Varying the s in s-step GMRES

David Imberti and Jocelyne Erhel

FLUMINANCE team, Inria Rennes, France

NL2A
Marseille, October 2016

$Ax = b$, $A \in \mathbb{R}^{n \times n}$ $x, b \in \mathbb{R}^n$ $B = AM^{-1}$ $x_0 \in \mathbb{R}^n$ $r_0 = b - Ax_0$

GMRES(m): a Krylov subspace method

- [Saad and Schultz 1986, Meurant's book 1999, Saad's book 2003, Simoncini and Szyld 2007, Erhel 2011, ...]
- $\mathcal{K}_m(B, r_0) = span\{r_0, Br_0, \ldots B^{m-1} r_0\}$
- Find $x_m \in x_0 + M^{-1}\mathcal{K}_m(B, r_0)$ such that
  $\|r_m\|_2 = \|b - Bx_m\|_2 = \min_{x \in x_0 + M^{-1}\mathcal{K}_m(B,r_0)} \|b - Bx\|_2$

Building blocks of GMRES(m)

- Build an orthonormal basis V_{k+1} of the Krylov subspace $\mathcal{K}_{k+1}$
  get the Arnoldi-like relation $BV_k = V_{k+1}H_k$ for $k = 1, \ldots, m$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_k y$ implies $r = r_0 - BV_k y = V_{k+1}(\beta e_1 - H_k y)$
  Solve the least-squares problem: $\min_{y \in \mathbb{R}^k} \|\beta e_1 - H_k y\|$
- Restart if not converged

$$x_0 = x_0 + M^{-1}V_m y_m$$

Preconditioned GMRES

$Ax = b, \quad A \in \mathbb{R}^{n \times n} \quad x, b \in \mathbb{R}^n \quad B = AM^{-1} \quad x_0 \in \mathbb{R}^n \quad r_0 = b - Ax_0$

GMRES(m): a Krylov subspace method

- [Saad and Schultz 1986, Meurant's book 1999, Saad's book 2003, Simoncini and Szyld 2007, Erhel 2011, ...]
- $\mathcal{K}_m(B, r_0) = span\{r_0, Br_0, \ldots B^{m-1} r_0\}$
- Find $x_m \in x_0 + M^{-1}\mathcal{K}_m(B, r_0)$ such that $\|r_m\|_2 = \|b - Bx_m\|_2 = \min_{x \in x_0 + M^{-1}\mathcal{K}_m(B,r_0)} \|b - Bx\|_2$

Building blocks of GMRES(m)

- Build an orthonormal basis V_{k+1} of the Krylov subspace $\mathcal{K}_{k+1}$ get the Arnoldi-like relation $BV_k = V_{k+1} H_k$ for $k = 1, \ldots, m$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1} V_k y$ implies $r = r_0 - BV_k y = V_{k+1}(\beta e_1 - H_k y)$ Solve the least-squares problem: $\min_{y \in \mathbb{R}^k} \|\beta e_1 - H_k y\|$
- Restart if not converged

$$x_0 = x_0 + M^{-1} V_m y_m$$

GMRES ... practical issues

Arnoldi process

```
1:  v_1 = r_0/||r_0||_2
2:  for k = 1, m do
3:      p = B v_k
4:      for i = 1 : k do
5:          h_ik = v_i^T p
6:          p = p − h_ik v_i
7:      end for
8:      h_{k+1,k} = ||p||_2
9:      v_{k+1} = p/h_{k+1,k}
10: end for
```

⇓

$$BV_m = V_{m+1}H_m$$

Preconditioning issues

⇒ use multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009, Smith et al's book 1996,...]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

⇒ Use deflation to recover possible loss of information

- Deflation by preconditioning [Erhel et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Granularity issues in parallel algorithms

⇒ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [De Sturler 1994, Erhel 1995, Sidje 1997]
- Compute the basis block by block [Hoemmen 2010, Demmel et al 2011]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam 2011, Nuentsa Wakam+Erhel+Gropp 2013, Nuentsa Wakam+Pacull 2013, Nuentsa Wakam+Erhel 2014]

GMRES ... practical issues

Arnoldi process

1: $v_1 = r_0 / \|r_0\|_2$
2: **for** $k = 1, m$ **do**
3: $p = Bv_k$
4: **for** $i = 1 : k$ **do**
5: $h_{ik} = v_i^T p$
6: $p = p - h_{ik} v_i$
7: **end for**
8: $h_{k+1,k} = ||p||_2$
9: $v_{k+1} = p / h_{k+1,k}$
10: **end for**

