Preconditioned Locally Minimal Residual Method for Computing Interior Eigenpairs of Symmetric Operators

Eugene Vecharynski[1] Andrew Knyazev[2]

[1]Department of Computer Science and Engineering
University of Minnesota

[2]Department of Mathematical and Statistical Sciences
University of Colorado Denver

June 15, 2011
Householder Symposium XVIII
Tahoe City, CA

The symmetric eigenvalue problem

Find a subset of eigenpairs (λ, v)

$$Av = \lambda Bv, \quad A = A^* \in \mathbb{R}^{n\times n}, \quad B = B^* \in \mathbb{R}^{n\times n}, \quad B > 0.$$

Assume that A and B are large, i.e., factorizations are infeasible.
"Modest" accuracy approximate solutions suffice.

For example,

- Compute several **extreme** eigenpairs (vibration modes analysis, buckling, electronic structure calculations, data clustering, etc).
  Methods: JD, Lanczos, LOBPCG, etc.
- Compute several **interior** eigenpairs (electronic structure calculations)
  Approaches: block iterations+deflation, shift-and-invert (SI), folded spectrum (FS).

The eigenvalue problem of interest

Given a symmetric eigenvalue problem

$$Av = \lambda Bv, \quad A = A^*, \quad B = B^* > 0.$$

We target the eigenpair (λ_q, v_q), which corresponds to the **smallest magnitude eigenvalue**; $|\lambda_q| = \min_j \{|\lambda_j|\}$.

- **For example:** find an eigenpair of the negative discrete Laplace operator L, which corresponds to the eigenvalue, closest to a given shift value c^2,

$$\left(L - c^2 I\right) v = \lambda v,$$

  i.e., $A \equiv L - c^2 I$, $B \equiv I$.

- Similar problems occur, e.g., in the study of large nanosystems ("band gap" calculations)

This talk

Two goals:

1. Introduce **preconditioning** for the eigenvalue problem
2. Construct a **preconditioned eigensolver**
   (No prior FS/SI transformation, matrix-free, low-dimensional trial subspaces, optimal or close to optimal convergence rate)

Outline:

1. From an eigenvalue problem to a linear system
   - Simple locally optimal preconditioned linear solvers
   - Absolute value preconditioning
2. From a linear solver to an eigenvalue solver
   - Choice of trial subspaces
   - Approximate eigenvector extraction procedure
   - Preconditioned Locally Minimal Residual method (PLMR)
3. Numerical examples for the model problem with the Laplacian
   - Absolute value preconditioning for existing schemes (LOBPCG)
   - Convergence behavior of PLMR

From an eigenvalue problem to a linear system

The symmetric eigenvalue problem

$$Av = \lambda Bv, \quad A = A^*, \quad B = B^* > 0.$$

- Assume that we know the exact eigenvalue $\lambda = \lambda_q$
- For simplicity, let λ_q be of multiplicity 1

The related homogeneous linear system

$$(A - \lambda_q B)x = 0, \quad A = A^*, \quad B = B^* > 0.$$

- Symmetric and (in general) indefinite
- Singular
- The solution is the eigenvector $x = v_q$ corresponding to λ_q

What are the simplest proved to be convergent preconditioned iterative schemes for approximating a nonzero null space vector?

How to define preconditioning for the linear system?

Solving $(A - \lambda_q B)\, v = 0$

Let T be a **symmetric positive definite (SPD) preconditioner**. The preconditioned system

$$T\,(A - \lambda_q B)\, x = 0$$

is symmetric in the T^{-1}-inner product; $(x, y)_{T^{-1}} = (x, T^{-1}y)$.

Consider a **preconditioned locally optimal scheme**

$$x^{(i+1)} \in x^{(i)} + u, \quad u \in \mathcal{U}^{(i)} = \text{span}\left\{ Tr^{(i)}, T\,(A - \lambda_q B)\, Tr^{(i)} \right\},$$

such that the residual vector $r^{(i)} = Ax^{(i)} - \lambda_q B x^{(i)}$ satisfies

$$\|r^{(i+1)}\|_T = \min_{u \in \mathcal{U}^{(i)}} \|r^{(i)} + (A - \lambda_q B)u\|_T.$$

(This can be viewed as the preconditioned minimal residual method for $(A - \lambda_q B)x = 0$ *restarted after every 2 steps)*

The accelerated scheme for $(A - \lambda_q B)x = 0$

The simple preconditioned locally optimal scheme

$$x^{(i+1)} = x^{(i)} + \alpha^{(i)} w^{(i)} + \beta^{(i)} T \left(A w^{(i)} - \lambda_q B w^{(i)} \right),$$

where $w^{(i)} = T \left(A x^{(i)} - \lambda_q B x^{(i)} \right)$ and T is SPD.

