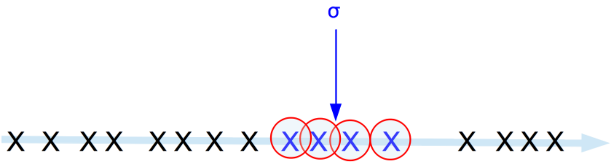
Symmetric indefinite systems, positive definite preconditioning, and interior eigenvalues

Eugene Vecharynski
Lawrence Berkeley National Laboratory
Computational Research Division

joint work with Andrew Knyazev (MERL)

14th Copper Mountain Conference on Iterative Methods,
March 25, 2016

The problem



Find several eigenvalues and (their eigenvectors) of a large, possibly sparse, Hermitian matrix A closest to the shift σ.

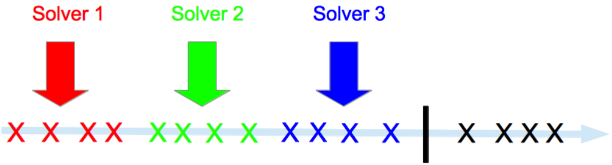
Motivation: HUGE **extreme** eigenvalue problems

Find a number of extreme (lowest) eigenpairs (λ, x) of

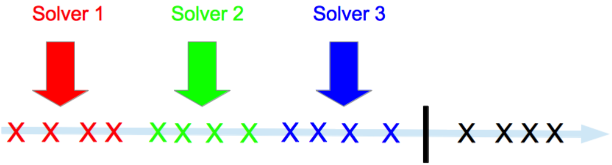
$$Av = \lambda v, \qquad A = A^*$$

- Matrix A is sparse or available implicitly
- The problem size n is well beyond 10^6
- The number of targeted eigenpairs $\sim 10^3$–10^5, or even more

Motivation: spectrum slicing eigensolvers



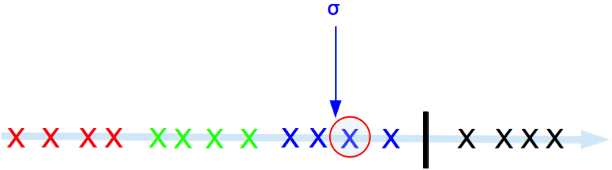
Motivation: spectrum slicing eigensolvers



Possible choices of interior eigensolvers

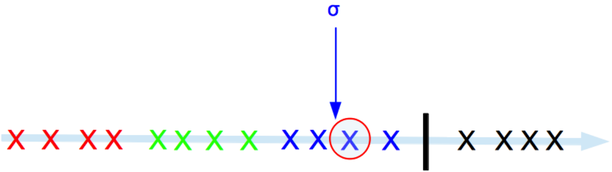
- Shift-Invert+Lanczos [Aktulga, et al PARCO'14]
- Filtering+Lanczos [Li, Xi, EV, Yang, Saad, submit. SISC'15]
- Block preconditioned interior eigensolvers

Preconditioned interior eigensolvers



Find an eigenvalue λ of A closest to the given shift σ and its associated eigenvector v.

Relation to linear systems



Assume that λ is known and we only need eigenvector v. Solve

$$(A - \lambda I)v = 0$$

Preconditioned linear solvers for $Cv = f$

C is Hermitian indefinite, T is HPD preconditioner

- Preconditioned MINRES (PMINRES)

$$\|r^{(i)}\|_T \leq 2 \left(\frac{\sqrt{|ad|} - \sqrt{|bc|}}{\sqrt{|ad|} + \sqrt{|bc|}} \right)^{[\frac{i}{2}]} \|r^{(0)}\|_T, \; \Lambda(TC) \subset [a, b] \cup [c, d]$$

"Globally" optimal Krylov subspace method, short-term recurrence

- Preconditioned steepest descent-like method (PSDI) ???...