$\Downarrow$

$BV_m = V_{m+1} H_m$

Preconditioning issues

$\Rightarrow$ use multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009, Smith et al's book 1996,...]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ Use deflation to recover possible loss of information

- Deflation by preconditioning [Erhel et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [De Sturler 1994, Erhel 1995, Sidje 1997]
- Compute the basis block by block [Hoemmen 2010, Demmel et al 2011]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam 2011, Nuentsa Wakam+Erhel+Gropp 2013, Nuentsa Wakam+Pacull 2013, Nuentsa Wakam+Erhel 2014]

GMRES ... practical issues

Arnoldi process

$v_1 = r_0/\|r_0\|_2$
for $k = 1, m$ **do**
  $p = Bv_k$
  for $i = 1 : k$ **do**
    $h_{ik} = v_i^T p$
    $p = p - h_{ik} v_i$
  end for
  $h_{k+1,k} = ||p||_2$
  $v_{k+1} = p/h_{k+1,k}$
end for

$\Downarrow$

$BV_m = V_{m+1} H_m$

Preconditioning issues

$\Rightarrow$ use multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009, Smith et al's book 1996,...]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ Use deflation to recover possible loss of information

- Deflation by preconditioning [Erhel et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [De Sturler 1994, Erhel 1995, Sidje 1997]
- Compute the basis block by block [Hoemmen 2010, Demmel et al 2011]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam 2011, Nuentsa Wakam+Erhel+Gropp 2013, Nuentsa Wakam+Pacull 2013, Nuentsa Wakam+Erhel 2014]

GMRES ... practical issues

Arnoldi process

$v_1 = r_0 / \|r_0\|_2$
for $k = 1, m$ **do**
  $p = Bv_k$
  for $i = 1 : k$ **do**
    $h_{ik} = v_i^T p$
    $p = p - h_{ik} v_i$
  end for
  $h_{k+1,k} = ||p||_2$
  $v_{k+1} = p / h_{k+1,k}$
end for

$\Downarrow$

$BV_m = V_{m+1} H_m$

Preconditioning issues

$\Rightarrow$ use multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009, Smith et al's book 1996,...]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ Use deflation to recover possible loss of information

- Deflation by preconditioning [Erhel et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [De Sturler 1994, Erhel 1995, Sidje 1997]
- Compute the basis block by block [Hoemmen 2010, Demmel et al 2011]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam 2011, Nuentsa Wakam+Erhel+Gropp 2013, Nuentsa Wakam+Pacull 2013, Nuentsa Wakam+Erhel 2014]

GMRES ... practical issues

Arnoldi process

$v_1 = r_0/\|r_0\|_2$
for $k = 1, m$ **do**
  $p = Bv_k$
  for $i = 1 : k$ **do**
    $h_{ik} = v_i^T p$
    $p = p - h_{ik} v_i$
  end for
  $h_{k+1,k} = ||p||_2$
  $v_{k+1} = p/h_{k+1,k}$
end for

$\Downarrow$

$BV_m = V_{m+1} H_m$

Preconditioning issues

$\Rightarrow$ use multilevel methods to deal with large systems

- Schwarz preconditioning [Atenekeng Kahou et al 2007, Dufaud+Tromeur-Dervout 2010, Giraud+Haidar 2009, Smith et al's book 1996,...]
- Filtering and Schur complement [Li et al 2003, Grigori et al 2011]
- Multilevel parallelism [Nuentsa Wakam et al 2011, Giraud et al 2010, ...]

Complexity and stagnation issues with restarted GMRES(m)

$\Rightarrow$ Use deflation to recover possible loss of information

- Deflation by preconditioning [Erhel et al 1996, Burrage et al 1998, Baglama et al 1998, ...]
- Deflation by augmented basis [Morgan 1995, Morgan 2002,...]

Granularity issues in parallel algorithms

$\Rightarrow$ Communication-avoiding strategies

- Generate the basis vectors [Reichel 1990, Bai et al 1994]
- Orthogonalize the basis [De Sturler 1994, Erhel 1995, Sidje 1997]
- Compute the basis block by block [Hoemmen 2010, Demmel et al 2011]

Strategy

Combine 'communication-avoiding' GMRES ... and Deflation ... and Domain Decomposition preconditioners
[Nuentsa Wakam 2011, Nuentsa Wakam+Erhel+Gropp 2013, Nuentsa Wakam+Pacull 2013, Nuentsa Wakam+Erhel 2014]

Basis computation and orthogonalization

- m-step GMRES(m)
- fixed s-step GMRES(m)
- variable s-step GMRES(m)

Building blocks of m-step GMRES(m)