We can further **accelerate the convergence by introducing an additional vector** $x^{(i)} - x^{(i-1)}$:

$$x^{(i+1)} = x^{(i)} + \alpha^{(i)} w^{(i)} + \beta^{(i)} T \left(A w^{(i)} - \lambda_q B w^{(i)} \right) + \gamma^{(i)} (x^{(i)} - x^{(i-1)})$$

- T-norm of the new residual is minimized over a larger subspace
- Can expect an accelerated convergence
- **Use this scheme as a "base" for deriving a preconditioned eigensolver**

How to define preconditioning?

Preconditioning for $(A - \lambda_q B)\, x = 0$

How to define an **SPD preconditioner** for

$$(A - \lambda_q B)\, x = 0?$$

- If $T = |A - \lambda_q B|^{\dagger}$ then the preconditioned locally optimal schemes **converges to the exact solution in exactly** 1 **step**.
  $\Rightarrow$ Construct SPD preconditioners $T \approx |A - \lambda_q B|^{\dagger}$
- The eigenvalue λ_q is the smallest in the absolute value. Assume that its magnitude is relatively negligible. Drop the term with $\lambda_q B$.
  $\Rightarrow$ Construct SPD preconditioners $T \approx |A|^{-1}$ (or $T \approx |A + \alpha I|^{-1}$ for a singular A).

Choose an SPD preconditioner T for system $(A - \lambda_q B)x = 0$ to "resemble" $|A|^{-1}$.

Use the same preconditioner for $Av = \lambda Bv$!

Absolute value preconditioning

Originally introduced for solving symmetric indefinite linear systems $Ax = b$.

Eugene Vecharynski and Andrew Knyazev:

Absolute value preconditioning for symmetric indefinite linear systems.

UCD CCM Report 298, submitted to SISC, 2011.

An SPD preconditioner T is an **absolute value preconditioner (AVP)**, if

$$\delta_0(v, T^{-1}v) \leq (v, |A|\, v) \leq \delta_1(v, T^{-1}v), \forall v \in \mathbb{R}^n,$$

with constants $\delta_1 \geq \delta_0 > 0$, such that $\delta_1/\delta_0 \geq 1$ is reasonably small and independent of problem size (for mesh problems).

We use AVP for $(A - \lambda_q B)x = 0$ and $Av = \lambda Bv$.

From a linear solver to an eigensolver

The choice of trial subspaces

The "base" linear solver for $(A - \lambda_q B)x = 0$

- Iteration of the form

$$x^{(i+1)} = x^{(i)} + \alpha^{(i)} w^{(i)} + \beta^{(i)} T\left(Aw^{(i)} - \lambda_q B w^{(i)}\right) + \gamma^{(i)}(x^{(i)} - x^{(i-1)})$$

- The new iterate $x^{(i+1)}$ is determined as an element of

$$\text{span}\left\{x^{(i)}, w^{(i)}, T(Aw^{(i)} - \lambda_q B w^{(i)}), x^{(i-1)}\right\}, \; w^{(i)} = T(Ax^{(i)} - \lambda_q B x^{(i)})$$

The eigensolver for $Av = \lambda Bv$

- Replace λ_q by the Rayleigh Quotient (RQ) $\lambda^{(i)}$
- **Extract eigenvector approximation $v^{(i+1)}$ from**

$$\text{span}\left\{v^{(i)}, w^{(i)}, T(Aw^{(i)} - \lambda^{(i)} B w^{(i)}), v^{(i-1)}\right\}, \; w^{(i)} = T(Av^{(i)} - \lambda^{(i)} B v^{(i)}),$$

where $\lambda^{(i)} = \frac{(v^{(i)}, Av^{(i)})}{(v^{(i)}, Bv^{(i)})}$, $v^{(-1)} = 0$.

