CCLRC
Rutherford Appleton Laboratory

# A Chebyshev-based two-stage iterative method as an alternative to the direct solution of linear systems

Mario Arioli

m.arioli@rl.ac.uk

CCLRC-Rutherford Appleton Laboratory

with Daniel Ruiz (E.N.S.E.E.I.H.T)

CCLRC Rutherford Appleton Laboratory

# Outline

- Problems with Preconditioning
- Techniques for improving the solvers
- The two phase algorithm
- Error Analysis
- Influence of the different parameters
- Numerical experiments

CCLRC Rutherford Appleton Laboratory

# Problems with Preconditioning

$$\mathbf{A}\mathbf{u} = \mathbf{b}$$

$$\mathbf{A} \in \mathbb{R}^{N \times N} \text{ SPD}$$

# Problems with Preconditioning

$$\mathbf{A}\mathbf{u} = \mathbf{b}$$

$$\mathbf{A} \in \mathbb{R}^{N \times N} \text{ SPD}$$

$\mathbf{M} = \mathbf{R}^T\mathbf{R}$ a SPD preconditioner

$$\mathbf{R}^{-T}\mathbf{A}\mathbf{R}^{-1}\mathbf{R}\mathbf{u} = \mathbf{R}^{-T}\mathbf{b}$$

$$\widehat{\mathbf{A}} = \mathbf{R}^{-T}\mathbf{A}\mathbf{R}^{-1},\ \mathbf{x} = \mathbf{R}\mathbf{u},\ \hat{\mathbf{b}} = \mathbf{R}^{-T}\mathbf{b}.$$

# Problems with Preconditioning

$$\mathbf{A}\mathbf{u} = \mathbf{b}$$

$$\mathbf{A} \in \mathbb{R}^{N \times N} \text{ SPD}$$

$\mathbf{M} = \mathbf{R}^T\mathbf{R}$ a SPD preconditioner

$$\mathbf{R}^{-T}\mathbf{A}\mathbf{R}^{-1}\mathbf{R}\mathbf{u} = \mathbf{R}^{-T}\mathbf{b}$$

$\widehat{\mathbf{A}} = \mathbf{R}^{-T}\mathbf{A}\mathbf{R}^{-1}$, $\mathbf{x} = \mathbf{R}\mathbf{u}$, $\hat{\mathbf{b}} = \mathbf{R}^{-T}\mathbf{b}$.

- Good clusterization but problem still ill-conditioned

# A PDE example

$\Omega \subset \mathbf{R}^2$ simply connected bounded polygonal. Let

$$a(\mathbf{u}, \mathbf{v}) \qquad \forall \mathbf{u}, \mathbf{v} \in H_0^1(\Omega)$$

be a continuous and coercive bilinear form, and

$$L(\mathbf{v}) \in H^{-1}(\Omega).$$

The problem

$$\begin{cases} \text{Find } \mathbf{u} \in H_0^1(\Omega) \text{ such that} \\ a(\mathbf{u}, \mathbf{v}) = L(\mathbf{v}), \quad \forall \mathbf{v} \in H_0^1(\Omega), \end{cases}$$

has a unique solution.

# A PDE example

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega \mathfrak{K}(\mathbf{x})\nabla\mathbf{u}\cdot\nabla\mathbf{v}d\mathbf{x}, \quad \forall\mathbf{u},\mathbf{v}\in H_0^1(\Omega)$$

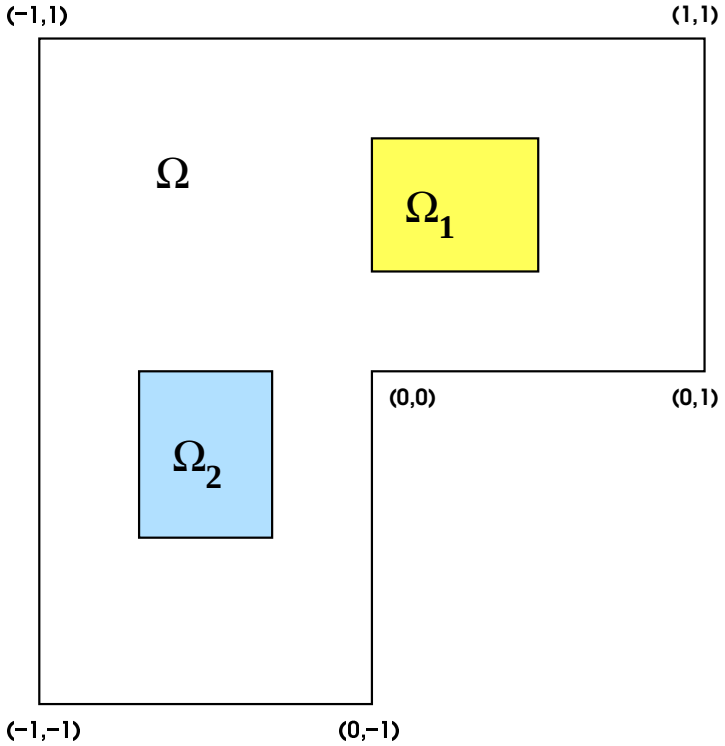
$$L(\mathbf{v}) = \int_\Omega 10\mathbf{v}d\mathbf{x}, \qquad \forall\mathbf{v}\in H_0^1(\Omega)$$

*

# A PDE example

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega \mathfrak{K}(\mathbf{x})\nabla\mathbf{u}\cdot\nabla\mathbf{v}d\mathbf{x}, \quad \forall \mathbf{u},\mathbf{v} \in H_0^1(\Omega)$$

$$L(\mathbf{v}) = \int_\Omega 10\mathbf{v}d\mathbf{x}, \qquad \forall \mathbf{v} \in H_0^1(\Omega)$$




$$\mathfrak{K}(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega \setminus \{\Omega_1 \cup \Omega_2\}, \\ 10^6 & \mathbf{x} \in \Omega_1, \\ 10^4 & \mathbf{x} \in \Omega_2. \end{cases}$$

*

# A PDE example

Using **pdetool**©, we generated a mesh satisfying the usual regularity conditions of Ciarlet and we computed a finite-element approximation of the problem with the use of continuous piece-wise linear elements. The approximated problem is equivalent to the following system of linear equations:

$$\mathbf{A}\mathbf{u} = \mathbf{b}.$$

In our mesh, the largest triangle has an area of $3.123 \times 10^{-4}$, therefore, the resulting linear system has $16256$ triangles, $8289$ nodes, and $7969$ degrees of freedom.

