PDE-constrained optimisation: why is it so challenging and some methods to overcome these challenges

Sue Thorne (née Dollar)

STFC Rutherford Appleton Laboratory

In collaboration with

- Nick Gould (STFC Rutherford Appleton Laboratory)
- Tyrone Rees (Oxford)
- Andy Wathen (Oxford)

PDE-constrained optimization

- Given f and boundary condition g, calculate u, where

$$\mathcal{L}u = f, \quad \alpha_1 u + \alpha_2 \frac{\partial u}{\partial n} = g \text{ on } \partial\Omega$$

on some domain Ω

PDE-constrained optimization

- Given f and boundary condition g, calculate u, where

$$\mathcal{L}u = f, \quad \alpha_1 u + \alpha_2 \frac{\partial u}{\partial n} = g \text{ on } \partial\Omega$$

  on some domain Ω
- Suppose given g and an approximation $\hat{u}$ to u on some domain $\hat{\Omega} \subset \Omega$. Want to calculate f such that $u \approx \hat{u}$: distributed control

PDE-constrained optimization

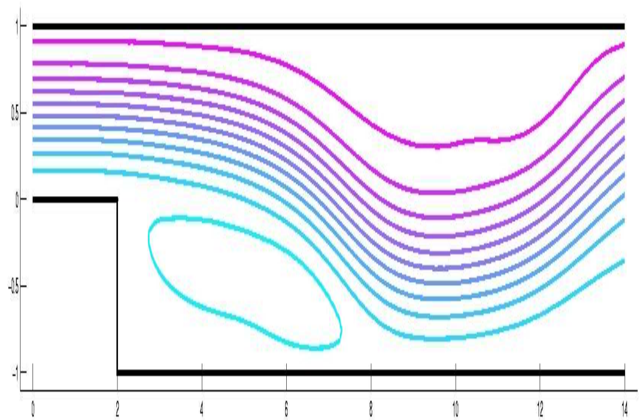
- Given f and boundary condition g, calculate u, where

$$\mathcal{L}u = f, \quad \alpha_1 u + \alpha_2 \frac{\partial u}{\partial n} = g \text{ on } \partial\Omega$$

  on some domain Ω

- Suppose given g and an approximation $\hat{u}$ to u on some domain $\hat{\Omega} \subset \Omega$. Want to calculate f such that $u \approx \hat{u}$: distributed control
- Suppose given f and an approximation $\hat{u}$ to u on some domain $\hat{\Omega} \subset \Omega$. Want to calculate g such that $u \approx \hat{u}$: boundary control

Different target temperatures



Reduce recirculation

Distributed control

$$\min_{\mathtt{u},\mathtt{f}} \frac{1}{2} \|\omega(x)\,(\mathtt{u} - \widehat{\mathtt{u}})\|_2^2 + \beta\,\|\mathtt{f}\|_2^2$$

subject to

$$\begin{aligned} \mathcal{L}\mathtt{u} &= \mathtt{f} \text{ in } \Omega \\ \mathtt{u} &= \mathtt{g} \text{ on } \delta\Omega \end{aligned}$$

Here

$$\omega(x) = \begin{cases} 1 & x \in \hat{\Omega} \\ 0 & \text{otherwise} \end{cases}$$

Distributed control

Discretize:

$$\mathtt{u_h} = \sum u_j \phi_j, \quad \mathtt{f_h} = \sum f_j \phi_j$$

$$\min_{\mathtt{u_h},\mathtt{f_h}} \frac{1}{2} \left\|\omega(x)\left(\mathtt{u_h} - \widehat{u}\right)\right\|_2^2 + \beta \left\|\mathtt{f_h}\right\|_2^2$$

subject to

$$\begin{aligned} \mathcal{L}\mathtt{u_h} &= \mathtt{f_h} \text{ in } \Omega \\ \mathtt{u_h} &= \mathsf{g} \text{ on } \delta\Omega \end{aligned}$$