From a linear solver to an eigensolver

The approximate eigenvector extraction

The "base" linear solver for $(A - \lambda_q B)x = 0$

Iteration parameters are chosen such that

$$x^{(i+1)} = \underset{u}{\operatorname{argmin}} \|Au - \lambda_q Bu\|_T,$$

$$u \in x^{(i)} + \operatorname{span}\left\{w^{(i)}, T\left(Aw^{(i)} - \lambda_q Bw^{(i)}\right), x^{(i)} - x^{(i-1)}\right\}.$$

The eigensolver for $Av = \lambda Bv$

- Extract an approximate eigenvector $v^{(i+1)}$ such that

$$v^{(i+1)} = \underset{u, \|u\|_B = 1}{\operatorname{argmin}} \|Au - \tilde{\lambda} Bu\|_T, \ \tilde{\lambda} \approx \lambda_q$$

  for $u \in \operatorname{span}\left\{v^{(i)}, w^{(i)}, T(Aw^{(i)} - \lambda^{(i)} Bw^{(i)}), v^{(i-1)}\right\}$, where $w^{(i)} = T\left(Ax^{(i)} - \lambda^{(i)} Bx^{(i)}\right)$

- Such an approach (w.r.t. 2-norm) is known as the **refined** extraction procedure (Jia'97)
- Results in the solution of a 4-by-4 symmetric eigenvalue problem

The PLMR method

The Preconditioned Locally Minimal Residual method (PLMR) for approximating the eigenpair (λ_q, v_q), where $|\lambda_q| = \min_j |\lambda_j|$:

$$\begin{aligned} \lambda^{(i)} &= \frac{(v^{(i)}, Av^{(i)})}{(v^{(i)}, Bv^{(i)})}, \\ w^{(i)} &= T\left(Av^{(i)} - \lambda^{(i)} Bv^{(i)}\right), \ p^{(i)} = v^{(i)} - \alpha^{(i-1)} v^{(i-1)}, \ p^{(0)} = 0, \\ v^{(i+1)} &= \alpha^{(i)} v^{(i)} + \beta^{(i)} w^{(i)} + \gamma^{(i)} T\left(Aw^{(i)} - \lambda^{(i)} Bw^{(i)}\right) + \delta^{(i)} p^{(i)}. \end{aligned}$$

- The preconditioner T is SPD (we suggest AVP)
- The new approximate eigenvector $v^{(i+1)}$ is extracted from

$$\mathcal{V} = \text{span}\left\{v^{(i)}, w^{(i)}, T(Aw^{(i)} - \lambda^{(i)} Bw^{(i)}), v^{(i-1)}\right\}$$

- The parameters $\alpha^{(i)}$, $\beta^{(i)}$, $\gamma^{(i)}$, and $\delta^{(i)}$ are such that

$$v^{(i+1)} = \underset{z \in \mathcal{V}, \|z\|_B = 1}{\text{argmin}} \ \|Az - \tilde{\lambda} Bz\|_T, \ \tilde{\lambda} \approx \lambda_q$$

- $\tilde{\lambda} \leftarrow \text{RR}(AB^{-1}Av = \lambda Bv, \mathcal{V})$, or $\tilde{\lambda} = \lambda^{(i)}$ near solution

A model problem

Find an eigenpair of the standard 2D negative discrete (FD) Laplace operator L, which corresponds to the eigenvalue closest to c^2.

1. If c^2 is reasonably small, then use LOBPCG to compute a sufficiently large number of extreme eigenpairs of

$$Lv = \lambda v$$

2. For larger c^2, use PLMR to compute the eigenpair corresponding to the smallest magnitude eigenvalue of

$$\left(L - c^2 I\right) v = \lambda v$$

- For both approaches choose T to resemble $\left|L - c^2 I\right|^{-1}$ (AVP), i.e.,

$$\delta_0(v, T^{-1}v) \leq (v, \left|L - c^2 I\right| v) \leq \delta_1(v, T^{-1}v), \ \forall v \in \mathbb{R}^n$$

- **An example of such a (MG) preconditioner has been constructed for a linear system** $\left(L - c^2 I\right) x = b$ with c^2 not too large (EV, Knyazev '11) Reuse this preconditioner for our model eigenvalue problem!

Absolute value preconditioner for the model problem

Algorithm (Two-grid absolute value preconditioner)

Input r, output $w = Tr$.

1. **Pre-smoothing**. *Apply ν pre-smoothing steps,*

$$w^{(i+1)} = w^{(i)} + M^{-1}(r - Lw^{(i)}),\ i = 0, \ldots, \nu - 1,\ w^{(0)} = 0.$$

This step results in the pre-smoothed vector $w^{pre} = w^{(\nu)}$, $\nu \geq 1$.

