GMRES preconditioned by a perturbed LDL^T decomposition with static pivoting

M. Arioli, I. S. Duff, S. Gratton, and S. Pralet

http://www.numerical.rl.ac.uk/people/marioli/marioli.html

Outline

- Multifrontal
- Static pivoting
- GMRES and Flexible GMRES: a roundoff error analysis
- Numerical experiments

Linear system

We wish to solve large sparse systems

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

where $\mathbf{A} \in \mathbf{R}^{\mathrm{N}\times\mathrm{N}}$ is symmetric indefinite

Linear system

A particular and important case arises in saddle-point problems where the coefficient matrix is of the form

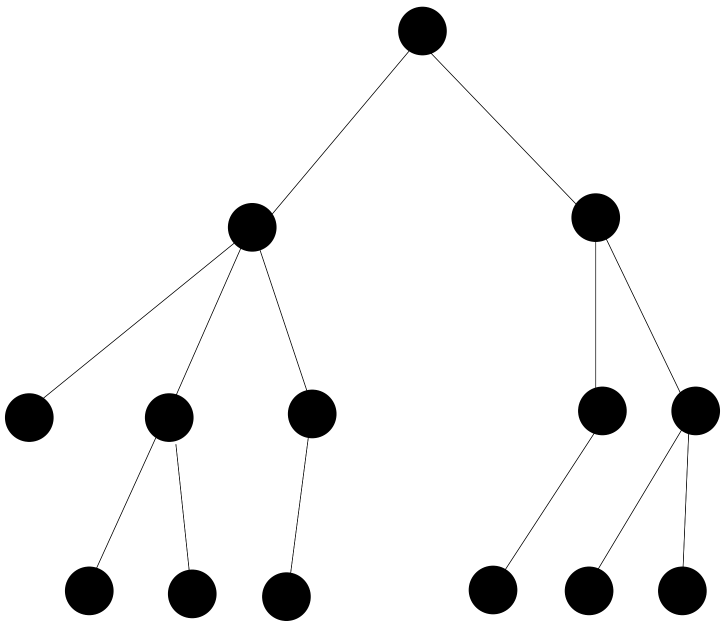
$$\begin{bmatrix} H & A \\ A^T & 0 \end{bmatrix}$$

Since we want accurate solutions, we would prefer to use a direct method of solution and our method of choice uses a multifrontal approach.

Science & Technology Facilities Council
Rutherford Appleton Laboratory

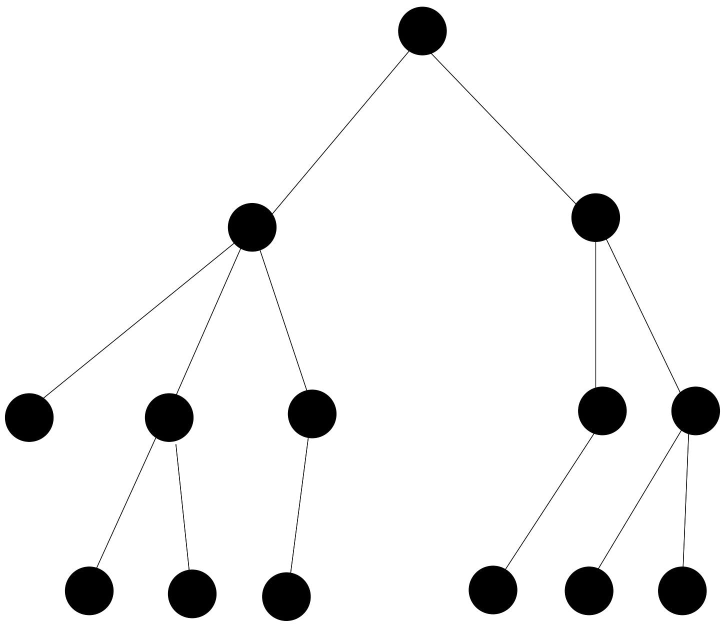
Multifrontal method

ASSEMBLY TREE



Multifrontal method

ASSEMBLY TREE

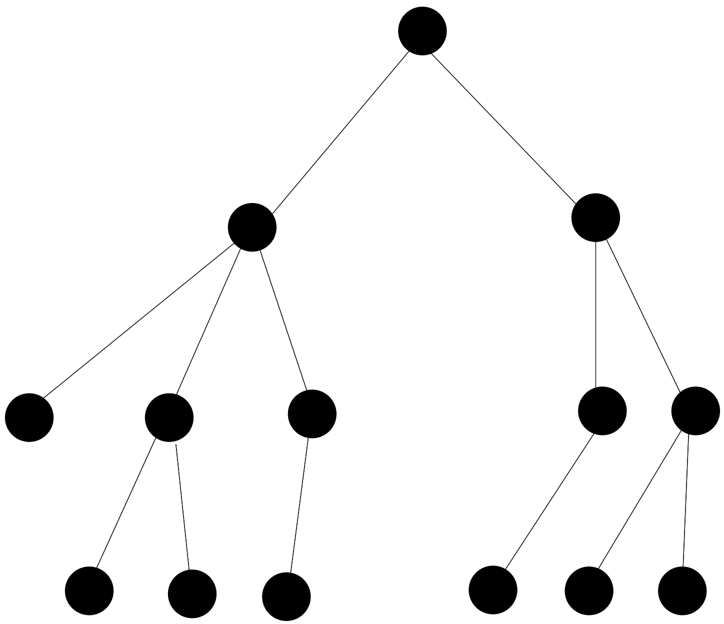


AT EACH NODE

$$\begin{bmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{bmatrix}$$

Multifrontal method

ASSEMBLY TREE



AT EACH NODE

$$\begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{12}^{T} & F_{22} \\ \hline \end{array}$$