The PSDI iteration for $Cv = f$

C is Hermitian indefinite, T is HPD preconditioner

- Restarting PMINRES at every step, **does not converge**

$$v^{(i+1)} \leftarrow v^{(i)} + b^{(i)} Tr^{(i)}, \quad r^{(i+1)} \perp_T CTr^{(i)}, \quad r^{(i)} = f - Cv^{(i)}$$

The PSDI iteration for $Cv = f$

C is Hermitian indefinite, T is HPD preconditioner

- Restarting PMINRES at every step, **does not converge**

$$v^{(i+1)} \leftarrow v^{(i)} + b^{(i)} Tr^{(i)}, \quad r^{(i+1)} \perp_T CTr^{(i)}, \quad r^{(i)} = f - Cv^{(i)}$$

- Restarting PMINRES at every other step

$$v^{(i+1)} \leftarrow v^{(i)} + b^{(i)} Tr^{(i)} + c^{(i)} TCTr^{(i)}, \quad r^{(i+1)} \perp_T C\mathcal{K},$$

where $\mathcal{K} = \operatorname{span}\left\{ Tr^{(i)}, TCTr^{(i)} \right\}$

Linear convergence: $\|r^{(i+1)}\|_T \leq \left(\dfrac{|ad| - |bc|}{|ad| + |bc|} \right) \|r^{(i)}\|_T$

[EV, Knyazev, submitted '15]

Preconditioned linear solvers for $Cv = f$

C is Hermitian indefinite, T is HPD preconditioner

- Preconditioned MINRES (PMINRES)

$$\|r^{(i)}\|_T \le 2\left(\frac{\sqrt{|ad|}-\sqrt{|bc|}}{\sqrt{|ad|}+\sqrt{|bc|}}\right)^{[\frac{i}{2}]}\|r^{(0)}\|_T,\ \Lambda(TC)\subset[a,b]\cup[c,d]$$

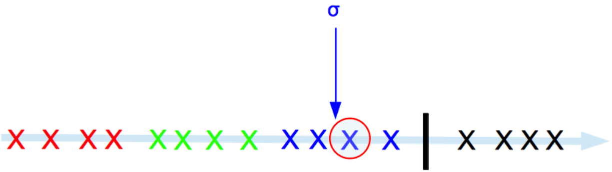
"Globally" optimal Krylov subspace method, short-term recurrence

- Preconditioned steepest descent-like method (PSDI)

$$\|r^{(i)}\|_T \le \left(\frac{|ad|-|bc|}{|ad|+|bc|}\right)\|r^{(i)}\|_T$$

"Locally" optimal, mathematically equivalent to PMINRES(2)

Back to interior eigenproblems ...



Assume that λ is known and we only need eigenvector v. Solve

$$\underbrace{(A - \lambda I)}_{C} v = \underbrace{0}_{f}$$

An ideal preconditioned eigensolver

- The PSDI iteration for $(A-\lambda I)v = 0$

$$v^{(i+1)} \leftarrow v^{(i)} + b^{(i)} Tr^{(i)} + c^{(i)} T\,(A-\lambda I)\,Tr^{(i)}, \quad r^{(i)} = (A-\lambda I)v^{(i)}$$

- Iteration coefficients chosen from

$$\underbrace{(A-\lambda I)v^{(i+1)}}_{r^{(i+1)}} \perp_T (A-\lambda I)\mathcal{K},$$

where $\mathcal{K} = \operatorname{span}\left\{ Tr^{(i)}, T\,(A-\lambda I)\,Tr^{(i)} \right\}$.

- T is HPD, $\quad (\cdot,\cdot)_T = (\cdot, T\cdot)$

- Converges linearly for singular consistent systems, cv rate determined by nonzero spectrum of $T(A-\lambda I)$ [EV, thesis'11]

From linear to eigenvalue solvers

Linear solver for $(A - \lambda I)v = 0$

$$v^{(i+1)} \leftarrow v^{(i)} + b^{(i)}Tr^{(i)} + c^{(i)}T(A - \lambda I)\,Tr^{(i)}, \quad r^{(i)} = (A - \lambda I)v^{(i)}$$

$$(A - \lambda I)v^{(i+1)} \perp_{T} (A - \lambda I)\mathcal{K}, \quad \mathcal{K} = \text{span}\left\{Tr^{(i)}, T(A - \lambda I)\,Tr^{(i)}\right\}$$