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

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

3. **Post-smoothing**. *Apply ν post-smoothing steps,*

$$w^{(i+1)} = w^{(i)} + M^{-*}(r - Lw^{(i)}),\ i = 0, \ldots, \nu - 1,\ w^{(0)} = w^{cgc}.$$

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

Absolute value preconditioner for the model problem

- The resulting preconditioner T is SPD under mild assumptions on the smoother, restriction and prolongation
- **In practice, we use the multilevel extension of the two-grid algorithm ("V-cycle")**

In our numerical tests

- The problem size $n = (2^7 - 1)^2 \approx 1.6 \times 10^4$ ($h = 2^{-7}$).
- Standard coarsening scheme. The coarsest grid has the mesh size 2^{-4}.
- Full weighting for the restriction and piecewise multilinear interpolation for the prolongation.
- The smoother: ω-damped Jacobi, $\omega = 4/5$.

Numerical examples. LOBPCG for $Lv = \lambda v$

Computing (λ_q, v_q) closest to 100. "Standard" MG preconditioner

- Eigenvalues of L: $\lambda_1 \approx 19.7382$, $\lambda_2 = \lambda_3 \approx 49.3396$, $\lambda_4 \approx 78.9410$, $\lambda_5 = \lambda_6 \approx 98.6554$, $\lambda_7 \approx 128.2568$ ... Target (λ_5, v_5)
- Use LOBPCG with block size 5 + "standard" MG preconditioner ($T \approx L^{-1}$)

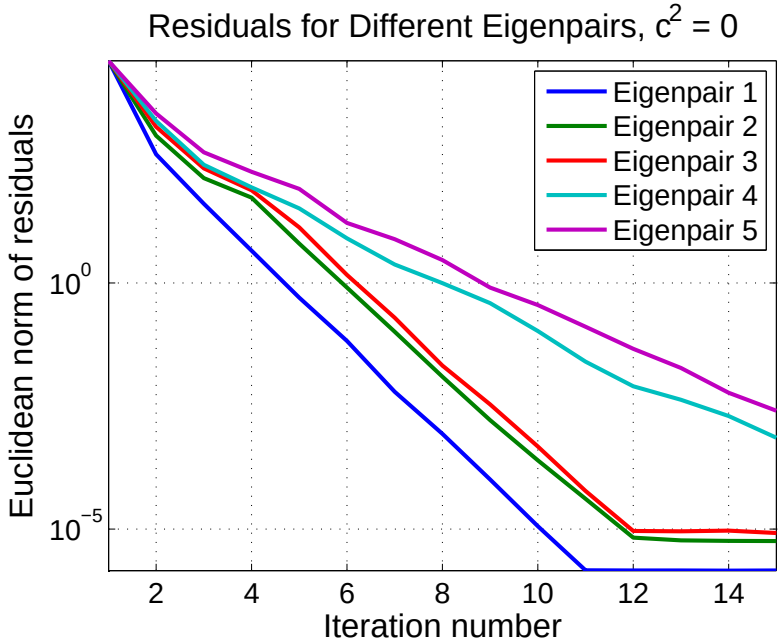


Figure: The typical convergence behavior within a block. **The convergence to (λ_5, v_5) is the slowest.**

Numerical examples. LOBPCG for $Lv = \lambda v$

Computing (λ_q, v_q) closest to 100. AVP, $c^2 = 50$

- Eigenvalues of L: $\lambda_1 \approx 19.7382$, $\lambda_2 = \lambda_3 \approx 49.3396$, $\lambda_4 \approx 78.9410$, $\lambda_5 = \lambda_6 \approx 98.6554$, $\lambda_7 \approx 128.2568$ ...
- Use LOBPCG with block size 5 + AVP ($T \approx |L - c^2 I|^{-1}$, $c^2 = 50$)