Let $\mathcal{L} = -\nabla^2$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Distributed control

$$
\begin{aligned}
\|\omega(x)\,(\mathtt{u_h} - \widehat{u})\|_2^2 &= \int_\Omega \omega(x)\,(\mathtt{u_h} - \widehat{u})^2 \\
&= \sum_i \sum_j u_i u_j \int_\Omega \omega_i \omega_j \phi_i \phi_j - 2 \sum_j u_j \int_\Omega \omega_j \phi_j \widehat{u} + \int_{\hat{\Omega}} \widehat{u}^2 \\
&= u^T \bar{M} u - u^T b + c \\
\|\mathtt{f_h}\|_2^2 &= f^T M f \\
K u &= M f
\end{aligned}
$$

where M is the mass matrix, K is the stiffness matrix, $\bar{M} = WMW$ and $W = \mathsf{diag}(\omega_i)$

Distributed control

$$\min_{u,f} \frac{1}{2} u^T \bar{M} u - u^T b + c + \beta f^T M f$$

subject to

$$Ku - Mf \quad = \quad d$$

Distributed control

$$\min_{u,f} \frac{1}{2} u^T \bar{M} u - u^T b + c + \beta f^T M f + l^T \left(Ku - Mf - d \right)$$

Optimality conditions:

$$\begin{bmatrix} \beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \begin{bmatrix} f \\ u \\ l \end{bmatrix} = \begin{bmatrix} 0 \\ b \\ d \end{bmatrix}$$

Direct vs Iterative Methods

	Direct Methods		Iterative Methods
$\surd$	Black box	$\surd$	Large problems
$\surd$	Robust (large $\kappa(A)$?)	$\surd$	Preconditioning – convergence
$\times$	Memory with large problems?	$\times$	Iterative method?
		$\times$	Preconditioner?

Definition: let $\kappa(\mathcal{A}) = ||\mathcal{A}||_2 ||\mathcal{A}^{-1}||_2$ be the condition number of $\mathcal{A}$

Spectral properties of linear system

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

If A is symmetric and positive definite, then $\lambda(A) \in I^- \cup I^+$, where

$$\begin{aligned} I^- &= \left[\frac{1}{2}\left(\lambda_{\min}(A) - \sqrt{\lambda_{\min}^2(A) + 4\|B\|^2}\right), \frac{1}{2}\left(\|A\| - \sqrt{\|A\|^2 + 4\sigma_{\min}^2(B)}\right)\right], \\ I^+ &= \left[\lambda_{\min}(A), \frac{1}{2}\left(\|A\| + \sqrt{\|A\|^2 + 4\|B\|^2}\right)\right], \end{aligned}$$

[Rusten and Winther 1992]

Spectral properties of linear system

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

If A is symmetric and positive semi-definite, then $\lambda(H) \in I^- \cup I^+$, where

$$\begin{aligned} I^- &= \left[\frac{1}{2}\left(\lambda_{\min}(A) - \sqrt{\lambda_{\min}^2(A) + 4\|B\|^2}\right), \frac{1}{2}\left(\|A\| - \sqrt{\|A\|^2 + 4\sigma_{\min}^2(B)}\right)\right], \\ I^+ &= \left[l(A,B), \frac{1}{2}\left(\|A\| + \sqrt{\|A\|^2 + 4\|B\|^2}\right)\right], \end{aligned}$$

$l(A,B)$ defined in Dollar 2009 (revised)

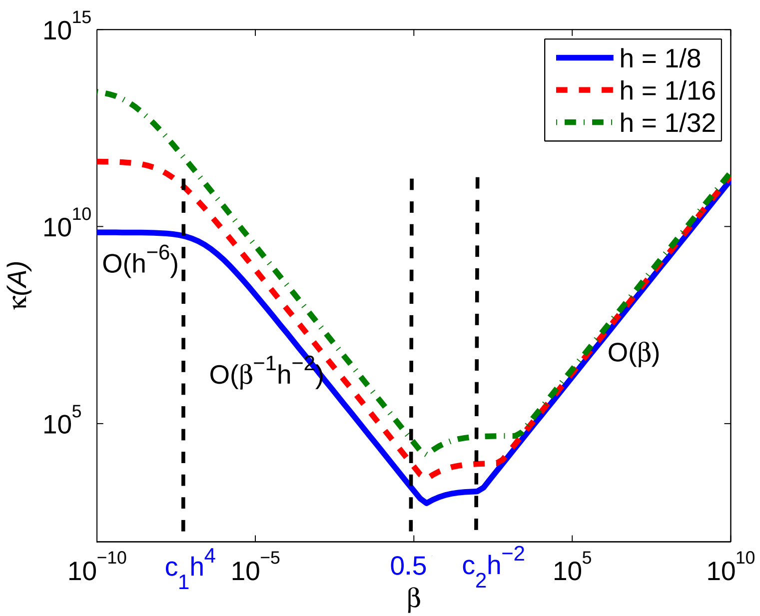
If A scaled so that $\lambda_{\max}(A) \le 1$, then $l(A,B) = Z^T AZ$, where $BZ = 0$ (Simoncini talk 2008)