From linear to eigenvalue solvers

Linear solver for $(A-\lambda I)v=0$

$$v^{(i+1)} \leftarrow v^{(i)} + b^{(i)}\,Tr^{(i)} + c^{(i)}\,T\,(A-\lambda I)\,Tr^{(i)}, \quad r^{(i)} = (A-\lambda I)v^{(i)}$$

$$(A-\lambda I)v^{(i+1)} \perp_{T} (A-\lambda I)\mathcal{K}, \quad \mathcal{K} = \text{span}\left\{ Tr^{(i)}, T\,(A-\lambda I)\,Tr^{(i)} \right\}$$

Eigenvalue solver for $Av=\lambda v$

$$v^{(i+1)} \leftarrow \alpha^{(i)} v^{(i)} + \beta^{(i)}\,Tr^{(i)} + \gamma^{(i)}\,T(A-\lambda^{(i)} I)\,Tr^{(i)}, \quad r^{(i)} = Av^{(i)} - \lambda^{(i)} v^{(i)}$$

From linear to eigenvalue solvers

Linear solver for $(A-\lambda I)v = 0$

$$v^{(i+1)} \leftarrow v^{(i)} + b^{(i)}\,Tr^{(i)} + c^{(i)}\,T\,(A-\lambda I)\,Tr^{(i)}, \quad r^{(i)} = (A-\lambda I)v^{(i)}$$

$$(A-\lambda I)v^{(i+1)} \perp_T (A-\lambda I)\mathcal{K}, \quad \mathcal{K} = \text{span}\left\{Tr^{(i)}, T\,(A-\lambda I)\,Tr^{(i)}\right\}$$

Eigenvalue solver for $Av = \lambda v$

$$v^{(i+1)} \leftarrow \alpha^{(i)} v^{(i)} + \beta^{(i)}\,Tr^{(i)} + \gamma^{(i)}\,T(A-\lambda^{(i)} I)\,Tr^{(i)}, \quad r^{(i)} = Av^{(i)} - \lambda^{(i)} v^{(i)}$$

$$Av^{(i+1)} - \theta v^{(i+1)} \perp_T (A-\sigma I)\mathcal{Z}, \quad \mathcal{Z} = \text{span}\left\{v^{(i)}, Tr^{(i)}, T(A-\lambda^{(i)} I)\,Tr^{(i)}\right\}$$

The T-harmonic Rayleigh–Ritz procedure

Find an approximate eigenpair $(\theta, v^{(i+1)})$, such that

$$Av^{(i+1)} - \theta v^{(i+1)} \perp_T (A - \sigma I)\mathcal{Z}, \quad v^{(i+1)} \in \mathcal{Z}$$

$$\Updownarrow$$

$$Z^*(A - \sigma I)T(A - \sigma I)Zy = \xi Z^*(A - \sigma I)TZy$$

- Matrix $Z = [v^{(i)},\ Tr^{(i)},\ T(A - \lambda^{(i)}I)Tr^{(i)}]$ and $y = (\alpha^{(i)}, \beta^{(i)}, \gamma^{(i)})^T$
- The T-harmonic Ritz pairs $\theta = \xi + \sigma$ and $v^{(i+1)} = Zy$
- Reduces to (standard) harmonic RR if $T = I$

Harmonic vs T-harmonic

The harmonic RR

$$A\tilde{v} - \theta\tilde{v} \perp (A - \sigma I)\mathcal{Z}, \quad \tilde{v} \in \mathcal{Z}$$

- A priori error bound [EV, LAA'16]

$$\sin\angle(v, \tilde{v}) \leq \kappa(A - \sigma I)\sqrt{1 + \frac{\gamma^2}{\delta^2}}\sin\angle(v, \mathcal{Z})$$

Harmonic vs T-harmonic

The harmonic RR

$$A\tilde{v} - \theta\tilde{v} \perp (A - \sigma I)\mathcal{Z}, \quad \tilde{v} \in \mathcal{Z}$$