# A PDE example

We used three kinds of preconditioners: the classical Jacobi diagonal matrix, $\mathbf{M} = diag(\mathbf{A})$, the incomplete Cholesky decomposition of $\mathbf{A}$ with zero fill-in, and the incomplete Cholesky decomposition of $\mathbf{A}$ with drop tolerance $10^{-2}$. Using the incomplete Cholesky decompositions, we computed the upper triangular matrix $\mathbf{R}$ such that $\mathbf{M} = \mathbf{R}^T\mathbf{R}$.

CCLRC Rutherford Appleton Laboratory

# A PDE example

We used three kinds of preconditioners: the classical Jacobi diagonal matrix, $\mathbf{M} = diag(\mathbf{A})$, the incomplete Cholesky decomposition of $\mathbf{A}$ with zero fill-in, and the incomplete Cholesky decomposition of $\mathbf{A}$ with drop tolerance $10^{-2}$. Using the incomplete Cholesky decompositions, we computed the upper triangular matrix $\mathbf{R}$ such that $\mathbf{M} = \mathbf{R}^T\mathbf{R}$.

| $\mathbf{M}$ | $\kappa(\mathbf{M}^{-1}\mathbf{A})$ | $\lambda_{\min}$ | $\lambda_{\max}$ |
|---|---|---|---|
| $\mathbf{I}$ | $2.6\ 10^{9}$ | $3.7\ 10^{-3}$ | $9.6\ 10^{6}$ |
| Jacobi | $6.8\ 10^{8}$ | $3.1\ 10^{-9}$ | 2.08 |
| Inc. Cholesky(0) | $9.4\ 10^{7}$ | $1.7\ 10^{-8}$ | 1.6 |
| Inc. Cholesky($10^{-2}$) | $6.2\ 10^{6}$ | $1.8\ 10^{-7}$ | 1.1 |

CCLRC Rutherford Appleton Laboratory

# A PDE example

| | Preconditioner M | | |
|---|---|---|---|
| $\mu$ | Jacobi | Inc. Cholesky(0) | Inc. Cholesky($10^{-2}$) |
| $\lambda_{\max}/10^3$ | 3 | | |
| $\lambda_{\max}/500$ | 5 | | |
| $\lambda_{\max}/200$ | 18 | | |
| $\lambda_{\max}/100$ | 43 | 3 | |
| $\lambda_{\max}/50$ | | 11 | |
| $\lambda_{\max}/20$ | | 32 | |
| $\lambda_{\max}/10$ | $>200$ | 68 | 3 |
| $\lambda_{\max}/5$ | | | 9 |
| $\lambda_{\max}/2$ | | | 40 |

Number of eigenvalues in $[\lambda_{\min}, \mu]$.

CCLRC Rutherford Appleton Laboratory

# Problems with Preconditioning

- Good clusterization but problem still ill-conditioned

CCLRC Rutherford Appleton Laboratory

# Problems with Preconditioning

- Good clusterization but problem still ill-conditioned
- Several right-hand sides in succession

CCLRC Rutherford Appleton Laboratory

# Problems with Preconditioning

- Good clusterization but problem still ill-conditioned
- Several right-hand sides in succession
- Implicit matrix

CCLRC Rutherford Appleton Laboratory

# Techniques for improving the solvers

CCLRC Rutherford Appleton Laboratory

# Techniques for improving the solvers

- Deflation: compute the invariant space corresponding to the smallest eigenvalues

CCLRC Rutherford Appleton Laboratory

# Techniques for improving the solvers

- Deflation: compute the invariant space corresponding to the smallest eigenvalues
- Filtering + Lanczos

CCLRC Rutherford Appleton Laboratory

# The two-phase iterative approach

- Fix $\mu$ such that $\lambda_{\min}(\widehat{\mathbf{A}}) < \mu < \lambda_{\max}(\widehat{\mathbf{A}})$

CCLRC Rutherford Appleton Laboratory

# The two-phase iterative approach

- Fix $\mu$ such that $\lambda_{\min}(\widehat{\mathbf{A}}) < \mu < \lambda_{\max}(\widehat{\mathbf{A}})$
- Start with an initial set of $s$ randomly generated vectors ($s$ is the block-size), and "*damp*", in these starting vectors, the eigenfrequencies associated with all the eigenvalues in $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$.