Spectral properties of linear system

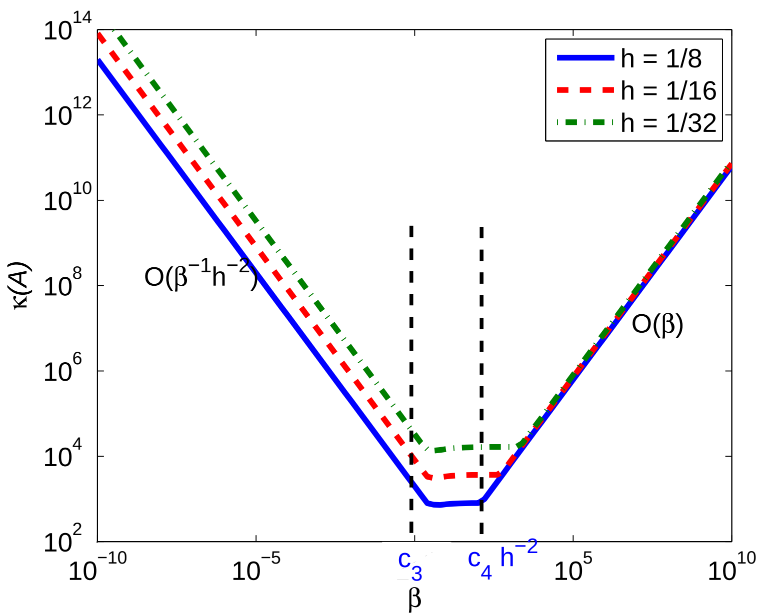
$$
\begin{aligned}
H_\beta &= \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \\
&= H_0 + \begin{bmatrix} 2\beta M & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}
$$