- A priori error bound [EV, LAA'16]

$$\sin\angle(v, \tilde{v}) \leq \kappa(A - \sigma I)\sqrt{1 + \frac{\gamma^2}{\delta^2}} \sin\angle(v, \mathcal{Z})$$

The T-harmonic RR

$$A\tilde{v} - \theta\tilde{v} \perp_T (A - \sigma I)\mathcal{Z}, \quad \tilde{v} \in \mathcal{Z}$$

- A priori error bound under "idealized" condition $TA = AT$

$$\sin\angle(v, \tilde{v}) \leq \kappa(T^{1/2}(A - \sigma I))\sqrt{1 + \frac{\gamma^2}{\delta^2}} \sin\angle(v, \mathcal{Z})$$

The PLHR algorithm

The Preconditioned Locally Harmonic Residual method [EV, Knyazev,'15]

Given an initial guess $v^{(0)}$ and an SPD preconditioner T, compute an eigenpair of A associated with λ closest to the shift σ

- $v \leftarrow v^{(0)}$; $v \leftarrow v/\|v\|$; $\lambda \leftarrow (v, Av)$; $p \leftarrow [\ ]$;
- **While** convergence not reached
    - Compute $w \leftarrow T(Av - \lambda v)$
    - Compute $s \leftarrow T(Aw - \lambda w)$
    - Set $Z \leftarrow [v,\ w,\ s,\ p]$
    - Find the eigenpair (ξ, y) associated with the smallest $|\xi|$ of $Z^*(A - \sigma I)T(A - \sigma I)Zy = \xi Z^*(A - \sigma I)TZy$
    - $v \leftarrow \alpha v + \beta w + \gamma s + \delta p$
    - $p \leftarrow \beta w + \gamma s + \delta p$
    - $v \leftarrow v/\|v\|$; $\lambda \leftarrow (v, Av)$
- **EndWhile**

The PLHR algorithm

The Preconditioned Locally Harmonic Residual method [EV, Knyazev,'15]

Given an initial guess $v^{(0)}$ and an SPD preconditioner T, compute an eigenpair of A associated with λ closest to the shift σ

- $v \leftarrow v^{(0)}$; $v \leftarrow v/\|v\|$; $\lambda \leftarrow (v, Av)$; $p \leftarrow [\ ]$;
- **While** convergence not reached
    - Compute $w \leftarrow T(Av - \lambda v)$
    - Compute $s \leftarrow T(Aw - \lambda w)$
    - Set $Z \leftarrow [v,\ w,\ s,\ p]$
    - Find the eigenpair (ξ, y) associated with the smallest $|\xi|$ of $Z^*(A - \sigma I)T(A - \sigma I)Zy = \xi Z^*(A - \sigma I)TZy$
    - $v \leftarrow \alpha v + \beta w + \gamma s + \delta p$
    - $p \leftarrow \beta w + \gamma s + \delta p$
    - $v \leftarrow v/\|v\|$; $\lambda \leftarrow (v, Av)$
- **EndWhile**

The PLHR algorithm

The Preconditioned Locally Harmonic Residual method [EV, Knyazev,'15]

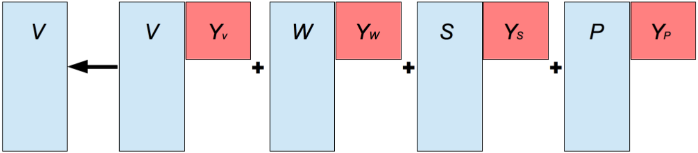
Given an initial guess $v^{(0)}$ and an SPD preconditioner T, compute an eigenpair of A associated with λ closest to the shift σ