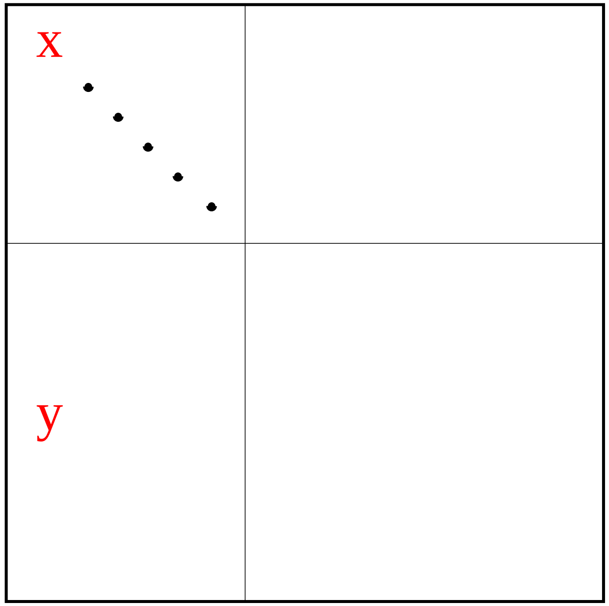
$$F_{22} \leftarrow F_{22} \; - \; F_{12}^{T} F_{11}^{-1} F_{12}$$

Multifrontal method

$$\begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{12}^{T} & F_{22} \\ \hline \end{array}$$

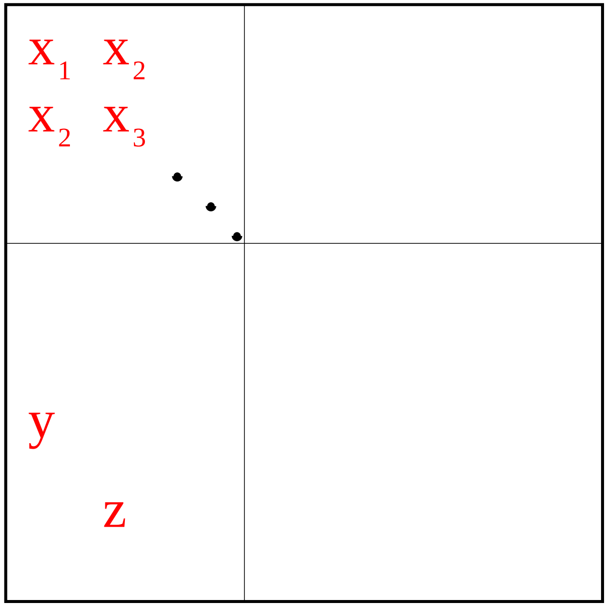
Pivot can only be chosen from F_{11} block since F_{22} is **NOT** fully summed.

Pivoting (1×1)



Choose x as 1×1 **pivot** if $|x| > u|y|$
where $|y|$ is the largest in column.

Pivoting (2×2)

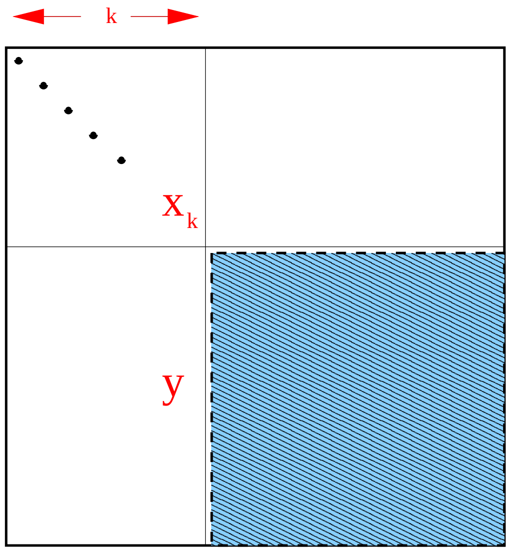


For the indefinite case, we can choose 2×2 **pivot** where we require

$$\left| \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix}^{-1} \right| \begin{bmatrix} |y| \\ |z| \end{bmatrix} \leq \begin{bmatrix} \frac{1}{u} \\ \frac{1}{u} \end{bmatrix}$$

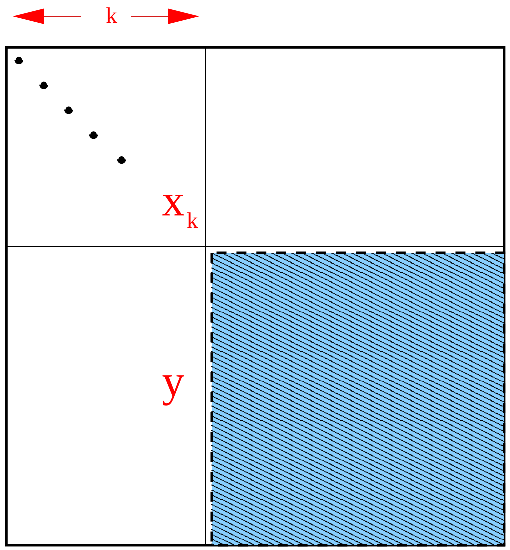
where again $|y|$ and $|z|$ are the largest in their columns.

Pivoting



If we assume that $k-1$ pivots are chosen but $|x_k| < u|y|$:

Pivoting

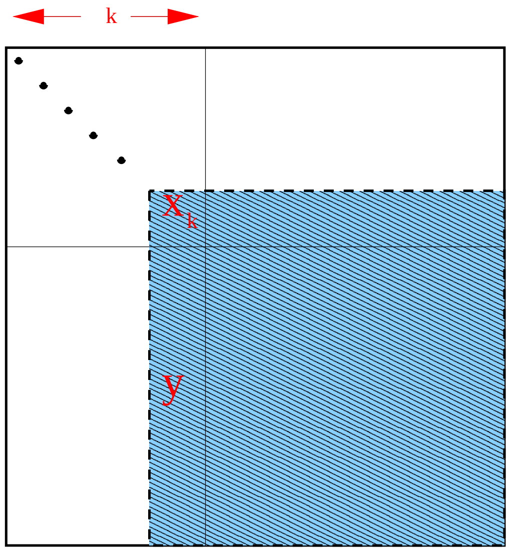


If we assume that $k-1$ pivots are chosen but $|x_k| < u|y|$:

- we can either take the **RISK** and use it or

Science & Technology Facilities Council
Rutherford Appleton Laboratory

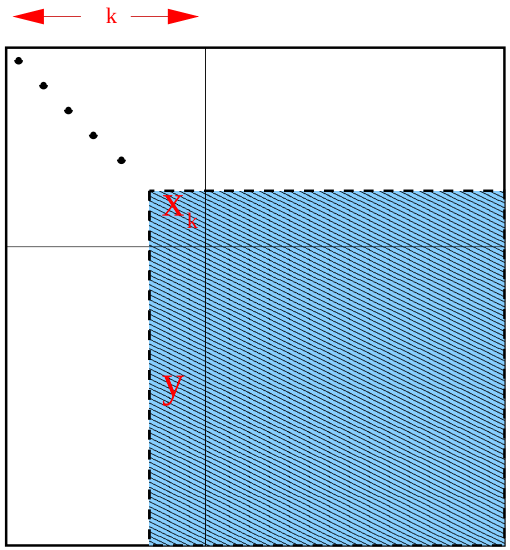
Pivoting



If we assume that $k-1$ pivots are chosen but $|x_k| < u|y|$:

- we can either take the **RISK** and use it or
- **DELAY** the pivot and then send to the parent a larger Schur complement.

