laplacian" equation**

$$-\Delta u(x,y) - c^2 u(x,y) = f(x,y), \quad (x,y) \in \Omega = (0,1) \times (0,1)$$
$$u|_\Gamma = 0.$$

- **The eigenvalue problem** (find eigenpairs near c^2)

$$-\Delta u(x,y) = \lambda u(x,y), \quad (x,y) \in \Omega = (0,1) \times (0,1)$$
$$u|_\Gamma = 0.$$

Model problems

After discretization (FD):

- **Symmetric indefinite linear system**

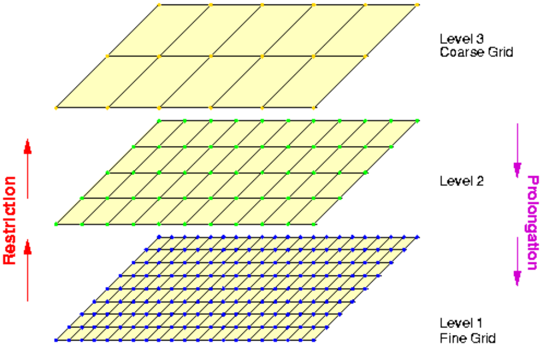
$$(L - c^2 I)x = b \;\Rightarrow\; T \approx |L - c^2 I|^{-1}$$

- **Symmetric eigenvalue problem** (find eigenpairs near c^2).

$$Lv = \lambda v \;\Rightarrow\; T \approx |L - c^2 I|^{-1}$$

$\Rightarrow$ Apply T by approximately solving $|L - c^2 I|w = r$. **Use MG!**

MG


Level 3
Coarse Grid
Level 2
Level 1
Fine Grid
Restriction
Prolongation

Two-grid scheme for $|L - c^2 I| w = r$

Algorithm (Two-grid scheme for $|L - c^2 I| w = r$)

Input r. Output $w = Tr$.

1. ***Presmoothing***. *Apply ν smoothing steps,*

$$w^{(i+1)} = w^{(i)} + \tau(r - |L - c^2 I| w^{(i)}), \; i = 0, \dots, \nu - 1, \; w^{(0)} = 0.$$

   Set $w^{pre} = w^{(\nu)}$, $\nu \geq 1$.

2. ***Coarse grid correction***. *Restrict residual to the coarse grid (R), apply $\left|L_H - c^2 I_H\right|^{-1}$, prolongate to the fine grid (P), and add to w^{pre},*

$$w^{cgc} = w^{pre} + P |L_H - c^2 I_H|^{-1} R \left(r - |L - c^2 I| w^{pre} \right).$$

3. ***Postsmoothing***. *Apply ν smoothing steps,*

$$w^{(i+1)} = w^{(i)} + \tau(r - |L - c^2 I| w^{(i)}), \; i = 0, \dots, \nu - 1, \; w^{(0)} = w^{cgc}.$$

   Return $w = w^{(\nu)}$.

Two-grid scheme for $|L - c^2 I| w = r$

Algorithm (Two-grid scheme for $|L - c^2 I| w = r$)

Input r. Output $w = Tr$.

1. ***Presmoothing**. Apply ν smoothing steps,*

$$w^{(i+1)} = w^{(i)} + \tau(r - |L - c^2 I| w^{(i)}), \; i = 0, \dots, \nu - 1, \; w^{(0)} = 0.$$

   Set $w^{pre} = w^{(\nu)}$, $\nu \geq 1$.

2. ***Coarse grid correction**. Restrict residual to the coarse grid (R), apply $\left|L_H - c^2 I_H\right|^{-1}$, prolongate to the fine grid (P), and add to w^{pre},*

$$w^{cgc} = w^{pre} + P |L_H - c^2 I_H|^{-1} R \left(r - |L - c^2 I| w^{pre} \right).$$

3. ***Postsmoothing**. Apply ν smoothing steps,*

$$w^{(i+1)} = w^{(i)} + \tau(r - |L - c^2 I| w^{(i)}), \; i = 0, \dots, \nu - 1, \; w^{(0)} = w^{cgc}.$$

   Return $w = w^{(\nu)}$.

Idea: Replace $|L - c^2 I| \approx B$, where $y = Bu$ is easy to compute

Two-grid AV preconditioner

Algorithm (Two-grid AV preconditioner)

Input r, $B \approx |L - c^2 I|$. Output $w = Tr$.