- $v \leftarrow v^{(0)}$; $v \leftarrow v/\|v\|$; $\lambda \leftarrow (v, Av)$; $p \leftarrow [\ ]$;
- **While** convergence not reached
  - Compute $w \leftarrow T(Av - \lambda v)$
  - Compute $s \leftarrow T(Aw - \lambda w)$
  - Set $Z \leftarrow [v,\ w,\ s,\ p]$
  - Find the eigenpair (ξ, y) associated with the smallest $|\xi|$ of $Z^*(A - \sigma I)T(A - \sigma I)Zy = \xi Z^*(A - \sigma I)TZy$
  - $v \leftarrow \alpha v + \beta w + \gamma s + \delta p$
  - $p \leftarrow \beta w + \gamma s + \delta p$
  - $v \leftarrow v/\|v\|$; $\lambda \leftarrow (v, Av)$
- **EndWhile**

The BPLHR algorithm for interior eigenpairs

The Block Preconditioned Locally Harmonic Residual method

[EV, Knyazev, SISC'15]



- $W \leftarrow T(AV - V\Lambda)$, $S \leftarrow T(AW - W\Lambda)$, $P \in \mathrm{col}\{V, V_{prev}\}$
- $Z \leftarrow [V,\ W,\ S,\ P]$
- Find eigenvectors Y associated with the k smallest magnitude eigenvalues of

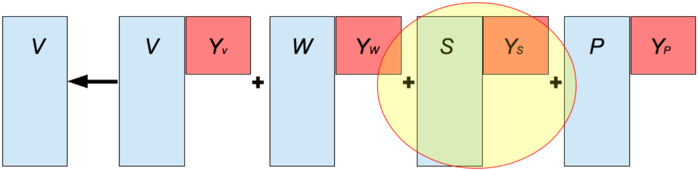
$$(Z^*(A - \sigma I)T(A - \sigma I)Z)Y = (Z^*(A - \sigma I)TZ)Y\Theta$$

- $V \leftarrow VY_V + WY_W + SY_S + PY_P$, $\Lambda \leftarrow \mathrm{diag}\{V^*AV\}$

The BPLHR algorithm for interior eigenpairs

The Block Preconditioned Locally Harmonic Residual method

[EV, Knyazev, SISC'15]



- $W \leftarrow T(AV - V\Lambda)$, $S \leftarrow T(AW - W\Lambda)$, $P \in \text{col}\{V, V_{prev}\}$
- $Z \leftarrow [V,\ W,\ S,\ P]$
- Find eigenvectors Y associated with the k smallest magnitude eigenvalues of

$$(Z^*(A - \sigma I)T(A - \sigma I)Z)Y = (Z^*(A - \sigma I)TZ)Y\Theta$$

- $V \leftarrow VY_V + WY_W + SY_S + PY_P$, $\Lambda \leftarrow \text{diag}\{V^*AV\}$

Diagonal preconditioning in planewave DFT calculations

- In the DFT based electronic structure calculations the Hamiltonian is of the form

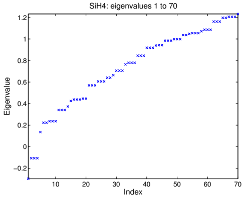
$$H = L + V$$

- The kinetic energy L is diagonal and SPD in planewave basis

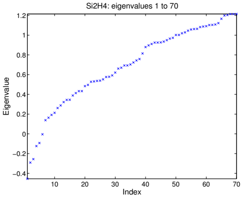
- Use $T \approx L^{-1}$ as a preconditioner [Teter et al, Phys Rev B '89]

- Standard preconditioning option in planewave codes [ABINIT, VASP, Quantum Espresso, QBox, etc.]

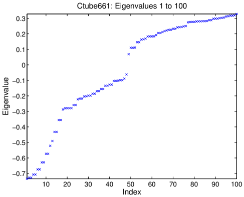
Several examples



(a) SiH4 ($\sigma = 0.8$)



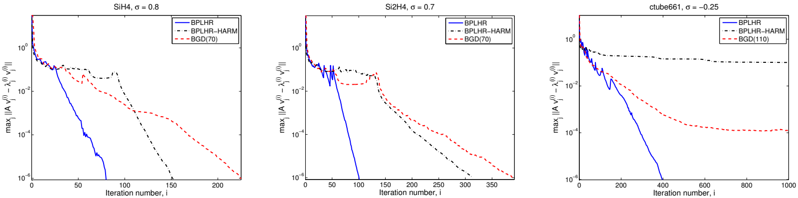
(b) Si2H4 ($\sigma = 0.7$)



(c) Ctube ($\sigma = -0.25$)

Figure: Parts of spectra of the Hamiltonian matrices.