- Build a basis W_m of the Krylov subspace $\mathcal{K}_m$
- Build an orthonormal basis V_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:

$$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$

- Alternative to minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:

$$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$

- Restart if not converged

Building blocks of m-step GMRES(m)

- Build a basis W_m of the Krylov subspace $\mathcal{K}_m$
- Build an orthonormal basis V_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$

- Alternative to minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_my\|$$

- Restart if not converged

Building blocks of m-step GMRES(m)

- Build a basis W_m of the Krylov subspace $\mathcal{K}_m$
- Build an orthonormal basis V_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_m R_m^{-1}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_m y$ implies $r = r_0 - BV_m y = V_{m+1}(\beta e_1 - H_m R_m^{-1} y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_m R_m^{-1} y\|$$
- Alternative to minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}W_m y$ implies $r = r_0 - BW_m y = V_{m+1}(\beta e_1 - H_m y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_m y\|$$
- Restart if not converged

Building blocks of m-step GMRES(m)

- Build a basis W_m of the Krylov subspace $\mathcal{K}_m$
- Build an orthonormal basis V_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_m R_m^{-1}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1}V_m y$ implies $r = r_0 - BV_m y = V_{m+1}(\beta e_1 - H_m R_m^{-1} y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_m R_m^{-1} y\|$$

- Alternative to minimize the residual in the Krylov subspace $x = x_0 + M^{-1}W_m y$ implies $r = r_0 - BW_m y = V_{m+1}(\beta e_1 - H_m y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_m y\|$$

- Restart if not converged

Building blocks of m-step GMRES(m)

- Build a basis W_m of the Krylov subspace $\mathcal{K}_m$
- Build an orthonormal basis V_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$
  get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$
  Solve the least-squares problem:

$$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$

- Alternative to minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$
  Solve the least-squares problem:

$$\min_{y\in\mathbb{R}^m} \|\beta e_1 - H_my\|$$

- Restart if not converged

Building blocks of m-step GMRES(m)

- Build a basis W_m of the Krylov subspace $\mathcal{K}_m$
- Build an orthonormal basis V_{m+1} of the Krylov subspace $\mathcal{K}_{m+1}$ get the Arnoldi-like relations $BW_m = V_{m+1}H_m$ and $BV_m = V_{m+1}H_mR_m^{-1}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1}V_my$ implies $r = r_0 - BV_my = V_{m+1}(\beta e_1 - H_mR_m^{-1}y)$ Solve the least-squares problem:

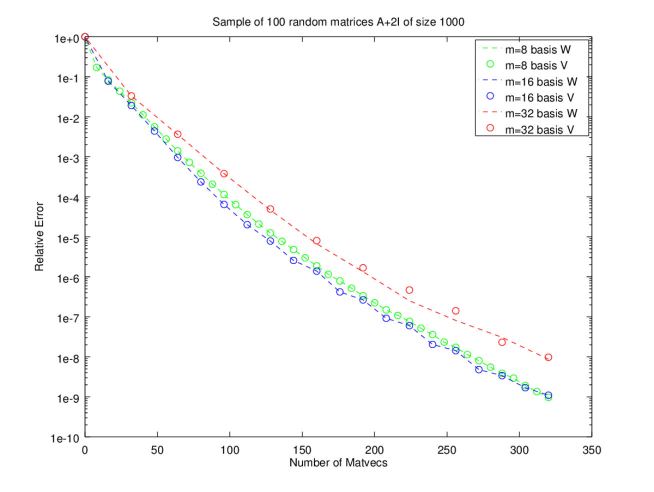
$$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_mR_m^{-1}y\|$$

- Alternative to minimize the residual in the Krylov subspace $x = x_0 + M^{-1}W_my$ implies $r = r_0 - BW_my = V_{m+1}(\beta e_1 - H_my)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_my\|$$

- Restart if not converged

Mean error curves with random matrices of size 1000



Fixed s step GMRES: SGMRES(m,s)

- Build a basis W_{sj} of the Krylov subspace $\mathcal{K}_{sj}$ for $1 \leq j \leq m/s$
- Build an orthonormal basis V_{sj+1} of the Krylov subspace $\mathcal{K}_{sj+1}$ get the Arnoldi-like relation $BW_{sj} = V_{sj+1}H_{sj}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1}W_{sj}y$ implies $r = r_0 - BW_{sj} = V_{sj+1}(\beta e_1 - H_{sj}y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{sj}} \|\beta e_1 - H_{sj}y\|$$

  Test convergence at each step j
- Restart if not converged at step m/s

$$x_0 = x_0 + M^{-1}W_m y_m$$

Fixed s step GMRES: SGMRES(m,s)