Pivoting



If we assume that $k-1$ pivots are chosen but $|x_k| < u|y|$:

- we can either take the **RISK** and use it or
- **DELAY** the pivot and then send to the parent a larger Schur complement.

This can cause more work and storage

Static Pivoting

An **ALTERNATIVE** is to use **Static Pivoting**, by replacing x_k by

$$x_k + \tau$$

and CONTINUE.

Static Pivoting

An **ALTERNATIVE** is to use **Static Pivoting**, by replacing x_k by

$$x_k + \tau$$

and CONTINUE.

This is even more important in the case of parallel implementation where static data structures are often preferred

Static Pivoting

An **ALTERNATIVE** is to use **Static Pivoting**, by replacing x_k by

$$x_k + \tau$$

and CONTINUE.

This is even more important in the case of parallel implementation where static data structures are often preferred

Several codes use (or have an option for) this device:

- SuperLU (Demmel and Li)
- PARDISO (Gärtner and Schenk)
- MA57 (Duff and Pralet)

Static Pivoting

We thus have factorized

$$A + E = LDL^T = M$$

where $|E| \leq \tau I$

Static Pivoting

We thus have factorized

$$A + E = LDL^T = M$$

where $|E| \leq \tau I$

The three codes then have an **Iterative Refinement** option.
IR will converge if $\rho(M^{-1}E) < 1$

Static Pivoting

If $\rho(M^{-1}E) > 1$ then

PLAN A (Iterative Refinement Algortithm) fails!!!

Static Pivoting

If $\rho(M^{-1}E) > 1$ then

PLAN A (Iterative Refinement Algortithm) fails!!!

PLEASE DO NOT PANIC !

Static Pivoting

If $\rho(M^{-1}E) > 1$ then

PLAN A (Iterative Refinement Algortithm) fails!!!

PLEASE DO NOT PANIC !

We have Plan B

Static Pivoting

If $\rho(M^{-1}E) > 1$ then

PLAN A (Iterative Refinement Algortithm) fails!!!

PLEASE DO NOT PANIC !

We have Plan B

Flexible GMRES

Right preconditioned GMRES and Flexible GMRES

```
procedure [x] = right_Prec_GMRES(A,M,b)
```

procedure [x] = right_Prec_GMRES(A,M,b)
 $x_0 = M^{-1}b, r_0 = b - Ax_0$ and $\beta = ||r_0||$
 $v_1 = r_0/\beta$; k = 0;
 while $||r_k|| > \mu(||b|| + ||A||\ ||x_k||)$
  $k = k + 1$;
  $z_k = M^{-1}v_k; w = Az_k$;
  for $i = 1, \dots, k$ do
   $h_{i,k} = v_i^T w$;
   $w = w - h_{i,k}v_i$;
  end for;
  $h_{k+1,k} = ||w||$;
  $v_{k+1} = w/h_{k+1,k}$;
  $V_k = [v_1, \dots, v_k]$;
  $H_k = \{h_{i,j}\}_{1 \le i \le j+1; 1 \le j \le k}$;
  $y_k = \arg\min_y ||\beta e_1 - H_k y||$;
  $x_k = x_0 + M^{-1}V_k y_k$ and $r_k = b - Ax_k$;
 end while ;
end procedure.

procedure [x] =FGMRES(A, M_i,b)
 $x_0 = M_0^{-1}b, r_0 = b - Ax_0$ and $\beta = ||r_0||$
 $v_1 = r_0/\beta$; k = 0;
 while $||r_k|| > \mu(||b|| + ||A||\ ||x_k||)$
  $k = k + 1$;
  $z_k = M_k^{-1}v_k; w = Az_k$;
  for $i = 1, \dots, k$ do
   $h_{i,k} = v_i^T w$;
   $w = w - h_{i,k}v_i$;
  end for;
  $h_{k+1,k} = ||w||$;
  $v_{k+1} = w/h_{k+1,k}$;
  $Z_k = [z_1, \dots, z_k]$; $V_k = [v_1, \dots, v_k]$;
  $H_k = \{h_{i,j}\}_{1 \le i \le j+1; 1 \le j \le k}$;
  $y_k = \arg\min_y ||\beta e_1 - H_k y||$;
  $x_k = x_0 + Z_k y_k$ and $r_k = b - Ax_k$;
 end while ;
end procedure.

Roundoff error 1

The computed $\hat{L}$ and $\hat{D}$ in floating-point arithmetic satisfy

$$\begin{cases} A + \delta A + \tau E = M \\ ||\delta A|| \leq c(n)\varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \\ ||E|| \leq 1. \end{cases}$$

The perturbation δA must have a norm smaller than τ, in order to not dominate the global error.

Roundoff error 1

The computed $\hat{L}$ and $\hat{D}$ in floating-point arithmetic satisfy

$$\begin{cases} A + \delta A + \tau E = M \\ ||\delta A|| \leq c(n)\varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \\ ||E|| \leq 1. \end{cases}$$

The perturbation δA must have a norm smaller than τ, in order to not dominate the global error.

A sufficient condition for this is $\boxed{n \, \varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \leq \tau}$

Roundoff error 1

The computed $\hat{L}$ and $\hat{D}$ in floating-point arithmetic satisfy

$$\begin{cases} A + \delta A + \tau E = M \\ ||\delta A|| \leq c(n)\varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \\ ||E|| \leq 1. \end{cases}$$

The perturbation δA must have a norm smaller than τ, in order to not dominate the global error.