Several examples



(a) SiH4 ($\sigma = 0.8$) (b) Si2H4 ($\sigma = 0.7$) (c) Ctube ($\sigma = -0.25$)

Figure: Computing 10 eigenpairs closest to σ.

A general definition of the PLHR preconditioner

Observation: The (restarted) PMINRES for $(A - \lambda I)v = 0$ with $T = |A - \lambda I|^{\dagger}$ converges to exact solution in at most two steps

A general definition of the PLHR preconditioner

Observation: The (restarted) PMINRES for $(A - \lambda I)v = 0$ with $T = |A - \lambda I|^{\dagger}$ converges to exact solution in at most two steps

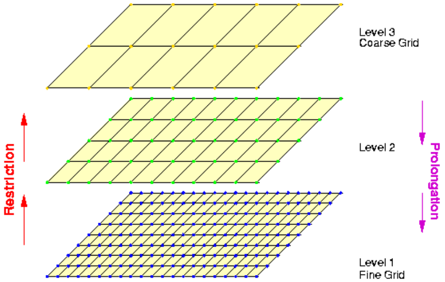
- Choose PLHR preconditioner as $T \approx |A - \sigma I|^{-1}$
- Absolute Value (AV) preconditioning [EV, Knyazev, SISC'13]
- Combines "shift-and-invert" and SPD'ness
- Can use $T \approx |C|^{-1}$ as a PMINRES preconditioner for $Cx = b$

Multigrid AV preconditioners for shifted Laplacian

Consider construction of

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

where A is a 2D Laplacian on a unit square.



Construct AV only on the coarsest grid! [EV, Knyazev, SISC'13]

PMINRES+AV for shifted Laplacian

Solve $(A - \sigma I)x = b$, MG preconditioner $T \approx |A - \sigma I|^{-1}$

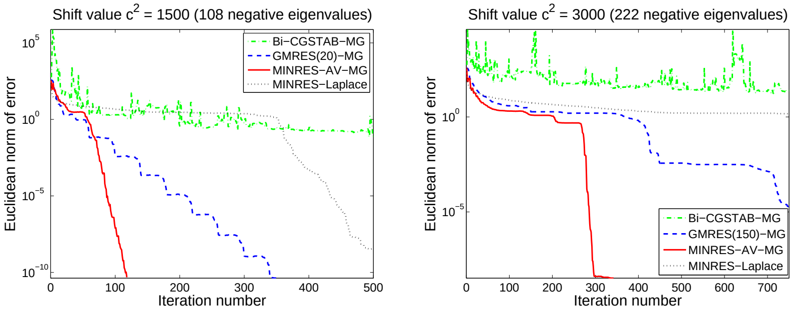


Figure: Coarsest problems of size 961; $\sigma = c^2$; $n = 65025$

BPLHR+AV for shifted Laplacian

Solve $Av = \lambda v$, MG preconditioner $T \approx |A - \sigma I|^{-1}$

			Shifts (σ)						
Method	Prec.	RR	400	450	500	550	600	650	700
BPLHR	AV	T-harm.	57	81	68	133	117	190	278
BPLHR	AV	harm.	563	-	493	635	-	-	-
BPLHR	Indef.	harm.	30	40	45	-	59	338	424
Davidson	Indef.	harm.	36	46	57	-	376	763	-

Table: Iteration numbers to converge to 10 eigenpairs of the Laplacian closest to shifts σ; $n = 16,129$.