- Build a basis W_{sj} of the Krylov subspace $\mathcal{K}_{sj}$ for $1 \leq j \leq m/s$
- Build an orthonormal basis V_{sj+1} of the Krylov subspace $\mathcal{K}_{sj+1}$ get the Arnoldi-like relation $BW_{sj} = V_{sj+1}H_{sj}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1}W_{sj}y$ implies $r = r_0 - BW_{sj} = V_{sj+1}(\beta e_1 - H_{sj}y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{sj}} ||\beta e_1 - H_{sj}y||$$

  Test convergence at each step j

- Restart if not converged at step m/s

$$x_0 = x_0 + M^{-1}W_m y_m$$

Fixed s step GMRES: SGMRES(m,s)

- Build a basis W_{sj} of the Krylov subspace $\mathcal{K}_{sj}$ for $1 \leq j \leq m/s$
- Build an orthonormal basis V_{sj+1} of the Krylov subspace $\mathcal{K}_{sj+1}$ get the Arnoldi-like relation $BW_{sj} = V_{sj+1}H_{sj}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1}W_{sj}y$ implies $r = r_0 - BW_{sj} = V_{sj+1}(\beta e_1 - H_{sj}y)$
  Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{sj}} \|\beta e_1 - H_{sj}y\|$$

  Test convergence at each step j
- Restart if not converged at step m/s

$$x_0 = x_0 + M^{-1}W_m y_m$$

Fixed s step GMRES: SGMRES(m,s)

- Build a basis W_{sj} of the Krylov subspace $\mathcal{K}_{sj}$ for $1 \le j \le m/s$
- Build an orthonormal basis V_{sj+1} of the Krylov subspace $\mathcal{K}_{sj+1}$ get the Arnoldi-like relation $BW_{sj} = V_{sj+1}H_{sj}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1}W_{sj}y$ implies $r = r_0 - BW_{sj} = V_{sj+1}(\beta e_1 - H_{sj}y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{sj}} \|\beta e_1 - H_{sj}y\|$$

  Test convergence at each step j

- Restart if not converged at step m/s

$$x_0 = x_0 + M^{-1}W_m y_m$$

Step j of SGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$

Block j **of size** s, with $1 \leq j \leq m/s$

- First step: s matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{s(j-1)+1}, C_j = \left[u, Bu, \ldots, B^{s-1}u\right]$$

  Define the Krylov basis by

$$W_{sj} = \left[W_{s(j-1)}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{sj+1}S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$[v_1, BW_{sj}] = V_{sj+1}R_{sj+1}, BW_{sj} = V_{sj+1}H_{sj}$$

Block computation and orthogonalization in SGMRES(m,s)

Step j of SGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s, with $1 \le j \le m/s$

- First step: s matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{s(j-1)+1}, C_j = \left[u, Bu, \dots, B^{s-1}u\right]$$

  Define the Krylov basis by

$$W_{sj} = \left[W_{s(j-1)}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{sj+1}S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$[v_1, BW_{sj}] = V_{sj+1}R_{sj+1}, BW_{sj} = V_{sj+1}H_{sj}$$

Block computation and orthogonalization in SGMRES(m,s)

Step j of SGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s, with $1 \leq j \leq m/s$

- First step: s matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{s(j-1)+1}, C_j = \left[u, Bu, \ldots, B^{s-1}u\right]$$

  Define the Krylov basis by

$$W_{sj} = \left[W_{s(j-1)}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{sj+1}S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$[v_1, BW_{sj}] = V_{sj+1}R_{sj+1}, BW_{sj} = V_{sj+1}H_{sj}$$

Step j of SGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s, with $1 \leq j \leq m/s$

- First step: s matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{s(j-1)+1}, C_j = \left[u, Bu, \ldots, B^{s-1}u\right]$$

  Define the Krylov basis by

$$W_{sj} = \left[W_{s(j-1)}, C_j\right]$$

- Second step: orthogonalization

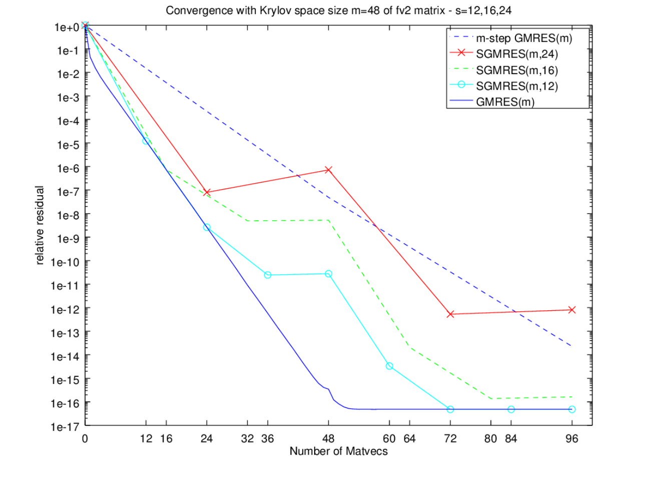
$$BC_j = V_{sj+1}S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$[v_1, BW_{sj}] = V_{sj+1}R_{sj+1}, BW_{sj} = V_{sj+1}H_{sj}$$