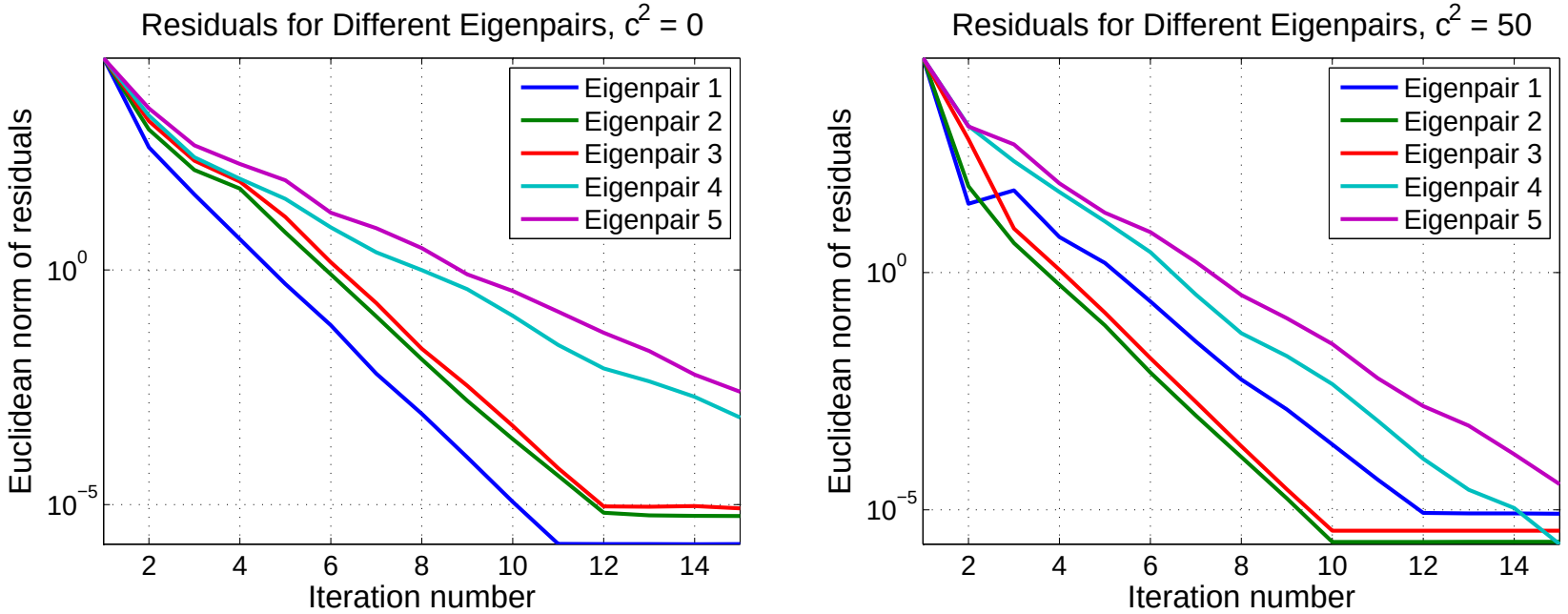


Figure: The convergence to (λ_q, v_q) closest to 50 is the fastest.

Further increase c^2 in the AVP...

Numerical examples. LOBPCG for $Lv = \lambda v$

Computing (λ_q, v_q) closest to 100. AVP, $c^2 = 100$

- Eigenvalues of L: $\lambda_1 \approx 19.7382$, $\lambda_2 = \lambda_3 \approx 49.3396$, $\lambda_4 \approx 78.9410$, $\lambda_5 = \lambda_6 \approx 98.6554$, $\lambda_7 \approx 128.2568$ ...
- Use LOBPCG with block size 5 + AVP ($T \approx |L - c^2 I|^{-1}$, $c^2 = 100$)