1. ***Presmoothing****. Apply ν smoothing steps,*
   $$w^{(i+1)} = w^{(i)} + \tau(r - Bw^{(i)}),\ i = 0, \ldots, \nu - 1,\ w^{(0)} = 0.$$
   Set $w^{pre} = w^{(\nu)}$, $\nu \geq 1$.
2. ***Coarse grid correction****. Restrict residual to the coarse grid (R), apply $\left|L_H - c^2 I_H\right|^{-1}$, prolongate to the fine grid (P), and add to w^{pre},*
   $$w^{cgc} = w^{pre} + P|L_H - c^2 I_H|^{-1} R\left(r - Bw^{pre}\right).$$
3. ***Postsmoothing****. Apply ν smoothing steps,*
   $$w^{(i+1)} = w^{(i)} + \tau(r - Bw^{(i)}),\ i = 0, \ldots, \nu - 1,\ w^{(0)} = w^{cgc}.$$
   Return $w = w^{(\nu)}$.

MG AV preconditioner (V-cycle)

Algorithm (AV-MG(r_l): the MG AV preconditioner)

Input r_l, output w_l.

1. *Set $B_l \approx |L_l - c^2 I_l|$.*
2. ***Presmoothing**. Apply ν smoothing steps,*

   $$w_l^{(i+1)} = w_l^{(i)} + \tau_l (r_l - B_l w_l^{(i)}), \; i = 0, \ldots, \nu - 1, \; w_l^{(0)} = 0.$$

   Set $w_l^{pre} = w_l^{(\nu)}$, $\nu \geq 1$.
3. ***Coarse grid correction**. Restrict residual to coarser grid (R_{l-1}), apply AV-MG on this level, prolongate to the finer grid (P_l), add to w_l^{pre},*

   $$\begin{aligned} w_l^{cgc} &= w_l^{pre} + P_l \, AV\text{-}MG(R_{l-1} \, (r_l - B_l w_l^{pre})), \; \text{if } l > 1; \\ w_1^{cgc} &= w_1^{pre} + P_1 |L_0 - c^2 I_0|^{-1} R_0 \left(r_1 - B_1 w_1^{pre}\right), \; \text{if } l = 1. \end{aligned}$$
4. ***Postsmoothing**. Apply ν smoothing steps,*

   $$w_l^{(i+1)} = w_l^{(i)} + \tau_l (r_l - B_l w_l^{(i)}), \; i = 0, \ldots, \nu - 1, \; w_l^{(0)} = w_l^{cgc}.$$

   Return $w_l = w_l^{(\nu)}$.

How to choose $B_l \approx |L_l - c^2 I_l|$?

Observations:

- On sufficiently fine grids ($ch_l << 1$), $|L_l - c^2 I_l| \approx L_l$.
- Since $|L_l - c^2 I_l| = \left(2h(L_l - c^2 I_l) - I\right)\left(L_l - c^2 I_l\right)$,

$$|L_l - c^2 I_l| \approx \underbrace{\left(2q_{m-1}(L_l - c^2 I_l) - I\right)\left(L_l - c^2 I_l\right)}_{p_m(L_l - c^2 I_l)},$$

  where q_{m-1} is a least-squares polynomial approximation of the Heaviside step function h (**polynomial filtering**).
- On coarser grids, evaluation of $y = p_m(L_l - c^2 I_l)u$ is inexpensive.

How to choose $B_l \approx |L_l - c^2 I_l|$?

Observations:

- On sufficiently fine grids ($ch_l << 1$), $|L_l - c^2 I_l| \approx L_l$.
- Since $|L_l - c^2 I_l| = \left(2h(L_l - c^2 I_l) - I\right)\left(L_l - c^2 I_l\right)$,

$$|L_l - c^2 I_l| \approx \underbrace{\left(2q_{m-1}(L_l - c^2 I_l) - I\right)\left(L_l - c^2 I_l\right)}_{p_m(L_l - c^2 I_l)},$$

  where q_{m-1} is a least-squares polynomial approximation of the Heaviside step function h (**polynomial filtering**).
- On coarser grids, evaluation of $y = p_m(L_l - c^2 I_l)u$ is inexpensive.

$\Rightarrow$ Given a "switching parameter" $\delta \in (0, 1)$, define

$$B_l = \begin{cases} L_l, & ch_l < \delta; \\ p_m(L_l - c^2 I_l), & \text{otherwise.} \end{cases}$$

Summary of theoretical analysis

- The preconditioner is SPD under mild assumptions on the smoother, restriction and prolongation.
- The size of the coarsest grid h_0 should be $\sim 1/c$.
- Two-grid preconditioner reduces the error for $|L - c^2 I| w = r$ in the directions of *almost* all Laplacian eigencomponents.