A sufficient condition for this is $\boxed{n \, \varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \leq \tau}$

$|| \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \approx \frac{n}{\tau} \Longrightarrow \varepsilon \leq \frac{\tau^2}{n^2}$

Roundoff error 1

The computed $\hat{L}$ and $\hat{D}$ in floating-point arithmetic satisfy

$$\begin{cases} A + \delta A + \tau E = M \\ ||\delta A|| \leq c(n)\varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \\ ||E|| \leq 1. \end{cases}$$

The perturbation δA must have a norm smaller than τ, in order to not dominate the global error.

A sufficient condition for this is $\boxed{n \, \varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \leq \tau}$

$|| \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || \approx \frac{n}{\tau} \Longrightarrow \varepsilon \leq \frac{\tau^2}{n^2}$

Moreover, we assume that $\boxed{\max\{||M^{-1}||, ||\bar{Z}_k||\} \leq \frac{\tilde{c}}{\tau}}$.

Roundoff error FGMRES

Theorem 1.

$$\boxed{\sigma_{\mathit{min}}(\bar{H}_k) > c_7(k,1)\varepsilon||\bar{H}_k|| + \mathcal{O}(\varepsilon^2) \quad \forall k,}$$

$$\boxed{|\bar{s}_k| < 1 - \varepsilon,\ \forall k,}$$

(where $\bar{s}_k$ are the sines computed during the Givens algorithm)

and

$$\boxed{2.12(n+1)\varepsilon < 0.01 \text{ and } 18.53\varepsilon n^{\frac{3}{2}}\kappa(C^{(k)}) < 0.1\ \forall k}$$

$$\exists \hat{k}, \qquad \hat{k} \leq n$$

such that, $\forall k \geq \hat{k}$*, we have*

$$||b - A\bar{x}_k|| \leq c_1(n,k)\varepsilon\Big(||b|| + ||A||\,||\bar{x}_0|| + ||A||\,||\bar{Z}_k||\,||\bar{y}_k||\Big) + \mathcal{O}(\varepsilon^2).$$

Roundoff error FGMRES

Moreover, if $M_i = M, \forall i$*,*

$$\boxed{\rho = 1.3\,||\hat{W}_k|| + c_2(k,1)\varepsilon||M||\,||\bar{Z}_k|| < 1 \quad \forall k < \hat{k},}$$

where

$$\boxed{\hat{W}_k = [M\bar{z}_1 - \bar{v}_1, \ldots, M\bar{z}_k - \bar{v}_k]\,,}$$

we have:

$$||b - A\bar{x}_k|| \leq c(n,k)\gamma\varepsilon(||b|| + ||A||\,||\bar{x}_0|| + ||A||\,||\bar{Z}_k||\,||M(\bar{x}_k - \bar{x}_0)||) + \mathcal{O}(\varepsilon^2)$$

$$\gamma = \frac{1.3}{1-\rho}.$$

Roundoff error FGMRES

Theorem 2

Under the Hypotheses of Theorem 1, and

$$\boxed{\mathbf{c}(n)\varepsilon \| \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, \| < \tau}$$

$$\boxed{c(n,k)\gamma\varepsilon \|A\| \, \|\bar{Z}_k\| < 1 \quad \forall k < \hat{k}}$$

$$\boxed{\max\{\|M^{-1}\|, \|\bar{Z}_k\|\} \le \frac{\tilde{c}}{\tau}}$$

we have

Roundoff error FGMRES

Theorem 2

Under the Hypotheses of Theorem 1, and

$$\boxed{\mathbf{c}(n)\varepsilon || \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, || < \tau}$$

$$\boxed{c(n,k)\gamma\varepsilon ||A|| \, ||\bar{Z}_k|| < 1 \quad \forall k < \hat{k}}$$

$$\boxed{\max\{||M^{-1}||, ||\bar{Z}_k||\} \leq \frac{\tilde{c}}{\tau}}$$

we have

$$||b - A\bar{x}_k|| \leq 2\mu\varepsilon(||b|| + ||A|| \, (||\bar{x}_0|| + ||\bar{x}_k||)) + \mathcal{O}(\varepsilon^2).$$

$$\mu = \frac{c(n,k)}{1 - c(n,k)\varepsilon ||A|| \, ||\bar{Z}_k||}$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Roundoff error right preconditioned GMRES

Theorem 3
We assume of applying Iterative Refinement for solving $M(\bar{x}_k - \bar{x}_0) = \bar{V}_k \bar{y}_k$ *at last step.*
Under the Hypotheses of Theorem 1 and $\boxed{c(n)\varepsilon\,\kappa(M) < 1}$

$$\exists \hat{k}, \qquad \hat{k} \le n$$

such that, $\forall k \ge \hat{k}$*, we have*

$$\begin{aligned} \|b - A\bar{x}_k\| \le \quad & c_1(n,k)\varepsilon \Big\{ \|b\| + \|A\|\,\|\bar{x}_0\| + \|A\|\,\|\bar{Z}_k\|\,\|M(\bar{x}_k - \bar{x}_0)\| + \\ & \|AM^{-1}\|\,\|M\|\,\|\bar{x}_k - \bar{x}_0\| + \\ & \|AM^{-1}\|\,\|\,|\hat{L}|\,|\hat{D}|\,|\hat{L}^T|\,\|\,\|M(\bar{x}_k - \bar{x}_0)\| \Big\} + \mathcal{O}(\varepsilon^2). \end{aligned}$$

Test Problems

	n	nnz	Description
CONT_201	80595	239596	KKT matrix Convex QP (M2)
CONT_300	180895	562496	KKT matrix Convex QP (M2)
TUMA_1	22967	76199	Mixed-Hybrid finite-element

Test problems

MA57 tests

	n	nnz(L)+nnz(D)	Factorization time
CONT_201	80595	9106766	9.0 sec
CONT_300	180895	22535492	28.8 sec