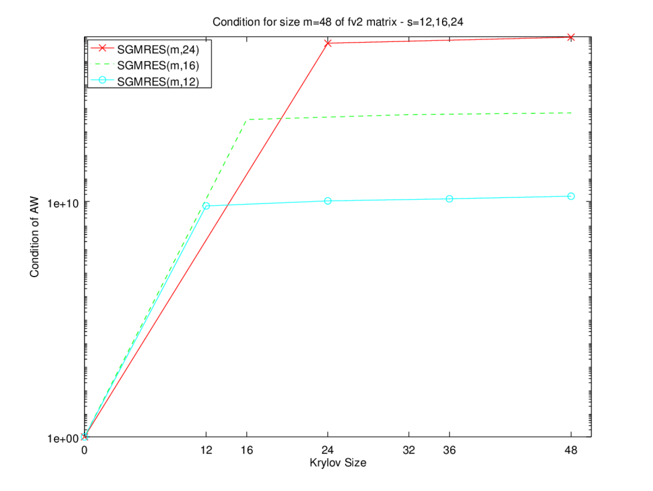
Symmetric matrix FV2 of size $n = 9801$ and nonzeros $nz = 87025$

Convergence curves with $m = 48$



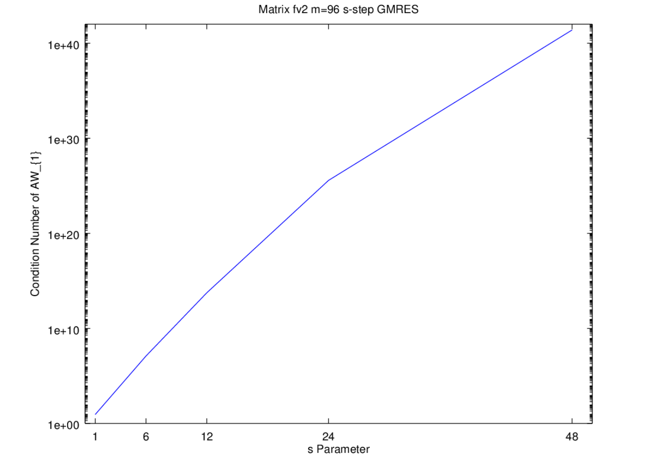
Symmetric matrix FV2 of size $n = 9801$ and nonzeros $nz = 87025$

Condition number of AW **with** $m = 48$



Symmetric matrix FV2 of size $n = 9801$ and nonzeros $nz = 87025$

Condition number of AW_1 with $m = 96$



Stability issues

Condition number of BC_j and BW_{sj}

$$W_{sj} = [C_1, C_2, \dots, C_j]$$

$$\kappa(BW_{sj}) \geq \max_{1 \leq k \leq j} \kappa(BC_k)$$

$$\kappa(BC_j) \geq cste\,|\lambda_1/\lambda_2|^{s-1}$$

Idea: variable block size s_j
increase gradually the size s_j and cap at s

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$ get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$

  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$

$$x_0 = x_0 + M^{-1} W_m y_m$$

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$
  get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace
  $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$
  Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$

  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$

$$x_0 = x_0 + M^{-1} W_m y_m$$

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$ get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$

  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$

$$x_0 = x_0 + M^{-1} W_m y_m$$

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$ get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1}H_{l_j}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1}W_{l_j}y$ implies $r = r_0 - BW_{l_j}y = V_{l_j+1}(\beta e_1 - H_{l_j}y)$
  Solve the least-squares problem:
  $$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j}y\|$$
  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$
  $$x_0 = x_0 + M^{-1}W_m y_m$$

Variable s step GMRES: VGMRES(m,s)

Variable block size s_j and Krylov size $l_1 = s_1, l_j = l_{j-1} + s_j, j \geq 2$