CCLRC Rutherford Appleton Laboratory

# The two-phase iterative approach

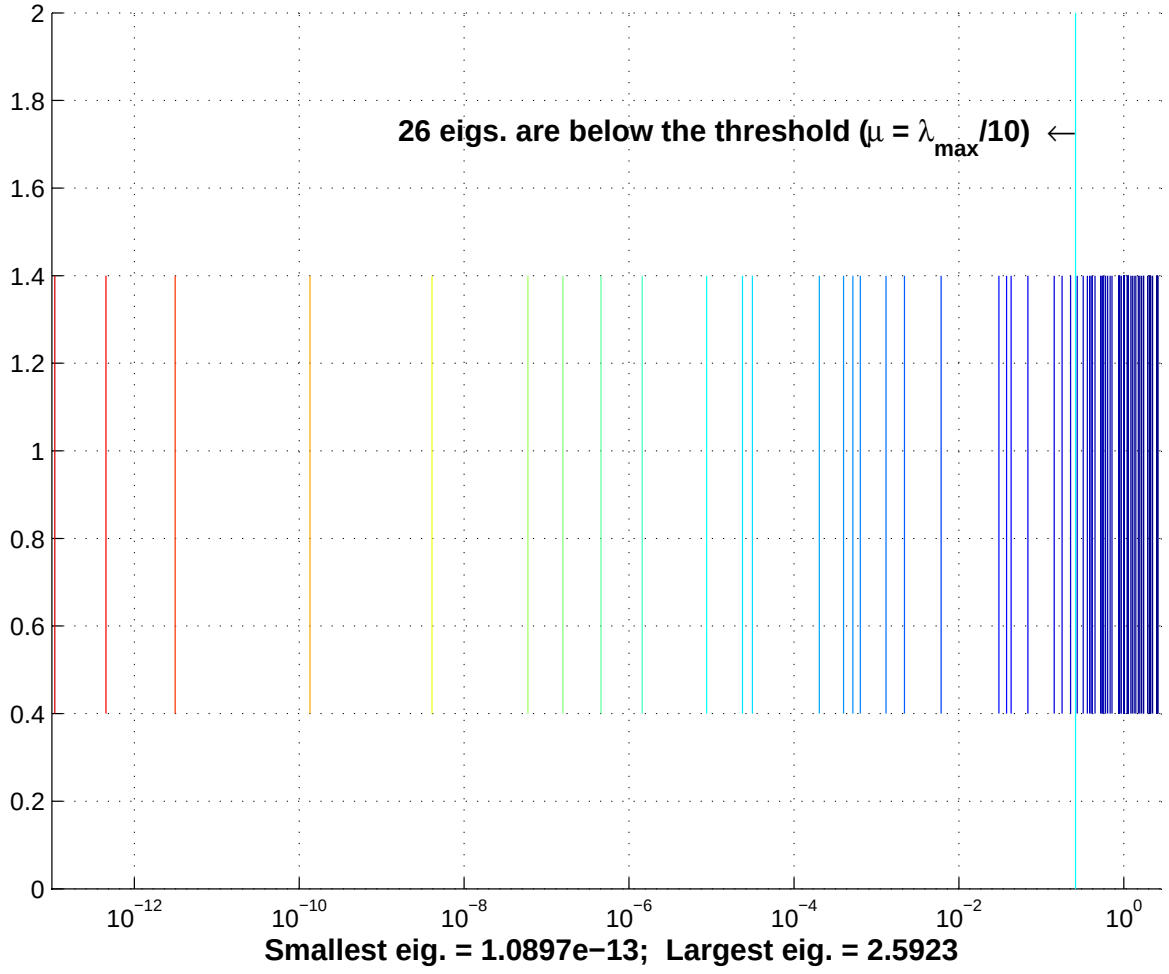
- Fix $\mu$ such that $\lambda_{\min}(\widehat{\mathbf{A}}) < \mu < \lambda_{\max}(\widehat{\mathbf{A}})$
- Start with an initial set of $s$ randomly generated vectors ($s$ is the block-size), and "*damp*", in these starting vectors, the eigenfrequencies associated with all the eigenvalues in $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$.
- Compute all the eigenvectors associated with all the eigenvalues in the range $[\lambda_{\min}(\widehat{\mathbf{A}}), \mu]$.

# The two-phase iterative approach

- Fix $\mu$ such that $\lambda_{\min}(\widehat{\mathbf{A}}) < \mu < \lambda_{\max}(\widehat{\mathbf{A}})$
- Start with an initial set of $s$ randomly generated vectors ($s$ is the block-size), and "*damp*", in these starting vectors, the eigenfrequencies associated with all the eigenvalues in $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$.
- Compute all the eigenvectors associated with all the eigenvalues in the range $[\lambda_{\min}(\widehat{\mathbf{A}}), \mu]$.

If the eigenvalues are well clustered the number of remaining eigenvalues in $[\lambda_{\min}(\widehat{\mathbf{A}}), \mu]$, with reasonable $\mu$ ($\lambda_{\max}/100$, or $\lambda_{\max}/10$), should be small compared to the size of the linear system.

# The two-phase iterative approach



# The two-phase iterative approach

We can use Chebyshev polynomials to *damp*

$$\begin{cases} T_0(\omega) = 1\,, \quad T_1(\omega) = \omega \\ T_{n+1}(\omega) = 2\omega T_n(\omega) - T_{n-1}(\omega) \qquad n \geq 1. \end{cases}.$$

If we consider $d > 1$ and

$$H_n(\omega) = \frac{T_n(\omega)}{T_n(d)},$$

then $H_n$ has minimum $l_\infty$ norm on the interval $[-1, 1]$ over all polynomials $Q_n$ of degree less than of equal to $n$ and satisfying the condition $Q_n(d) = 1$, and we have

$$\max_{\omega \in [-1,1]} H_n(\omega) = \frac{1}{T_n(d)}.$$

# The two-phase iterative approach

$$\lambda \in \mathbb{R} \longmapsto \omega_\mu(\lambda) = \frac{\lambda_{\max}(\widehat{\mathbf{A}}) + \mu - 2\lambda}{\lambda_{\max}(\widehat{\mathbf{A}}) - \mu},$$

with $\lambda_{\min}(\widehat{\mathbf{A}}) < \mu < \lambda_{\max}(\widehat{\mathbf{A}})$ given above.
This transformation maps $\lambda_{\max}(\widehat{\mathbf{A}})$ to $-1$, $\mu$ to $1$, and $0$ to

$$\omega_\mu(0) = d_\mu = \frac{\lambda_{\max}(\widehat{\mathbf{A}}) + \mu}{\lambda_{\max}(\widehat{\mathbf{A}}) - \mu} > 1.$$

Then,

$$\mathcal{P}_n(\lambda) = \frac{T_n(\omega_\mu(\lambda))}{T_n(d_\mu)},$$

# The two-phase iterative approach

$$\widehat{\mathbf{A}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

the eigendecomposition of the SPD matrix $\widehat{\mathbf{A}}$, with $\mathbf{\Lambda} = \mathrm{diag}(\lambda_i)$

# The two-phase iterative approach

$$\widehat{\mathbf{A}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

the eigendecomposition of the SPD matrix $\widehat{\mathbf{A}}$, with $\mathbf{\Lambda} = \mathrm{diag}(\lambda_i)$

$$\mathbf{z} = \sum_{i=1}^{n} \mathbf{u}_i \xi_i,$$

with $\xi_i = \mathbf{u}_i^T \mathbf{z}$, $i = 1, \ldots, n$,

# The two-phase iterative approach

$$\widehat{\mathbf{A}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

the eigendecomposition of the SPD matrix $\widehat{\mathbf{A}}$, with $\mathbf{\Lambda} = \mathrm{diag}(\lambda_i)$

$$\mathbf{z} = \sum_{i=1}^{n} \mathbf{u}_i \xi_i,$$

with $\xi_i = \mathbf{u}_i^T \mathbf{z}$, $i = 1, \ldots, n$,

$$\mathbf{v} = \mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{z} = \sum_{i=1}^{n} \mathbf{u}_i \left(\mathcal{P}_n(\lambda_i)\xi_i\right),$$