Spectral properties of linear system

$\hat{\Omega} = \Omega$



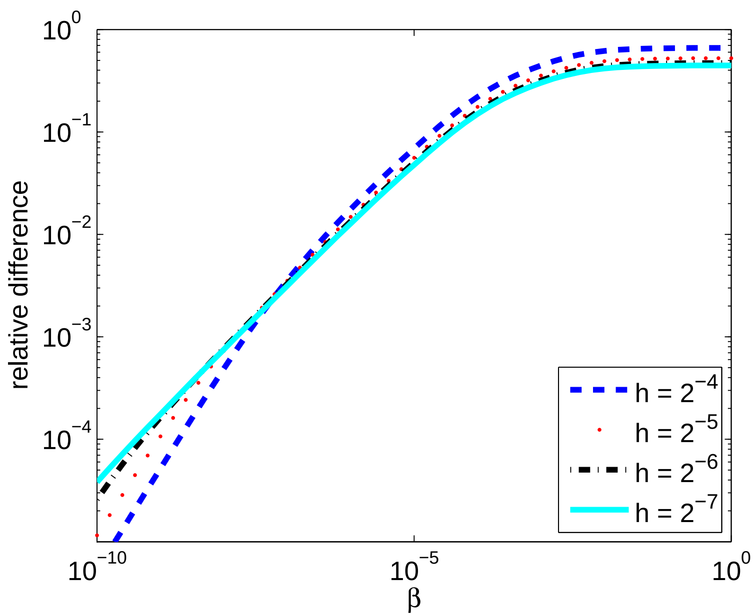
$\hat{\Omega} \neq \Omega$



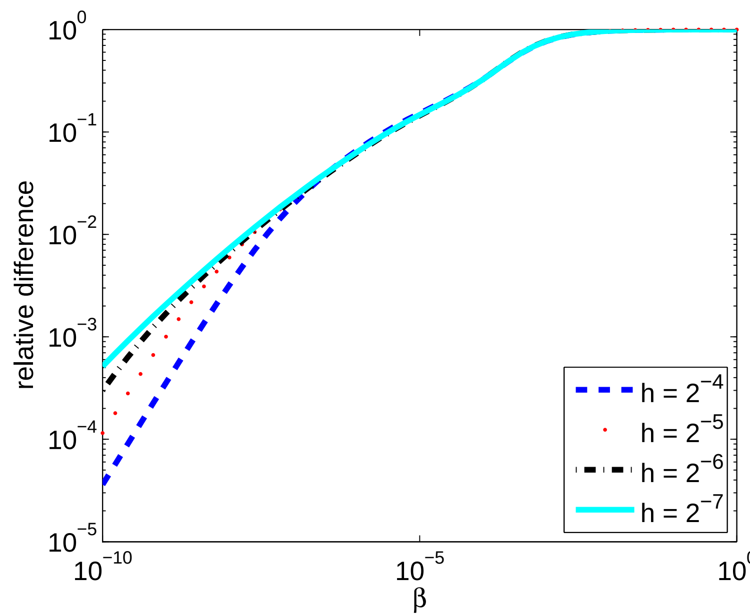
$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{u}(x,y)\vert_{\hat{\Omega}_1}$	$\hat{u}(x,y)\vert_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega/\hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x,y) : (x-\frac{5}{8})^2 + (y-\frac{3}{4})^2 \leq \frac{1}{25}\right\}$	$\partial\Omega$	2	0

Spectral properties of linear system

$\hat{\Omega} = \Omega$



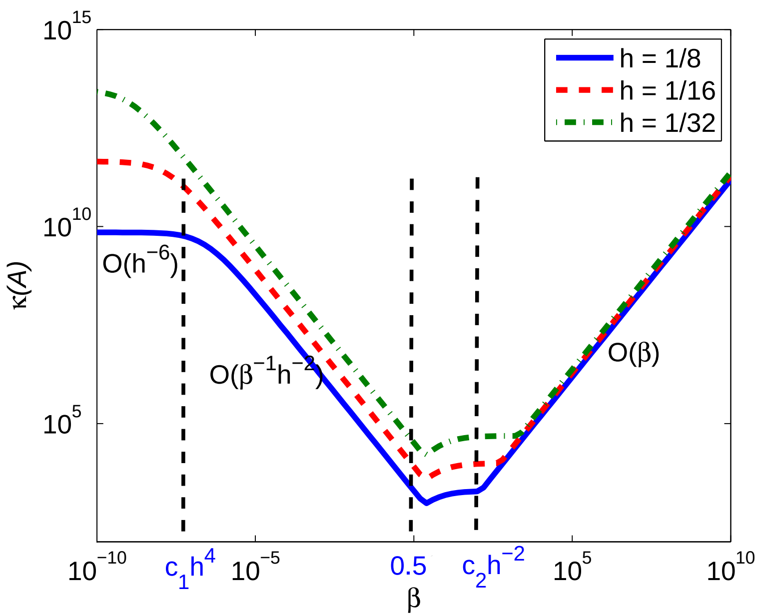
$\hat{\Omega} \neq \Omega$



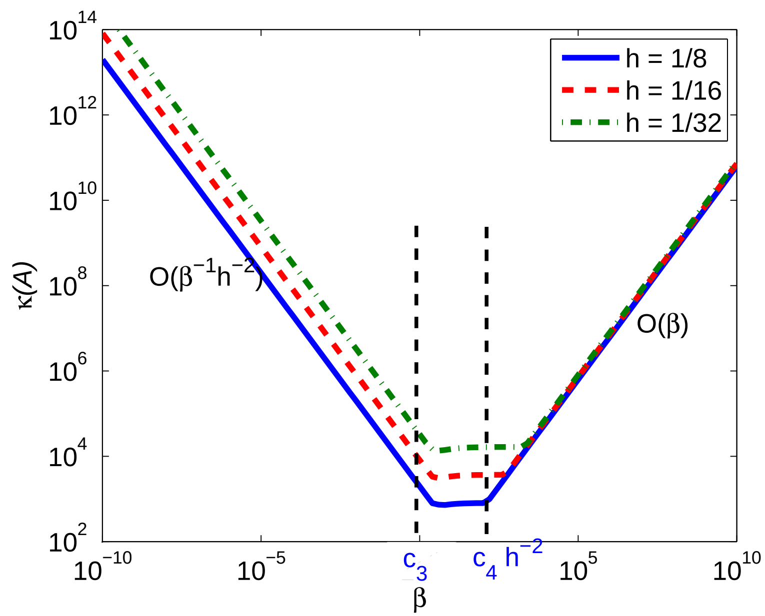
$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{u}(x,y)\vert_{\hat{\Omega}_1}$	$\hat{u}(x,y)\vert_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega / \hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x,y) : (x - \frac{5}{8})^2 + (y - \frac{3}{4})^2 \leq \frac{1}{25}\right\}$	$\partial\Omega$	2	0