For more details, see

Eugene Vecharynski and Andrew Knyazev:
Absolute value preconditioning for symmetric indefinite linear systems.
Available at http://arxiv.org/abs/1104.4530, submitted

Numerical experiments

In our tests

- Standard coarsening
- Full weighting for restriction and piecewise multilinear interpolation for prolongation
- Choose $\delta = 1/3$ and $m = 10$, so that

  $$B_l = \begin{cases} L_l, & ch_l < 1/3; \\ p_{10}(L_l - c^2 I_l), & \text{otherwise.} \end{cases}$$

  Computations of p_{10} are based on filtering with Chebyshev polynomials.
- Smoother for $B_l = L_l$: 1 Richardson's iteration, $\tau_l = h_l^2/5$
- Smoother for $B_l = p_m(L_l - c^2 I_l)$: 5 Richardson's iterations, $\tau_l = h_l^2/(5 - c^2 h_l^2)$

Numerical experiments

AV preconditioning for a model linear systems

Spectra of preconditioned matrices

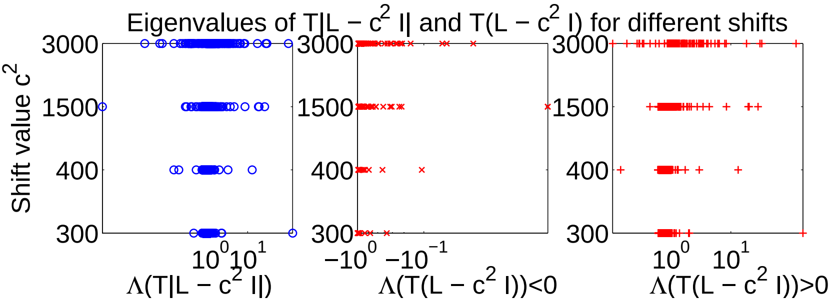


Figure : Spectrum of $T\left|L - c^2 I\right|$ (left), negative eigenvalues of $T\left(L - c^2 I\right)$ (center), and positive eigenvalues of $T\left(L - c^2 I\right)$ (right); $n = 16129$

Available preconditioning for $(L - c^2 I)x = b$

Compare AV-MG with

- SPD preconditioning: $T = L^{-1}$ (Bayliss, Goldstein, Turkel 83)
- Indefinite preconditioning: $T =$ MG V-cycle for $(L - c^2 I)x = b$

Many related works: e.g., Elman, Ernst, O'Leary 01; Bramble, Leyk, Pasciak 93; van Gijzen, Erlangga, Vuik 07, etc.

Comparisons with available preconditioned schemes

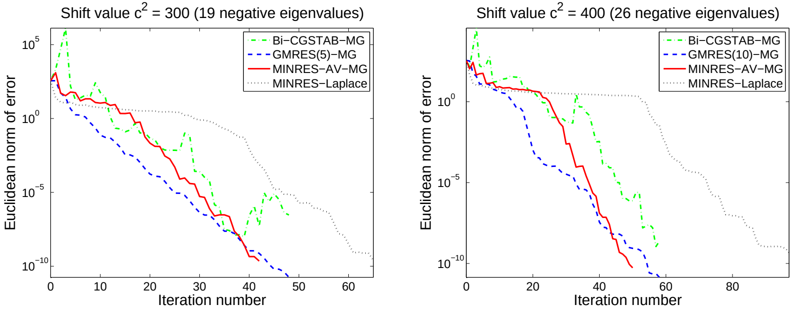


Figure : Coarsest problems of size 225; $n = 65025$

Comparisons with available preconditioned schemes

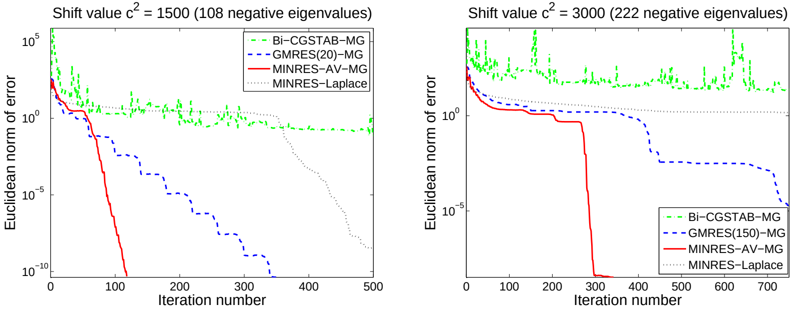


Figure : Coarsest problems of size 961; $n = 65025$

Mesh-independent convergence

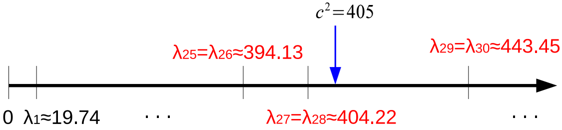
Number of steps performed to decrease error norm by 10^{-8}.