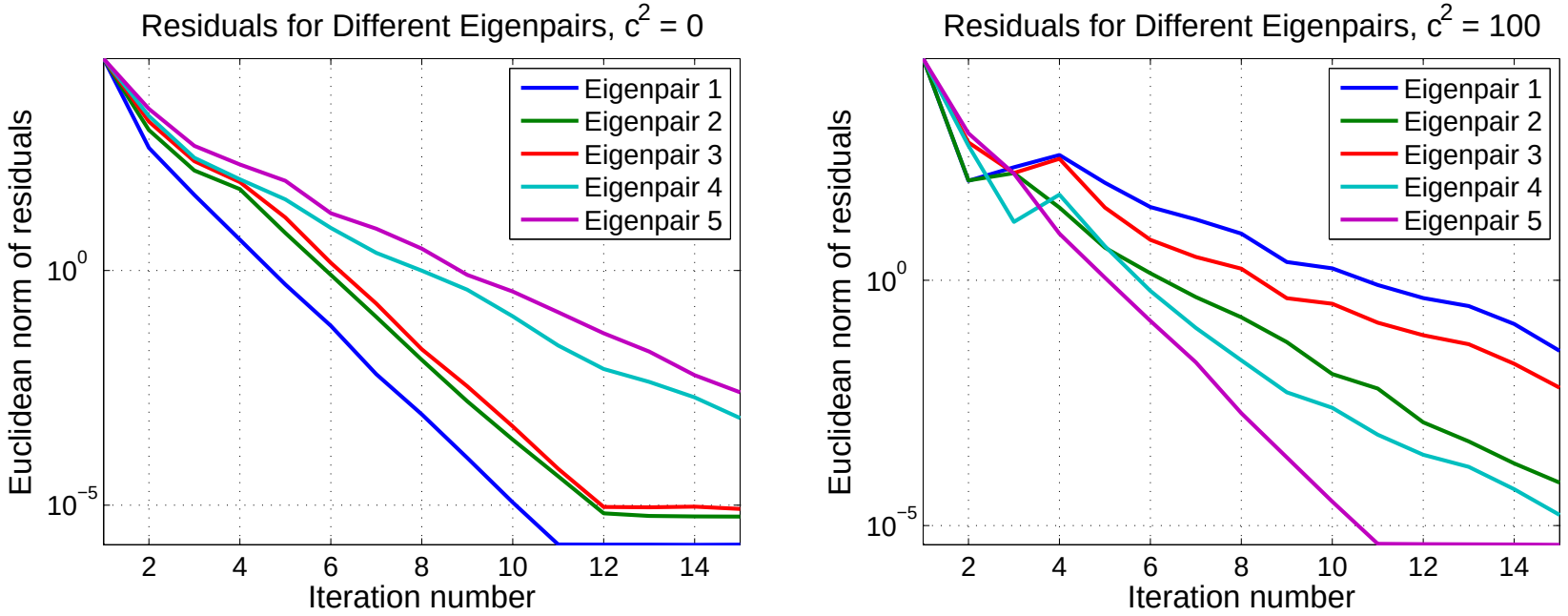


Figure: The convergence to (λ_5, v_5), closest to 100, is the fastest.

AVP accelerates convergence of LOBPCG to the desired eigenpairs.

Numerical examples. PLMR for $(L - c^2 I)v = \lambda v$

Computing (λ_q, v_q) closest to $c^2 = 200, 250$

For larger c^2 the AVP-LOBPCG becomes inefficient $\Rightarrow$ use PLMR for $(L - c^2 I)v = \lambda v$