# The two-phase iterative approach

$$\widehat{\mathbf{A}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

the eigendecomposition of the SPD matrix $\widehat{\mathbf{A}}$, with $\mathbf{\Lambda} = \mathrm{diag}(\lambda_i)$

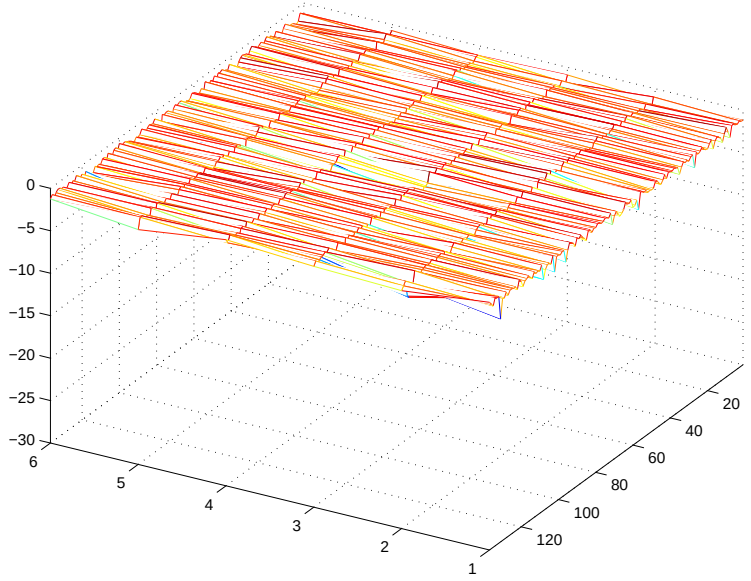
$$\mathbf{z} = \sum_{i=1}^{n} \mathbf{u}_i \xi_i,$$

with $\xi_i = \mathbf{u}_i^T \mathbf{z}$, $i = 1, \ldots, n$,

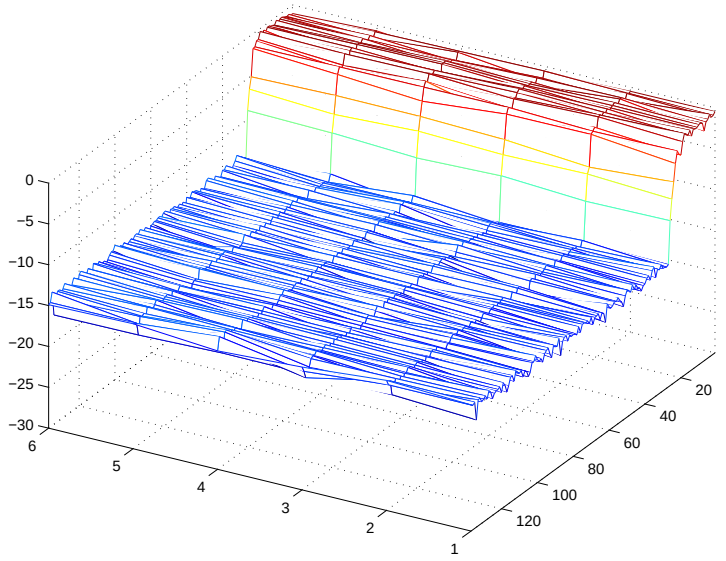
$$\mathbf{v} = \mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{z} = \sum_{i=1}^{n} \mathbf{u}_i \left(\mathcal{P}_n(\lambda_i)\xi_i\right),$$

The eigencomponents of the resulting vector $\mathbf{v}$ are close to that of the initial vector $\mathbf{z}$ for all $i$ such that $\lambda_i$ is close to $0$ and relatively much smaller for large enough degree $n$ and all $i$ such that $\lambda_i \in [\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$.

# The two-phase iterative approach




Eigencomponents of the starting set of orthonormal vectors (6 vectors, randomly generated, and orthonormalized).




Eigencomponents of this set of vectors after 51 Chebyshev iterations (the resulting 6 vectors have also been reorthonormalized).

- The eigenvectors of $\mathbf{A}$ are indexed on the X-axis (from 1 to 136) in increasing order of their corresponding eigenvalue.
- The indexes (from 1 to 6) of the vectors in the current set are indicated on the Y-axis.
- The Z-axis indicates the logarithm of the absolute values of the eigencomponents in each of the 6 vectors.

CCLRC Rutherford Appleton Laboratory

# The two-phase iterative approach

We can interpret the use of Chebyshev polynomials as a *filtering* tool that increases the degree of colinearity with some selected eigenvectors. Then, after "*filtering*" the initial starting vectors, we obtain a set of $s$ vectors with eigencomponents below a certain level $\varepsilon$ for those eigenvalues in the range $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$, and relatively much bigger eigencomponents linked with the smallest eigenvalues in $\widehat{\mathbf{A}}$.

# The two-phase iterative approach

We can interpret the use of Chebyshev polynomials as a *filtering* tool that increases the degree of colinearity with some selected eigenvectors. Then, after "*filtering*" the initial starting vectors, we obtain a set of $s$ vectors with eigencomponents below a certain level $\varepsilon$ for those eigenvalues in the range $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$, and relatively much bigger eigencomponents linked with the smallest eigenvalues in $\widehat{\mathbf{A}}$.

The block-size $s$ v.s. the number $k$ of remaining eigenvalues outside the interval $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$.

# The two-phase iterative approach

We can interpret the use of Chebyshev polynomials as a *filtering* tool that increases the degree of colinearity with some selected eigenvectors. Then, after "*filtering*" the initial starting vectors, we obtain a set of $s$ vectors with eigencomponents below a certain level $\varepsilon$ for those eigenvalues in the range $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$, and relatively much bigger eigencomponents linked with the smallest eigenvalues in $\widehat{\mathbf{A}}$.