	$h = 2^{-7}$	$h = 2^{-8}$	$h = 2^{-9}$	$h = 2^{-10}$	$h = 2^{-11}$
$c^2 = 300$	31	30	30	30	30
$c^2 = 400$	38	37	37	37	37
$c^2 = 1500$	97	89	88	89	90
$c^2 = 3000$	222	279	256	257	256

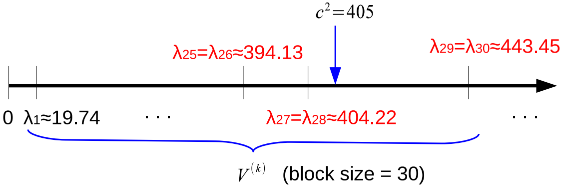
Table : Mesh-independent convergence of MINRES-AV-MG.

AV preconditioning for a model eigenvalue problem

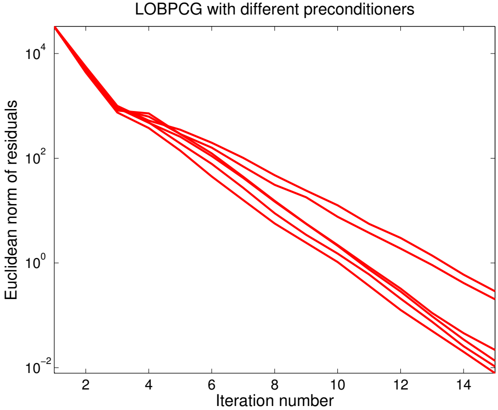
The eigenvalue problem $Lv = \lambda v$



The eigenvalue problem $Lv = \lambda v$

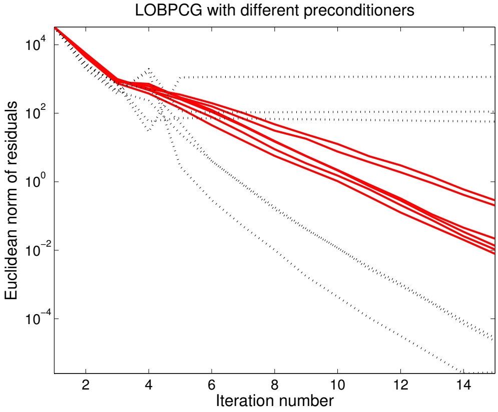


Convergence to eigenpairs (λ_{25}, v_{25}) through (λ_{30}, v_{30})



— "Standard" SPD preconditioning: $T \approx L^{-1}$ (MG V-cycle for L)

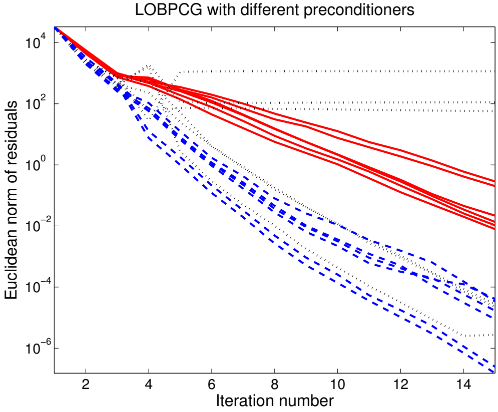
Convergence to eigenpairs (λ_{25}, v_{25}) through (λ_{30}, v_{30})



— "Standard" SPD preconditioning: $T \approx L^{-1}$ (MG V-cycle for L)

— Indefinite preconditioning: $T \approx (L - c^2 I)^{-1}$ (MG V-cycle for $L - c^2 I$)

Convergence to eigenpairs (λ_{25}, v_{25}) through (λ_{30}, v_{30})



— "Standard" SPD preconditioning: $T \approx L^{-1}$ (MG V-cycle for L)
— Indefinite preconditioning: $T \approx (L - c^2 I)^{-1}$ (MG V-cycle for $L - c^2 I$)
— AV preconditioning: $T \approx |L - c^2 I|^{-1}$ (AV-MG)

Current and future work

- Domain decomposition AV preconditioning
- Algebraic MG type AV preconditioning
- AV preconditioning for saddle point problems
- Construction of novel AV preconditioned iterative schemes for interior eigenvalue problems
- Application to electronic structure calculations
- HPC software development

Thank you!