BPLHR+AV for shifted Laplacian

Solve $Av = \lambda v$, MG preconditioner $T \approx |A - \sigma I|^{-1}$

			Shifts (σ)					
Method	Prec.	RR	800	900	1000	1100	1200	1300
BPLHR	AV	T-harm.	270	168	177	344	365	363
BPLHR	AV	harm.	590	417	377	625	437	217
BPLHR	Indef.	harm.	-	-	-	-	-	-
Davidson	Indef.	harm.	230	818	837	-	-	-

Table: Iteration numbers to converge to 20 eigenpairs of the Laplacian closest to shifts σ; $n = 16,129$.

Current and future work

- Development of AV preconditioners (DD, AMG, etc.)
- Generalization of PLHR on non-Hermitian matrices, indefinite preconditioning [EV, Yang, Xue, SISC'16]
- Tools for slicing the spectrum [EV, Yang, submitted '16], post-processing (missing eigs, orthogonality, etc.)
- A lot of CS: load balance, scheduling, parallel code, etc.
- Large-scale applications

Current and future work

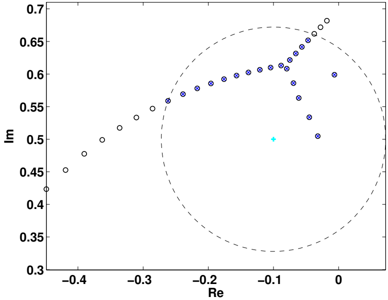
- Development of AV preconditioners (DD, AMG, etc.)
- Generalization of PLHR on non-Hermitian matrices, indefinite preconditioning [EV, Yang, Xue, SISC'16]
- Tools for slicing the spectrum [EV, Yang, submitted '16], post-processing (missing eigs, orthogonality, etc.)
- A lot of CS: load balance, scheduling, parallel code, etc.
- Large-scale applications

Thank you!

The Generalized PLHR (GPLHR) algorithm

Compute a subset of k right eigenpairs (λ, x) of a non-Hermitian matrix pair (A, B) that are closest to a given shift $\sigma \in \mathbb{C}$

$$Ax = \lambda Bx, \quad A, B \in \mathbb{C}^{n \times n}$$



A, B can can be symmetric

T can be indefinite

Larger trial subspace

[EV, Yang, Xue, to appear in SISC'16]

Preconditioned linear solvers for $Cv = b$

C is Hermitian positive definite (HPD), T is HPD preconditioner

- Preconditioned Conjugate Gradient (PCG)

$$\|e^{(i)}\|_C \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^i \|e^{(0)}\|_C, \quad \kappa = \lambda_{\max}(TC)/\lambda_{\min}(TC)$$

"Globally" optimal Krylov subspace method, short-term recurrence

- Preconditioned steepest descent (PSD)

$$v^{(i+1)} \leftarrow v^{(i)} + \alpha^{(i)} T(b - Cv^{(i)}), \quad \|e^{(i+1)}\|_C \le \left(\frac{\kappa-1}{\kappa+1}\right)\|e^{(i)}\|_C$$

"Locally" optimal, mathematically equivalent to PCG(1)

The Block PLHR (BPLHR) algorithm

Given an initial guess $V^{(0)}$ and an SPD preconditioner T, compute eigenpairs of A associated with k eigenvalues closest to the shift σ

- $V \leftarrow V^{(0)}$; normalize V; $\lambda_j \leftarrow (v_j, Av_j)$; $\Lambda \leftarrow \text{diag}\{\lambda_j\}$, $P \leftarrow [\ ]$;
- **While** convergence not reached
  - Compute $W \leftarrow T(AV - V\Lambda)$
  - Compute $S \leftarrow T(AW - W\Lambda)$
  - Set $Z \leftarrow [V,\ W,\ S,\ P]$
  - Find k eigenpairs (ξ, y) associated with the smallest $|\xi|$ of $Z^*(A - \sigma I)T(A - \sigma I)Zy = \xi Z^*(A - \sigma I)TZy$
  - $V \leftarrow VY_V + WY_W + SY_S + PY_P$
  - $P \leftarrow WY_W + SY_S + PY_P$
  - Normalize columns of V; $\lambda_j \leftarrow (v_j, Av_j)$; $\Lambda \leftarrow \text{diag}\{\lambda_j\}$
- **EndWhile**