The block-size $s$ v.s. the number $k$ of remaining eigenvalues outside the interval $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$.

- $k \leq s$, a Ritz decomposition enables us to get the $k$ wanted eigenvectors.

# The two-phase iterative approach

We can interpret the use of Chebyshev polynomials as a *filtering* tool that increases the degree of colinearity with some selected eigenvectors. Then, after "*filtering*" the initial starting vectors, we obtain a set of $s$ vectors with eigencomponents below a certain level $\varepsilon$ for those eigenvalues in the range $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$, and relatively much bigger eigencomponents linked with the smallest eigenvalues in $\widehat{\mathbf{A}}$.

The block-size $s$ v.s. the number $k$ of remaining eigenvalues outside the interval $[\mu, \lambda_{\max}(\widehat{\mathbf{A}})]$.

- $k \leq s$, a Ritz decomposition enables us to get the $k$ wanted eigenvectors.
- $k > s$, we use a Block-Lanczos type of approach to build a Krylov basis starting with these filtered vectors and to stop when appropriate.

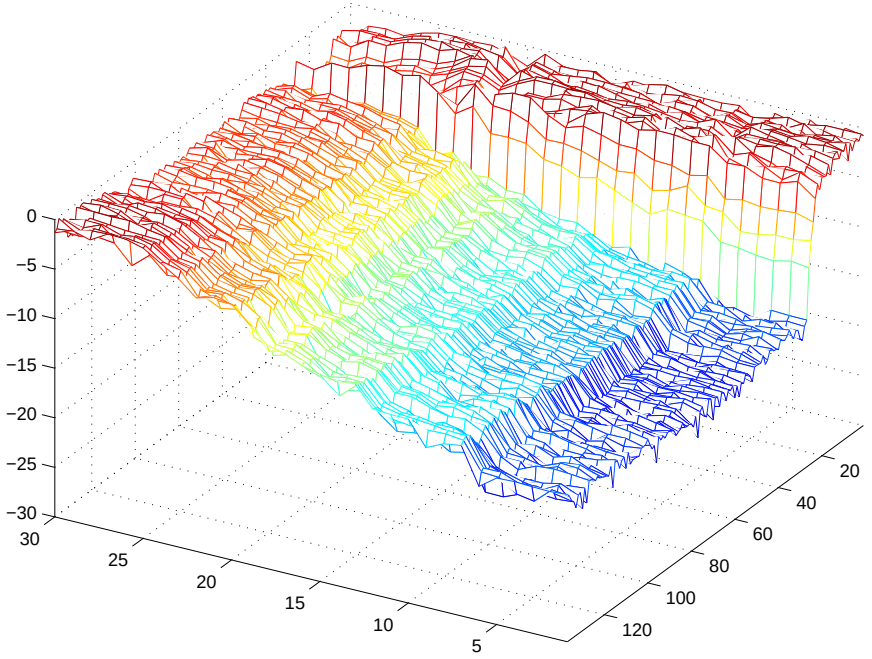
CCLRC Rutherford Appleton Laboratory

# The two-phase iterative approach

Lanczos/Block-Lanczos algorithm does not maintain the nice property of the filtered vectors.

# The two-phase iterative approach

Lanczos/Block-Lanczos algorithm does not maintain the nice property of the filtered vectors.




Eigendecomposition of the Krylov basis obtained after 5 block-Lanczos steps (Block Lanczos with block size 6 and full classical Gram-Schmidt reorthogonalization).

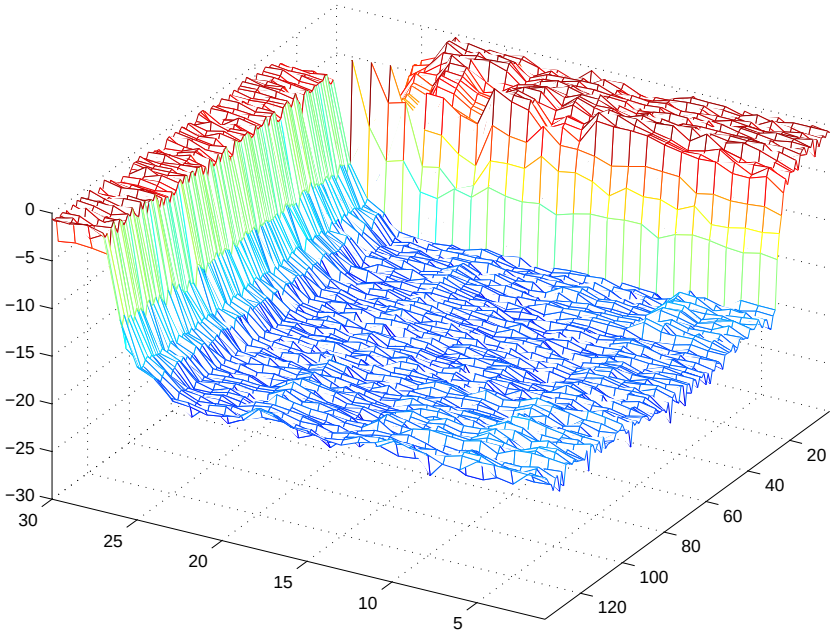
CCLRC Rutherford Appleton Laboratory

# Lanczos/Orthodir approach

To maintain the level of the unwanted eigenfrequencies in the orthonormal block Krylov basis under $\varepsilon$, perform, at each Block-Lanczos iteration, a few extra Chebyshev iterations on the newly generated Block Lanczos vectors $\mathbf{V}^{(k+1)}$.

# Lanczos/Orthodir approach

To maintain the level of the unwanted eigenfrequencies in the orthonormal block Krylov basis under $\varepsilon$, perform, at each Block-Lanczos iteration, a few extra Chebyshev iterations on the newly generated Block Lanczos vectors $\mathbf{V}^{(k+1)}$.