- Build a basis W_{l_j} of the Krylov subspace $\mathcal{K}_{l_j}$ for $1 \leq j \leq J$
- Build an orthonormal basis V_{l_j+1} of the Krylov subspace $\mathcal{K}_{l_j+1}$ get the Arnoldi-like relation $BW_{l_j} = V_{l_j+1} H_{l_j}$
- Minimize the residual in the Krylov subspace $x = x_0 + M^{-1} W_{l_j} y$ implies $r = r_0 - BW_{l_j} y = V_{l_j+1}(\beta e_1 - H_{l_j} y)$ Solve the least-squares problem:

$$\min_{y \in \mathbb{R}^{l_j}} \|\beta e_1 - H_{l_j} y\|$$

  Test convergence at each step j
- Restart if not converged at step J with $l_J = m$

$$x_0 = x_0 + M^{-1} W_m y_m$$

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$

Block j **of size** s_j, with $1 \leq j \leq J$

- First step: s_j matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{l_{j-1}+1}, C_j = \left[u, Bu, \ldots, B^{s_j-1}u\right]$$

  Define the Krylov basis by

$$W_{l_j} = \left[W_{l_{j-1}}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{l_j+1}S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$\left[v_1, BW_{l_j}\right] = V_{l_j+1}R_{l_j+1}, BW_{l_j} = V_{l_j+1}H_{l_j}$$

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s_j, with $1 \leq j \leq J$

- First step: s_j matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{l_{j-1}+1}, C_j = \left[u, Bu, \ldots, B^{s_j-1}u\right]$$

  Define the Krylov basis by

$$W_{l_j} = \left[W_{l_{j-1}}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{l_j+1}S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$\left[v_1, BW_{l_j}\right] = V_{l_j+1}R_{l_j+1}, BW_{l_j} = V_{l_j+1}H_{l_j}$$

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s_j, with $1 \leq j \leq J$

- First step: s_j matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{l_{j-1}+1}, C_j = \left[u, Bu, \ldots, B^{s_j-1}u\right]$$

  Define the Krylov basis by

$$W_{l_j} = \left[W_{l_{j-1}}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{l_j+1}S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$\left[v_1, BW_{l_j}\right] = V_{l_j+1}R_{l_j+1}, BW_{l_j} = V_{l_j+1}H_{l_j}$$

Block computation and orthogonalization with VGMRES(m,s)

Step j of VGMRES(m,s)

Initialization: $r_0 = b - Ax_0, \beta = \|r_0\|, v_1 = r_0/\beta, W_0 = \emptyset, V_1 = [v_1]$
Block j of size s_j, with $1 \leq j \leq J$

- First step: s_j matrix vector products
  Parallel preconditioning $t = M^{-1}u$ then parallel matrix-vector product At
  Compute the blocks C_j and BC_j where

$$u = v_{l_{j-1}+1}, C_j = \left[u, Bu, \ldots, B^{s_j-1}u\right]$$

  Define the Krylov basis by

$$W_{l_j} = \left[W_{l_{j-1}}, C_j\right]$$

- Second step: orthogonalization

$$BC_j = V_{l_j+1} S_j$$

  RODDEC [Sidje 1997, Erhel 1995] or TSQR [Demmel et al 2011]
  By induction, get the Arnoldi-like relation

$$\left[v_1, BW_{l_j}\right] = V_{l_j+1} R_{l_j+1}, BW_{l_j} = V_{l_j+1} H_{l_j}$$

Condition number of BC_j and BW_{l_j}

$$W_{l_j} = [C_1, C_2, \ldots, C_j]$$

$$\kappa(BW_{l_j}) \geq \max_{1 \leq k \leq j} \kappa(BC_k)$$

$$\kappa(BC_j) \geq cste\, |\lambda_1/\lambda_2|^{s_j - 1}$$

Objective: small condition numbers of the first blocks

Volume of communication related to the number of steps j
Objective: reduce the number of steps

Number of steps in SGMRES and VGMRES

SGMRES(m,s): m/s steps
VGMRES(m,s): J steps with $l_J = m$
Idea: Fibonacci sequence capped at s