MA57 without static pivot

MA57 tests

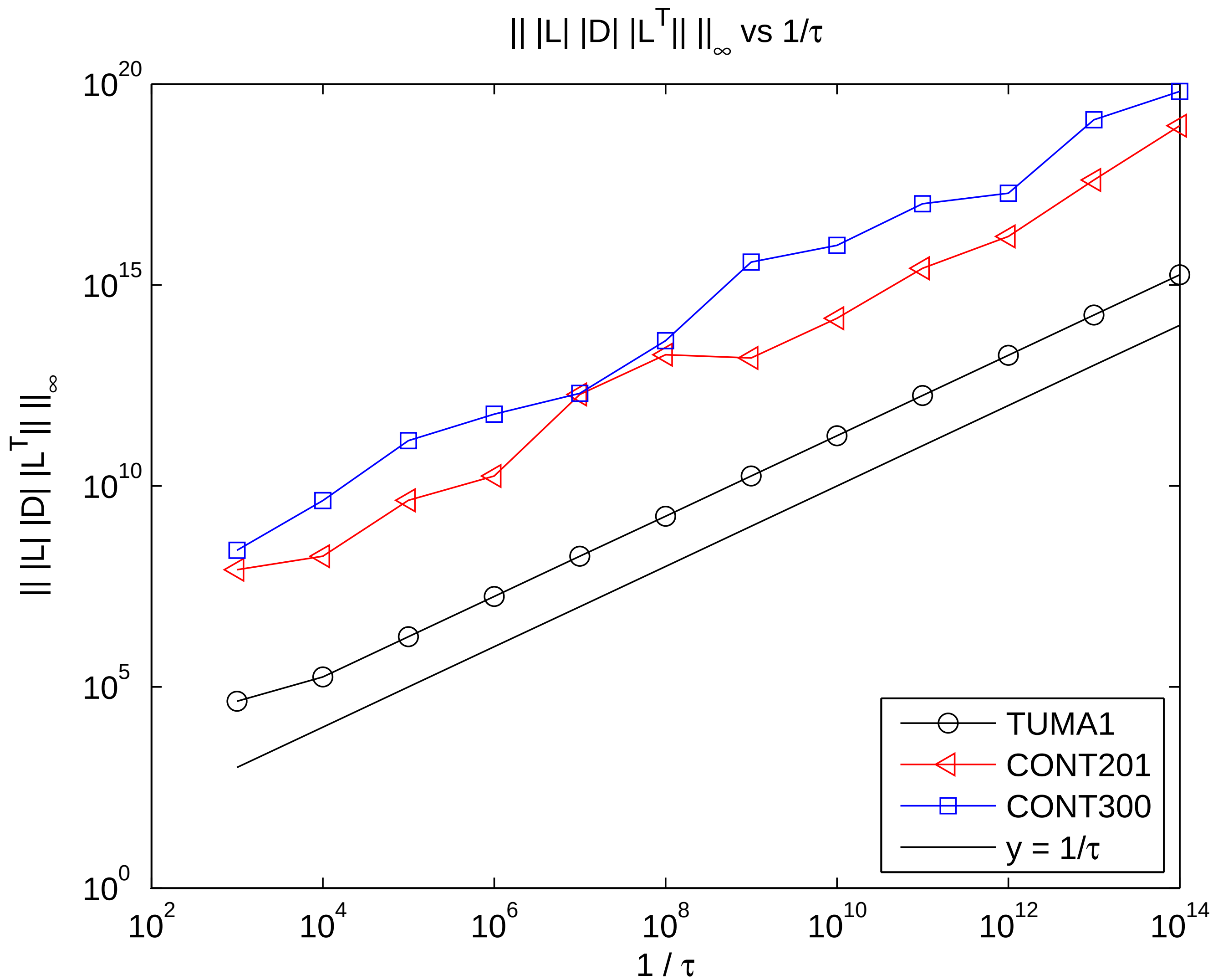
	n	nnz(L)+nnz(D)	Factorization time
CONT_201	80595	9106766	9.0 sec
CONT_300	180895	22535492	28.8 sec

MA57 without static pivot

	nnz(L)+nnz(D)+ FGMRES (#it)	Factorization time	# static pivots
CONT_201	5563735 (6)	3.1 sec	27867
CONT_300	12752337 (8)	8.9 sec	60585