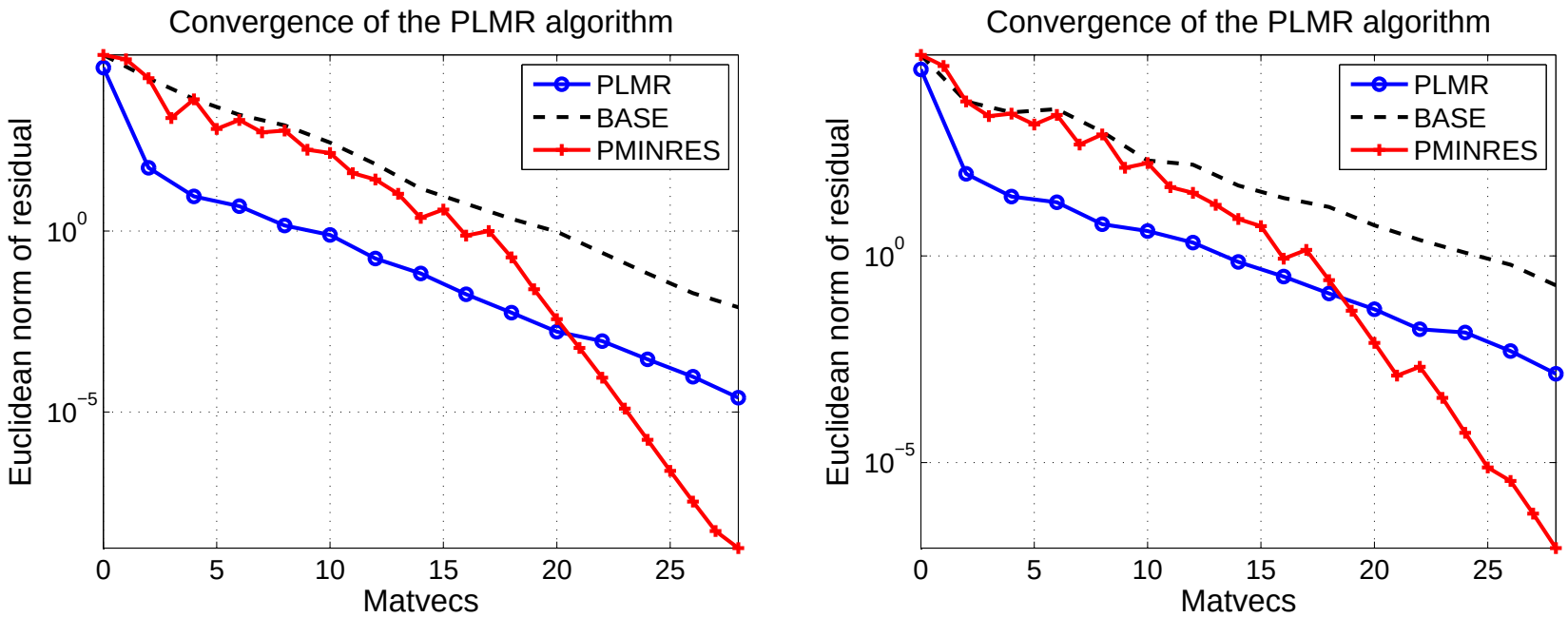
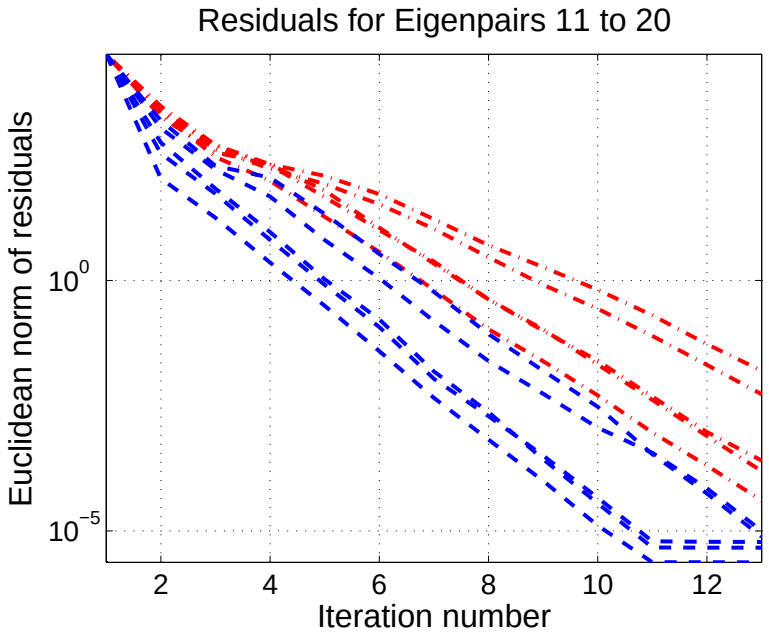


Figure: Comparison of PLMR with linear solvers for $((L - c^2 I) - \lambda_q I)x = 0$. The targeted eigenpairs correspond to λ_{13} (left, $c^2 = 200$) and λ_{15} (right, $c^2 = 250$).

Numerical examples. Further applications of AVP

Computing a large number of extreme eigenpairs of $Lv = \lambda v$

- Target 20 smallest eigenpairs of L. Set block size to 10.
- Use LOBPCG + "standard" MG preconditioner ($T \approx L^{-1}$) to compute the first 10 smallest eigenpairs. Put them into the constraints (deflation). Repeat the run with AVP $T \approx |L - c^2 I|^{-1}$ ($c^2 = \mathrm{RQ}(v_{10})$).
- Compare with the "standard" LOBPCG run ($T \approx L^{-1}$) after deflation



AVP can be used to accelerate the convergence of extreme eigenpairs.

Conclusions, current and future work

Conclusions

- Introduced a new approach for computing the eigenpair corresponding to the smallest magnitude eigenvalue
- Extended the concept of the absolute value preconditioning to the eigenvalue problem
- Good results for the model problem with the Laplacian for moderately small shift values

Current and future work

- Extension to the subspace iteration
- Practical applications
- Theoretical study
- Software

Thank you!