Eigendecomposisiton of the Krylov basis obtained after 5 block-Lanczos/Orthodir steps with additional Chebyshev filtering at each Lanczos iteration

# Partial Spectral Factorization phase

$$\mathbf{Z} = \text{Chebyshev\_Filter}(\mathbf{Y}, \varepsilon, [\mu, \lambda_{\max}], \widehat{\mathbf{A}})$$

is the application of a Chebyshev polynomial in $\widehat{\mathbf{A}}$ to $\mathbf{Y}$, viz.

$$\mathbf{Z} = \mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{Y},$$

where $\mathcal{P}_n$ has a degree such that its $\mathrm{L}_\infty$ norm over the interval $[\mu, \lambda_{\max}]$ is less than $\varepsilon$.
Fixing the block size to $s$, the cut-off eigenvalue $\mu$, $\lambda_{\min}(\widehat{\mathbf{A}}) < \mu < \lambda_{\max}(\widehat{\mathbf{A}})$, and the filtering level $\varepsilon \ll 1$:

CCLRC Rutherford Appleton Laboratory

# Partial Spectral Factorization phase

$\mathbf{P}^{(0)} = \text{random}(n, s)$; and $\mathbf{Q}^{(0)} = \text{orthonormalize}(\mathbf{P}^{(0)})$

$\mathbf{Z}^{(0)} = \text{Chebyshev\_Filter}(\mathbf{Q}^{(0)}, \varepsilon, [\mu, \lambda_{\max}], \widehat{\mathbf{A}})$

$[\mathbf{Y}^{(0)}, \mathbf{\Sigma}_1^{(0)}, \mathbf{W}] = \text{SVD}(\mathbf{Z}^{(0)}, 0)$; and $\delta_0 = \sigma_{\min}(\mathbf{\Sigma}_1^{(0)})$

$\widehat{\mathbf{Z}}^{(0)} = \text{Chebyshev\_Filter}(\mathbf{Y}^{(0)}, \delta_0, [\mu, \lambda_{\max}], \widehat{\mathbf{A}})$; and $\mathbf{V}^{(0)} = \text{orthonormalize}(\widehat{\mathbf{Z}}^{(0)})$

$\mathbf{V} = \mathbf{V}^{(0)}$; and set $\delta_2 = 1$

for $k = 0, 1, 2, \ldots,$ until convergence do :

$\mathbf{P}^{(k+1)} = \widehat{\mathbf{A}}\mathbf{V}^{(k)} - \mathbf{V}\mathbf{V}^T\widehat{\mathbf{A}}\mathbf{V}^{(k)}$; $[\mathbf{V}^{(k+1)}, \mathbf{\Sigma}_1^{(k+1)}, \mathbf{W}] = \text{SVD}(\mathbf{P}^{(k+1)}, 0)$;

set $\delta_1 = \sigma_{\min}(\mathbf{\Sigma}_1^{(k+1)})/\lambda_{\max}$ and $\delta = \max(\varepsilon, \delta_1 \times \delta_2)$

for $i = 1, 2$

$\mathbf{Z}^{(k+1)} = \text{Chebyshev\_Filter}(\mathbf{V}^{(k+1)}, \delta, [\mu, \lambda_{\max}], \widehat{\mathbf{A}})$

$\mathbf{Y}^{(k+1)} = \mathbf{Z}^{(k+1)} - \mathbf{V}\mathbf{V}^T\mathbf{Z}^{(k+1)}$

$[\mathbf{V}^{(k+1)}, \mathbf{\Sigma}_2^{(k+1)}, \mathbf{W}] = \text{SVD}(\mathbf{Y}^{(k+1)}, 0)$

$\delta_2 = \sigma_{\min}(\mathbf{\Sigma}_2^{(k+1)})$ and set $\delta = \delta_2$

end

$\mathbf{V} = [\mathbf{V}; \mathbf{V}^{(k+1)}]$

# Solution of the system

Once the *near*-invariant subspace linked to the smallest eigenvalues is obtained, we can use it for the computation of further solutions. The idea is to perform an oblique projection of the initial residual ($\hat{\mathbf{r}}_0 = \hat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{x}_0$) onto this *near* invariant subspace in order to get the eigencomponents in the solution corresponding to the smallest eigenvalues, viz.

$$\hat{\mathbf{r}}_1 = \hat{\mathbf{r}}_0 - \widehat{\mathbf{A}}\mathbf{V}\left(\mathbf{V}^T\widehat{\mathbf{A}}\mathbf{V}\right)^{-1}\mathbf{V}^T\hat{\mathbf{r}}_0,$$

$$\text{with} \quad \mathbf{x}_1 = \mathbf{x}_0 + \mathbf{V}\left(\mathbf{V}^T\widehat{\mathbf{A}}\mathbf{V}\right)^{-1}\mathbf{V}^T\hat{\mathbf{r}}_0.$$

To compute the remaining part of the solution vector $\widehat{\mathbf{A}}\mathbf{x}_2 = \hat{\mathbf{r}}_1$, one can then use the classical Chebyshev algorithm with eigenvalue bounds given by $\mu$ and $\lambda_{\max}(\widehat{\mathbf{A}})$.

CCLRC Rutherford Appleton Laboratory

# Solution phase

$$[\hat{\mathbf{r}}_1, \mathbf{x}_1] = \text{Chebyshev\_Solve}(\hat{\mathbf{r}}_0, \mathbf{x}_0, \delta, [\mu, \lambda_{\max}], \widehat{\mathbf{A}})$$

# Solution phase

$$[\hat{\mathbf{r}}_1, \mathbf{x}_1] = \text{Chebyshev\_Solve}(\hat{\mathbf{r}}_0, \mathbf{x}_0, \delta, [\mu, \lambda_{\max}], \widehat{\mathbf{A}})$$

## Solution phase

For any right-hand side vector $\hat{\mathbf{b}}$, set $\mathbf{x}_0$ and $\hat{\mathbf{r}}_0 = \hat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{x}_0$, and perform the following consecutive steps:

$$[\hat{\mathbf{r}}_1, \mathbf{x}_1] = \text{Chebyshev\_Solve}(\hat{\mathbf{r}}_0, \mathbf{x}_0, \varepsilon, [\mu, \lambda_{\max}], \widehat{\mathbf{A}})$$

$$\hat{\mathbf{r}} = \hat{\mathbf{r}}_1 - \widehat{\mathbf{A}}\mathbf{V}\left(\mathbf{V}^T\widehat{\mathbf{A}}\mathbf{V}\right)^{-1}\mathbf{V}^T\hat{\mathbf{r}}_1; \quad \text{and}$$