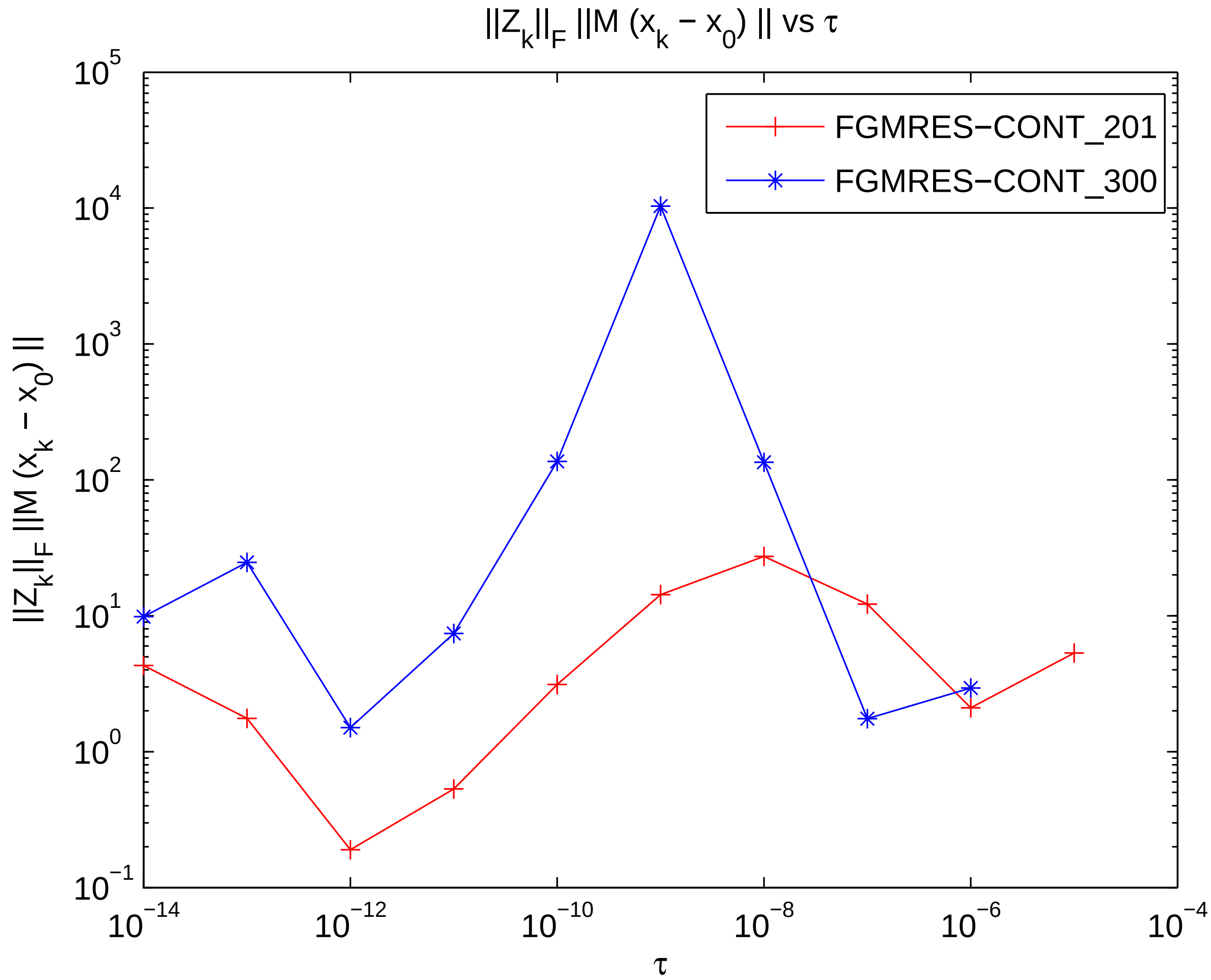
MA57 with static pivot $\tau = 10^{-8}$

$\| \, |\hat{L}| \, |\hat{D}| \, |\hat{L}^T| \, \|$ vs $1/\tau$

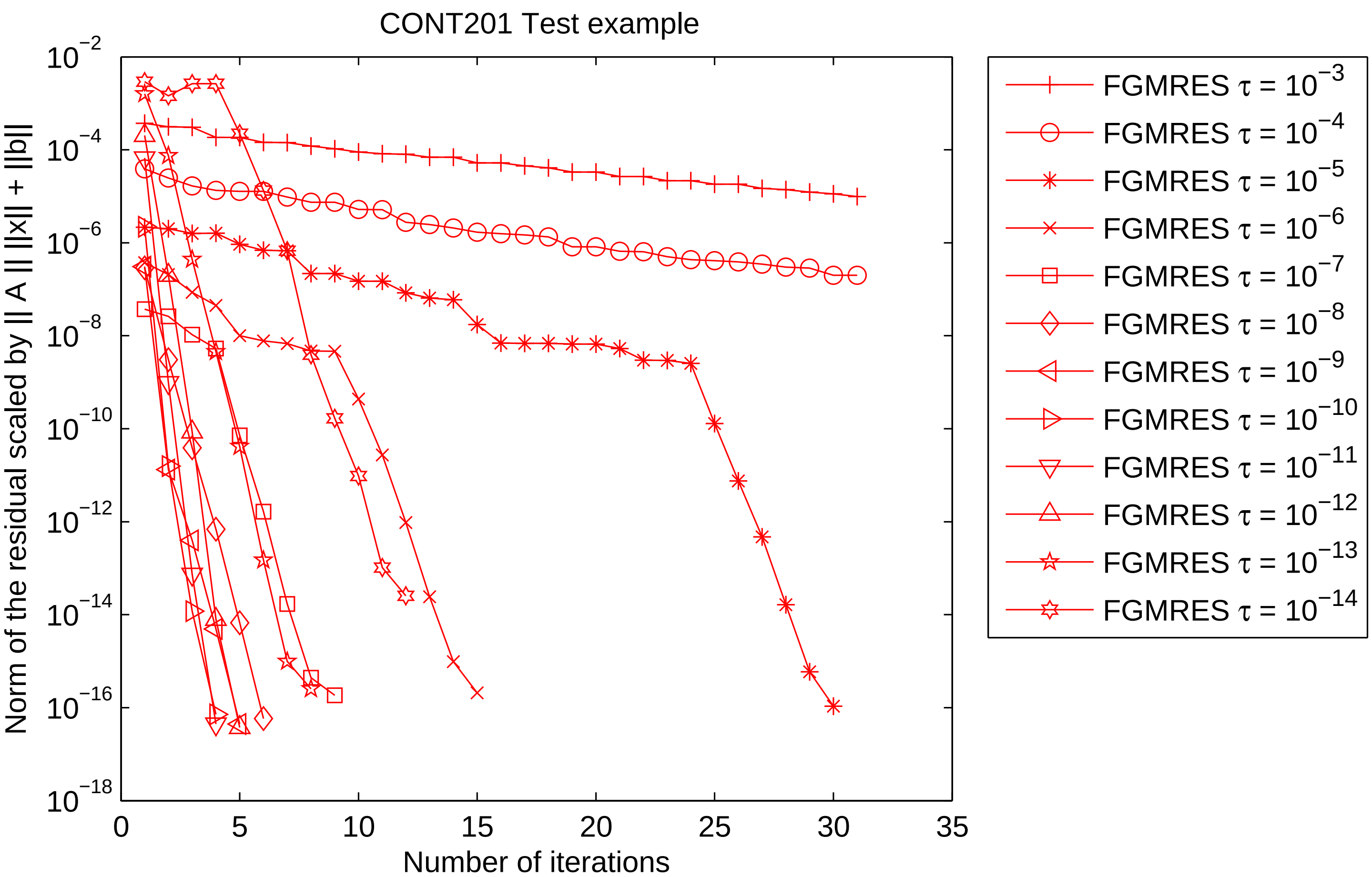


Science & Technology Facilities Council
Rutherford Appleton Laboratory

$||\bar{Z}_k||_F \, ||M(x_k - x_0)||$ vs τ

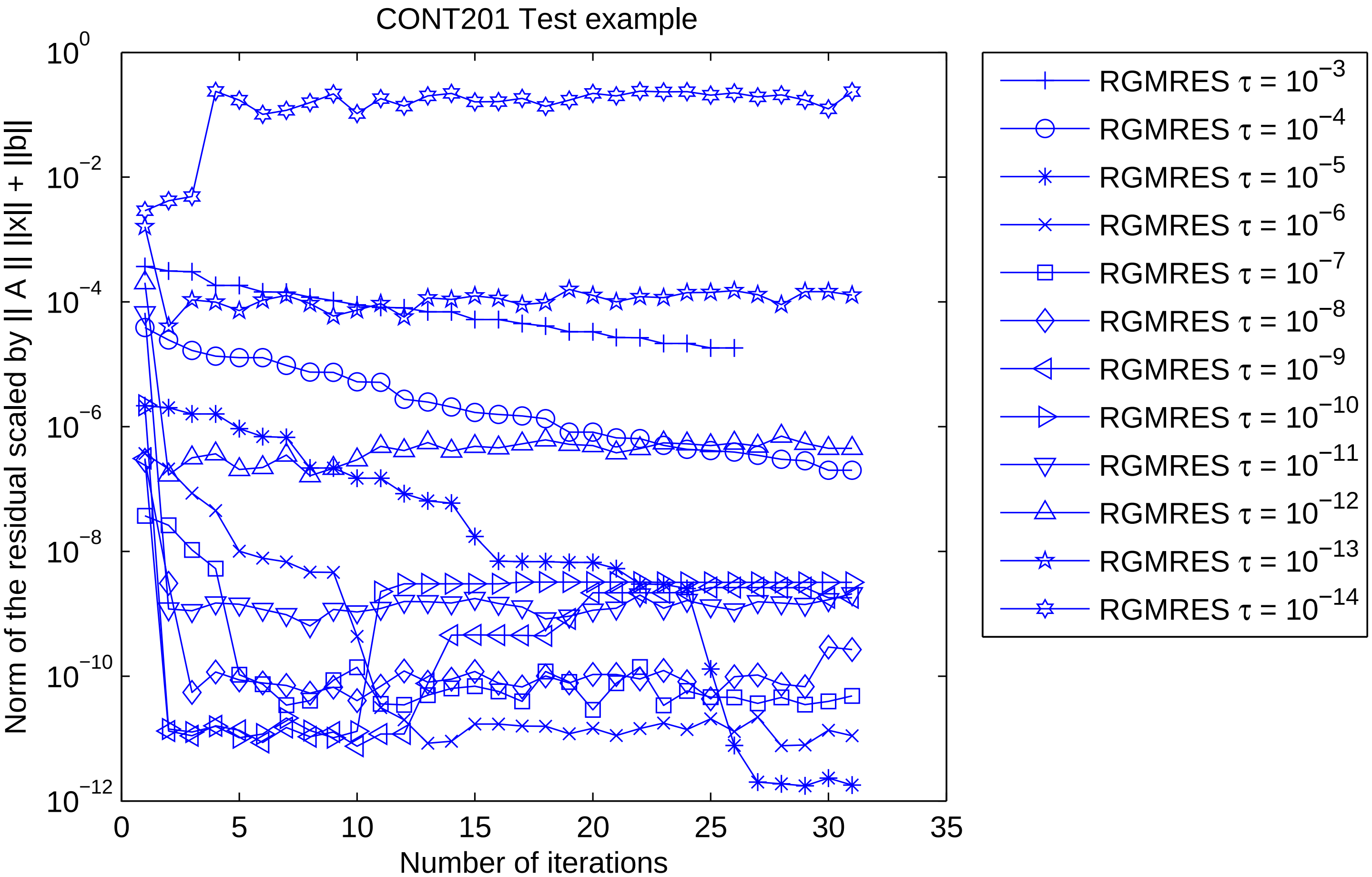


Numerical experiments



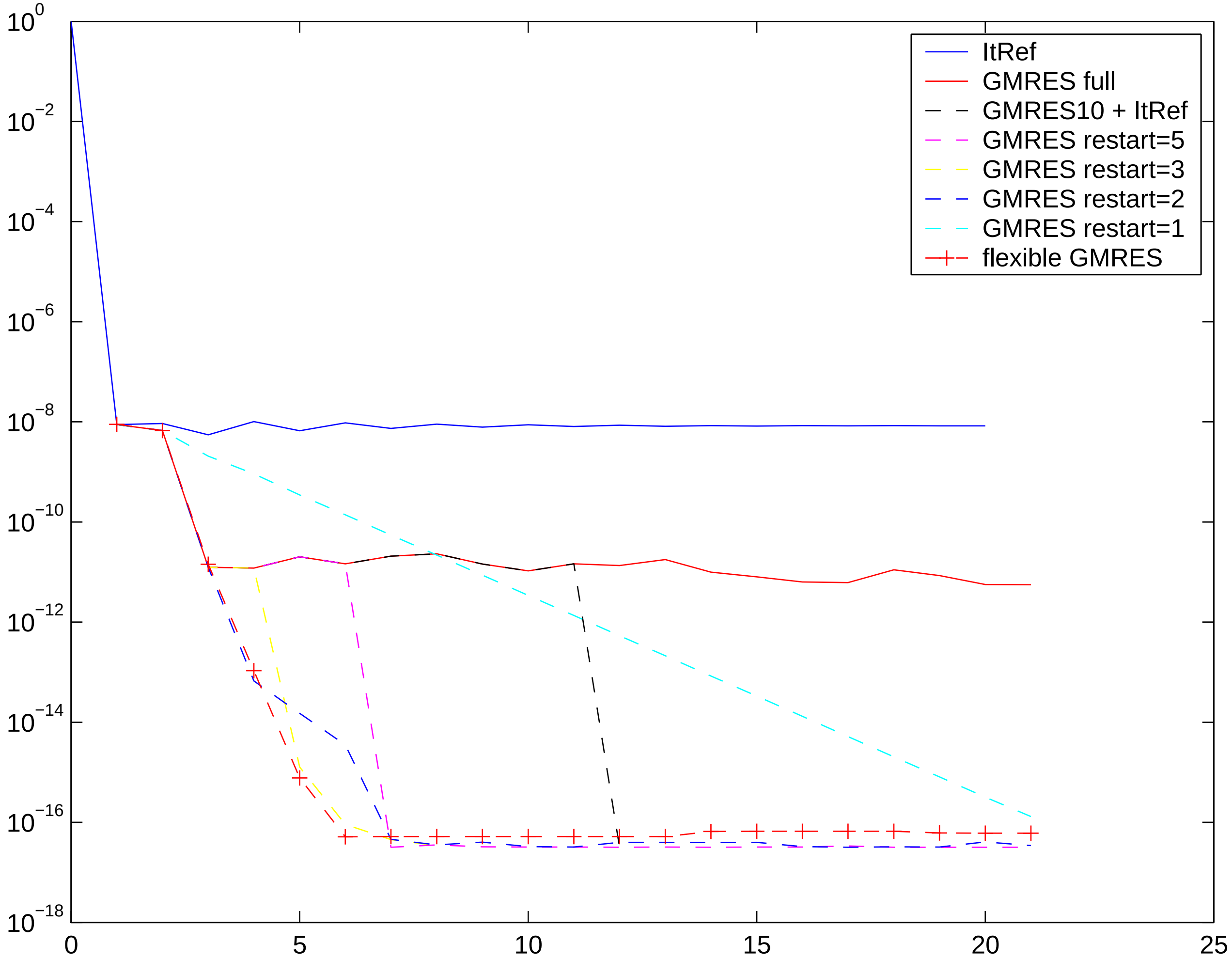
FGMRES on CONT-201 test example

Numerical experiments



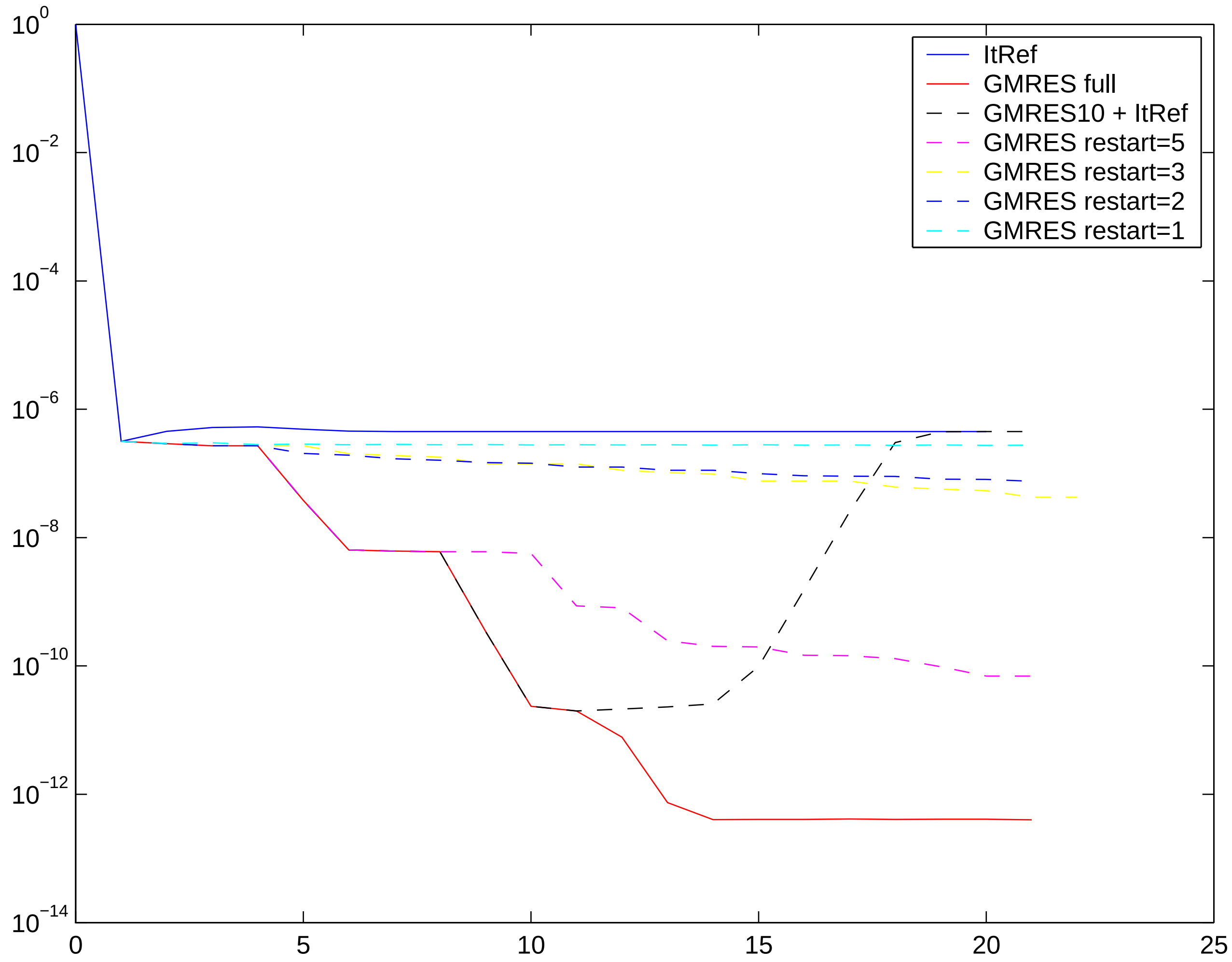
GMRES on CONT-201 test example

Numerical experiments



Restarted GMRES vs. FGMRES on CONT-201 test example: $\tau = 10^{-8}$

Numerical experiments