$$J = J_1 + J_2, J_2 = O(m/s), J_1 = O(\log_\phi(s))$$

Sequences with $m = 48$ and $s = 16$

j	1	2	3	4	5	6	7
s_j	1	2	3	5	8	13	16
l_j	1	3	6	11	19	32	48

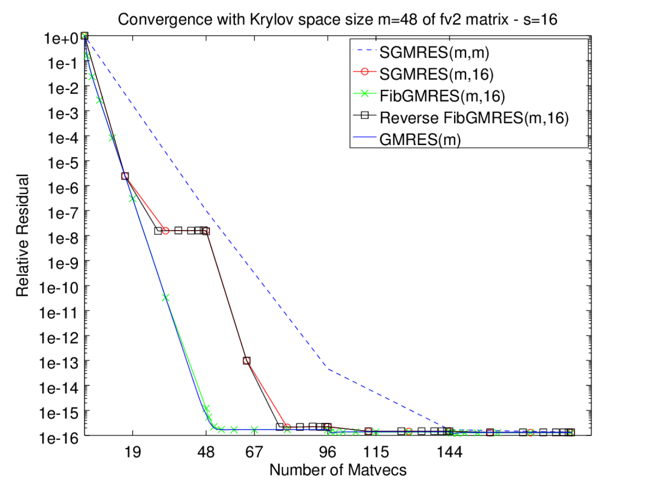
Variable increasing block size for $m = 48$ and $s = 16$.

j	1	2	3	4	5	6	7
s_j	16	13	8	5	3	2	1
l_j	16	29	37	42	45	47	48

Variable decreasing block size for $m = 48$ and $s = 16$.

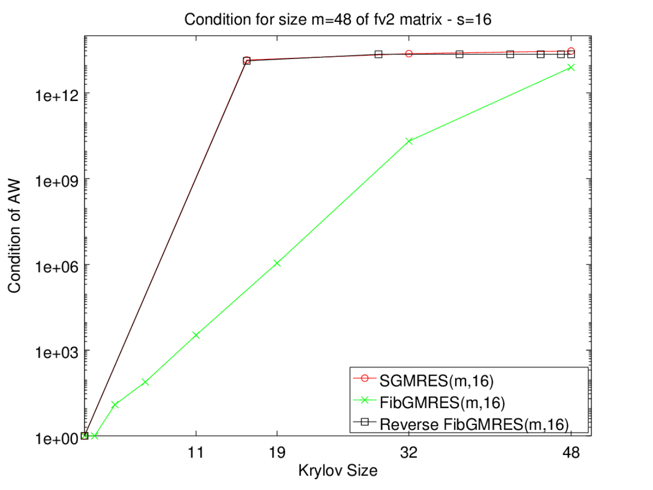
Symmetric matrix FV2 of size $n = 9801$ and nonzeros $nz = 87025$

Convergence curves with $m = 48$ and $s = 16$



Symmetric matrix FV2 of size $n = 9801$ and nonzeros $nz = 87025$

Condition numbers of AW with $m = 48$ and $s = 16$



Sequences of FibGMRES(m,s) with $m = 96$

j	1	2	3	4	5	6	7	8	9	10
s_j	1	2	3	5	8	13	16	16	16	16
l_j	1	3	6	11	19	32	48	64	80	96

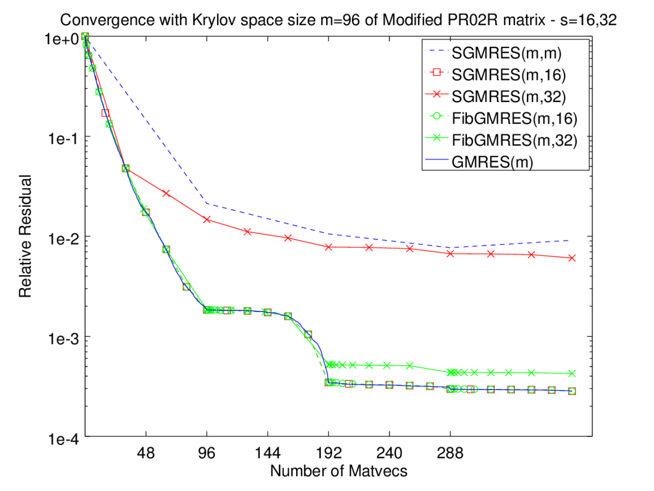
Variable increasing block size for $m = 96$ **and** $s = 16$

j	1	2	3	4	5	6	7	8	9
s_j	1	2	3	5	8	13	14	18	32
l_j	1	3	6	11	19	32	46	64	96

Variable increasing block size for $m = 96$ **and** $s = 32$

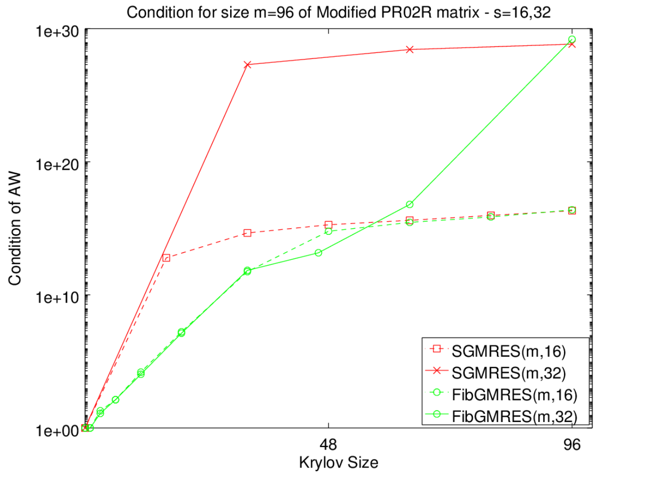
Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