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{V}\left(\mathbf{V}^T\widehat{\mathbf{A}}\mathbf{V}\right)^{-1}\mathbf{V}^T\hat{\mathbf{r}}_1$$

# Error Analysis

$$
\begin{aligned}
||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}} &= ||\mathbf{x}^* - \mathbf{x}||_{\widehat{\mathbf{A}}} \qquad \mathbf{e}_0 = \widehat{\mathbf{A}}^{-1}\hat{\mathbf{r}}_0 = \mathbf{x}^* - \mathbf{x}_0 \\
\mathbf{v} &= \mathcal{P}_n(\widehat{\mathbf{A}})\hat{\mathbf{r}}_0 / ||\mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{e}_0||_2 .
\end{aligned}
$$

# Error Analysis

$$
\begin{aligned}
||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}} &= ||\mathbf{x}^* - \mathbf{x}||_{\widehat{\mathbf{A}}} \qquad \mathbf{e}_0 = \widehat{\mathbf{A}}^{-1}\hat{\mathbf{r}}_0 = \mathbf{x}^* - \mathbf{x}_0 \\
\mathbf{v} &= \mathcal{P}_n(\widehat{\mathbf{A}})\hat{\mathbf{r}}_0 / ||\mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{e}_0||_2 .
\end{aligned}
$$

$$
\frac{||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}}}{||\hat{\mathbf{r}}_0||_{\widehat{\mathbf{A}}^{-1}}} \leq ||\widehat{\mathbf{A}}^{\frac{1}{2}}||_2 ||\widehat{\mathbf{A}}^{-\frac{1}{2}}||_2 ||(\mathbf{I} - \mathbf{V}\mathbf{V}^T)\mathbf{v}||_2 .
$$

# Error Analysis

$$
\begin{aligned}
||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}} &= ||\mathbf{x}^* - \mathbf{x}||_{\widehat{\mathbf{A}}} \qquad \mathbf{e}_0 = \widehat{\mathbf{A}}^{-1}\hat{\mathbf{r}}_0 = \mathbf{x}^* - \mathbf{x}_0 \\
\mathbf{v} &= \mathcal{P}_n(\widehat{\mathbf{A}})\hat{\mathbf{r}}_0 / ||\mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{e}_0||_2 .
\end{aligned}
$$

$$
\frac{||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}}}{||\hat{\mathbf{r}}_0||_{\widehat{\mathbf{A}}^{-1}}} \leq ||\widehat{\mathbf{A}}^{\frac{1}{2}}||_2 ||\widehat{\mathbf{A}}^{-\frac{1}{2}}||_2 ||(\mathbf{I} - \mathbf{V}\mathbf{V}^T)\mathbf{v}||_2 .
$$

Theorem Let $\mathbf{U}_1 \in \mathbb{R}^{n \times m}$ be the matrix of the eigenvectors of $\widehat{\mathbf{A}}$ corresponding to $\mathbf{\Lambda}_1$ and $\mathbf{U}_2 \in \mathbb{R}^{n \times (n-m)}$ be the matrix of the remaining eigenvectors of $\widehat{\mathbf{A}}$. Let $\mathbf{V} \in \mathbb{R}^{n \times \ell}$ be the full basis generated by the Algorithm using a filtering level $\varepsilon$ and $\mathbf{v} = \mathcal{P}_n(\widehat{\mathbf{A}})\hat{\mathbf{r}}_0 / ||\mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{e}_0||_2$. If $m\varepsilon \ll 1$ and $\ell \geq m$ then

$$
||(\mathbf{I} - \mathbf{V}\mathbf{V}^T)\mathbf{v}||_2 \leq 2\varepsilon\sqrt{m} + 2\varepsilon^2 m.
$$

# Error Analysis

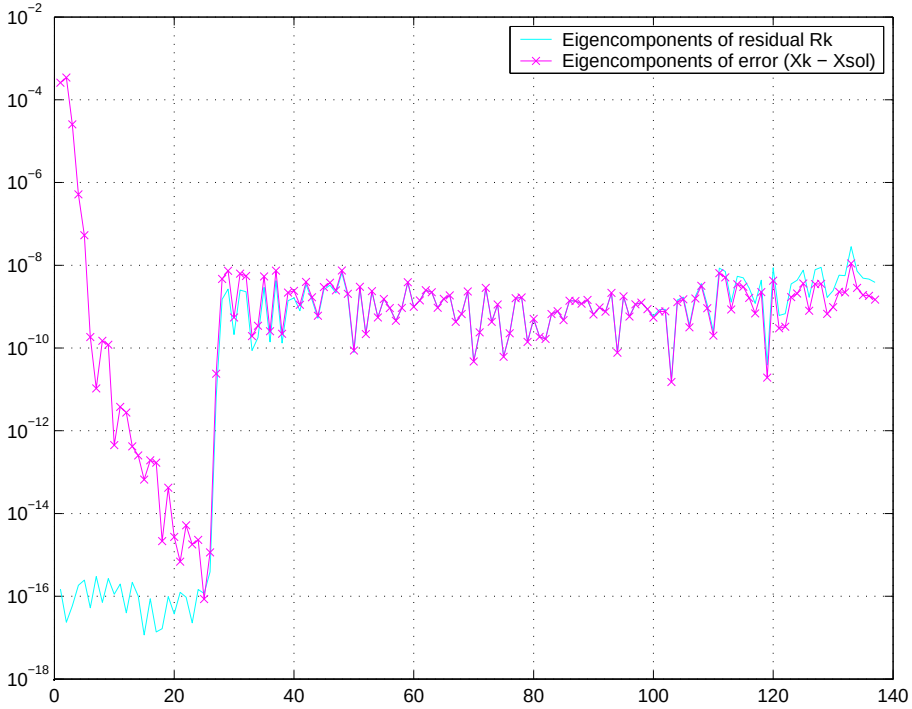
$$
\begin{aligned}
||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}} &= ||\mathbf{x}^* - \mathbf{x}||_{\widehat{\mathbf{A}}} \qquad \mathbf{e}_0 = \widehat{\mathbf{A}}^{-1}\hat{\mathbf{r}}_0 = \mathbf{x}^* - \mathbf{x}_0 \\
\mathbf{v} &= \mathcal{P}_n(\widehat{\mathbf{A}})\hat{\mathbf{r}}_0/||\mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{e}_0||_2.
\end{aligned}
$$

$$
\frac{||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}}}{||\hat{\mathbf{r}}_0||_{\widehat{\mathbf{A}}^{-1}}} \leq 4\sqrt{m}\varepsilon||\widehat{\mathbf{A}}^{\frac{1}{2}}||_2||\widehat{\mathbf{A}}^{-\frac{1}{2}}||_2.
$$