Restarted GMRES on CONT-201 test example: $\tau = 10^{-6}$

Summary

- IR with static pivoting is very sensitive to τ and not robust

Summary

- IR with static pivoting is very sensitive to τ and not robust
- GMRES is also sensitive and not robust

Summary

- IR with static pivoting is very sensitive to τ and not robust
- GMRES is also sensitive and not robust
- FGMRES is robust and less sensitive (see roundoff analysis)

Summary

- IR with static pivoting is very sensitive to τ and not robust
- GMRES is also sensitive and not robust
- FGMRES is robust and less sensitive (see roundoff analysis)
- Gains from restarting. Makes GMRES more robust, saves storage in FGMRES (but not really needed)

Summary

- IR with static pivoting is very sensitive to τ and not robust
- GMRES is also sensitive and not robust
- FGMRES is robust and less sensitive (see roundoff analysis)
- Gains from restarting. Makes GMRES more robust, saves storage in FGMRES (but not really needed)
- Understanding of why $\tau \approx \sqrt{\varepsilon}$ is best.

Summary

- IR with static pivoting is very sensitive to τ and not robust
- GMRES is also sensitive and not robust
- FGMRES is robust and less sensitive (see roundoff analysis)
- Gains from restarting. Makes GMRES more robust, saves storage in FGMRES (but not really needed)
- Understanding of why $\tau \approx \sqrt{\varepsilon}$ is best.
- PLAN B is working