Convergence curves with $m = 96$ and $s = 16$ or $s = 32$



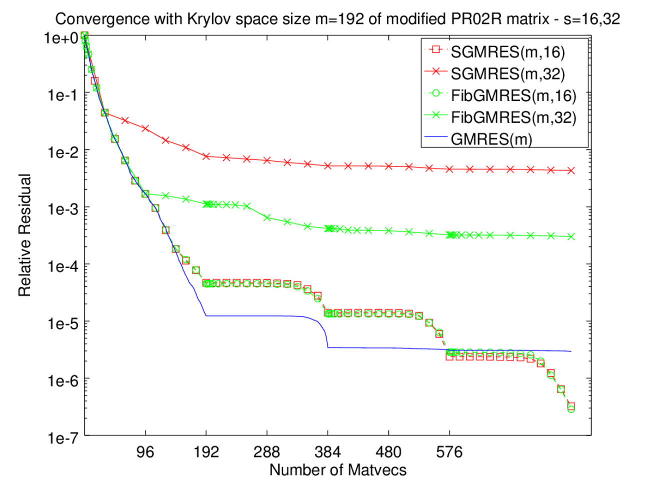
Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

Condition numbers of AW **with** $m = 96$ **and** $s = 16$ **or** $s = 32$



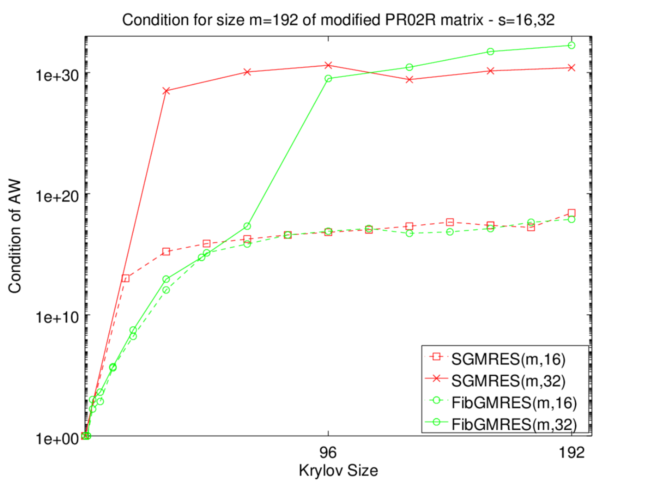
Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

Convergence curves with $m = 192$ and $s = 16$ or $s = 32$



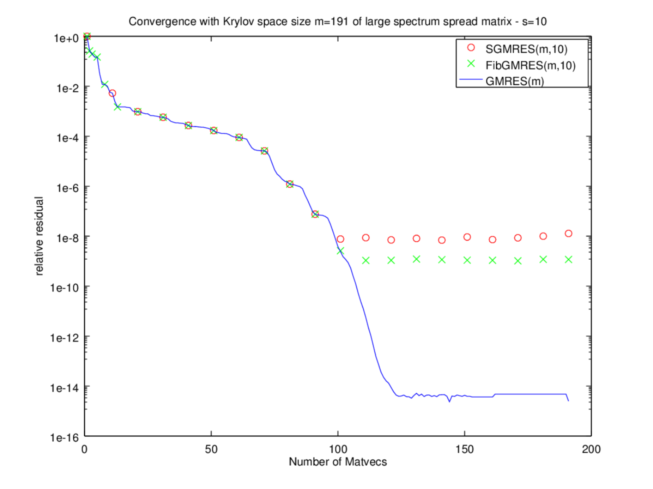
Nonsymmetric matrix (PR02R + 1000 I) with $n = 161070$ and $nz = 8185136$

Condition numbers of AW with $m = 192$ and $s = 16$ or $s = 32$



Large spectrum spread matrix

Convergence curves with $m = 192$ and $s = 10$



Conclusion

FibGMRES(m,s)

- A variable block size seems to converge faster than a fixed size
- Convergence seems related to the condition number of the largest block

Future work

- Adaptive variation of s
- Blocks computed via a Newton basis
- Schwarz preconditioning
- Coarse grid correction or augmented basis (deflation)
- Parallel computations