Theorem Let $\mathbf{U}_1 \in \mathbb{R}^{n\times m}$ be the matrix of the eigenvectors of $\widehat{\mathbf{A}}$ corresponding to $\mathbf{\Lambda}_1$ and $\mathbf{U}_2 \in \mathbb{R}^{n\times(n-m)}$ be the matrix of the remaining eigenvectors of $\widehat{\mathbf{A}}$. Let $\mathbf{V} \in \mathbb{R}^{n\times\ell}$ be the full basis generated by the Algorithm using a filtering level $\varepsilon$ and $\mathbf{v} = \mathcal{P}_n(\widehat{\mathbf{A}})\hat{\mathbf{r}}_0/||\mathcal{P}_n(\widehat{\mathbf{A}})\mathbf{e}_0||_2$. If $m\varepsilon \ll 1$ and $\ell \geq m$ then

$$
||(\mathbf{I} - \mathbf{V}\mathbf{V}^T)\mathbf{v}||_2 \leq 2\varepsilon\sqrt{m} + 2\varepsilon^2 m.
$$

# Error Analysis




Computation of the solution of a linear system with a right-hand side vector $\mathbf{b}$ corresponding to a given random exact solution vector $\mathbf{x}^*$. The filtering level $\varepsilon$ has been fixed at $10^{-8}$ in both phases of the algorithm.

CCLRC Rutherford Appleton Laboratory

# Influence of the different parameters

- The choice of the starting block size $s$

CCLRC Rutherford Appleton Laboratory

# Influence of the different parameters

- The choice of the starting block size $s$
- The choice of the cut-off eigenvalue $\mu$

CCLRC Rutherford Appleton Laboratory

# Influence of the different parameters

- The choice of the starting block size $s$
- The choice of the cut-off eigenvalue $\mu$
- The choice of the filtering level $\varepsilon$

CCLRC Rutherford Appleton Laboratory

# Influence of the different parameters

- The choice of the starting block size $s$
- The choice of the cut-off eigenvalue $\mu$
- The choice of the filtering level $\varepsilon$
  - $\varepsilon = 10^{-8}$

CCLRC Rutherford Appleton Laboratory

# Influence of the different parameters

- The choice of the starting block size $s$
- The choice of the cut-off eigenvalue $\mu$
- The choice of the filtering level $\varepsilon$
  - $\varepsilon = 10^{-8}$
  - $$\frac{||\hat{\mathbf{r}}||_{\widehat{\mathbf{A}}^{-1}}}{||\hat{\mathbf{r}}_0||_{\widehat{\mathbf{A}}^{-1}}} \leq \tau,$$
    $$\varepsilon = \frac{\tau}{\sqrt{\kappa(\widehat{\mathbf{A}})}}.$$

$\kappa(\widehat{\mathbf{A}}) = \|\widehat{\mathbf{A}}\|\|\widehat{\mathbf{A}}^{-1}\|$ "*a posteriori*" good approximation

CCLRC Rutherford Appleton Laboratory

# Influence of the different parameters

| | Number of Chebyshev Iterations ($s = 6$) | | | | | |
|---|---|---|---|---|---|---|
| | $\mu = \lambda_{\max}/5$ | | $\mu = \lambda_{\max}/10$ | | $\mu = \lambda_{\max}/100$ | |
| Block Krylov Iteration | Value of $\varepsilon$ | | Value of $\varepsilon$ | | Value of $\varepsilon$ | |
| | $10^{-14}$ | $10^{-8}$ | $10^{-14}$ | $10^{-8}$ | $10^{-14}$ | $10^{-8}$ |
| Start | 35 + 2 | 20 + 3 | 51 + 4 | 30 + 3 | 165 + 14 | 96 + 14 |
| 1 | 10 | 9 | 15 | 15 | 59 | 56 |
| 2 | 11 | 11 | 20 | 18 | 88 | 90 |
| 3 | 16 | 15 | 32 | 30 | 165 | 96 |
| 4 | 25 | 20 | 43 | 30 | - | - |
| 5 | 35 | 20 | - | - | - | - |

In the case $\mu = \lambda_{\max}/5$, there are 33 eigenvectors to capture.

In the case $\mu = \lambda_{\max}/10$, there are 26 eigenvectors to capture.

In the case $\mu = \lambda_{\max}/100$, there are 19 eigenvectors to capture.

# Numerical experiments

Back to the starting example

CCLRC Rutherford Appleton Laboratory

# Numerical experiments

| $\mu = \lambda_{\max}/\gamma$ | Spectral Factorization | | Solution phase | |
|---|---|---|---|---|
| $\gamma$ | Tot. Chebyshev Iterations | Size **V** | Chebyshev Iterations | Error Energy Norm |
| Jacobi | | | | |
| 1000 | 1030 (1004) | 3 | 231 | $7\ 10^{-3}$ ($2.6\ 10^{-5}$) |
| 500 | 1101 (1114) | 5 | 163 | $6\ 10^{-4}$ ($7.0\ 10^{-6}$) |
| 200 | 2234 (2615) | 1 | 103 | $2.5\ 10^{-4}$ ($4\ 10^{-6}$) |
| Inc. Cholesky(0) | | | | |
| 100 | 433 (248) | 5 (3) | 68 | $7.4\ 10^{-3}$ ($1.8\ 10^{-5}$) |
| 50 | 462 (503) | 9 | 48 | $3\ 10^{-3}$ ($1.3\ 10^{-5}$) |
| Inc. Cholesky($10^{-2}$) | | | | |
| 10 | 55 (70) | 3 | 19 | $8.2\ 10^{-3}$ ($3.3\ 10^{-6}$) |

Summary of the results

CCLRC
Rutherford Appleton Laboratory

# THE END