Spectral properties of linear system

$\hat{\Omega} = \Omega$



$\hat{\Omega} \neq \Omega$



$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{u}(x,y)\|_{\hat{\Omega}_1}$	$\hat{u}(x,y)\|_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega / \hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x,y) : (x-\frac{5}{8})^2 + (y-\frac{3}{4})^2 \le \frac{1}{25}\right\}$	$\partial\Omega$	2	0

Effect of ill-conditioning on direct methods

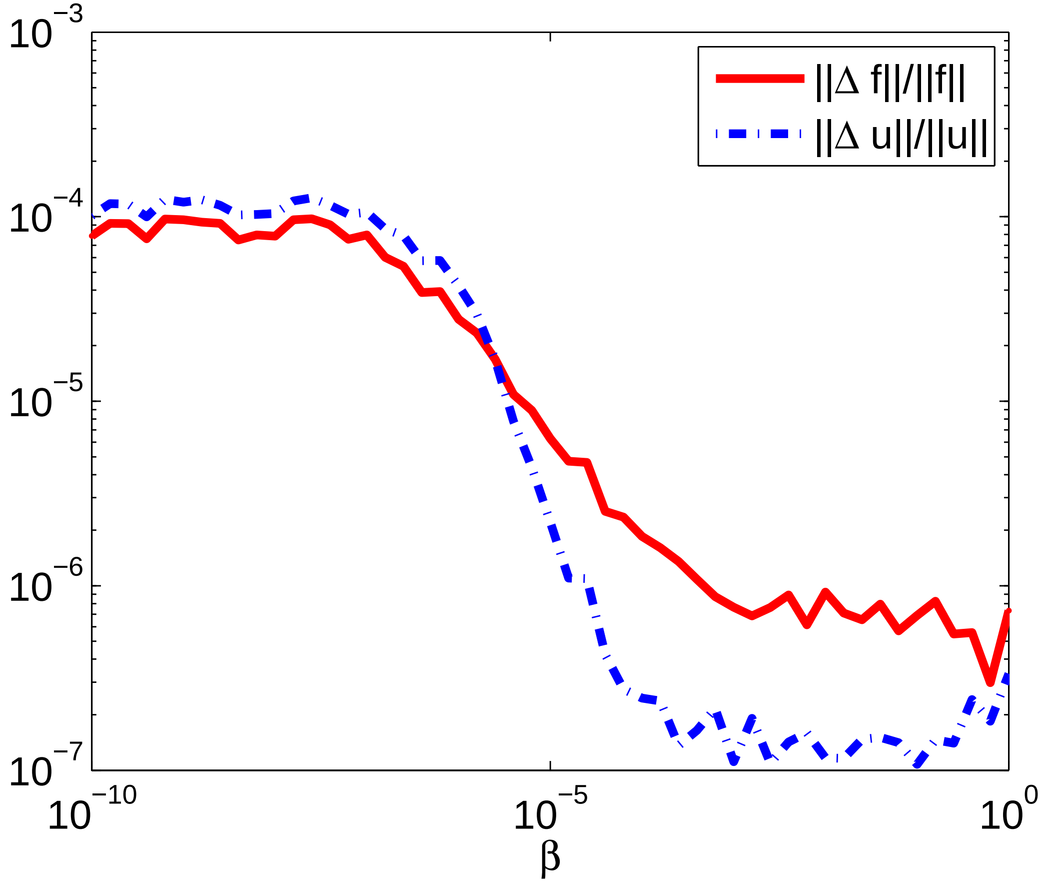
Consider solving $\mathcal{A}s = b$ with backward-stable method, then

$$||\Delta s||_2 \leq \mathsf{u}\gamma_N \kappa(\mathcal{A}) ||s||_2$$

Effect of ill-conditioning on direct methods

Consider solving $\mathcal{A}s = b$ with backward-stable method, then

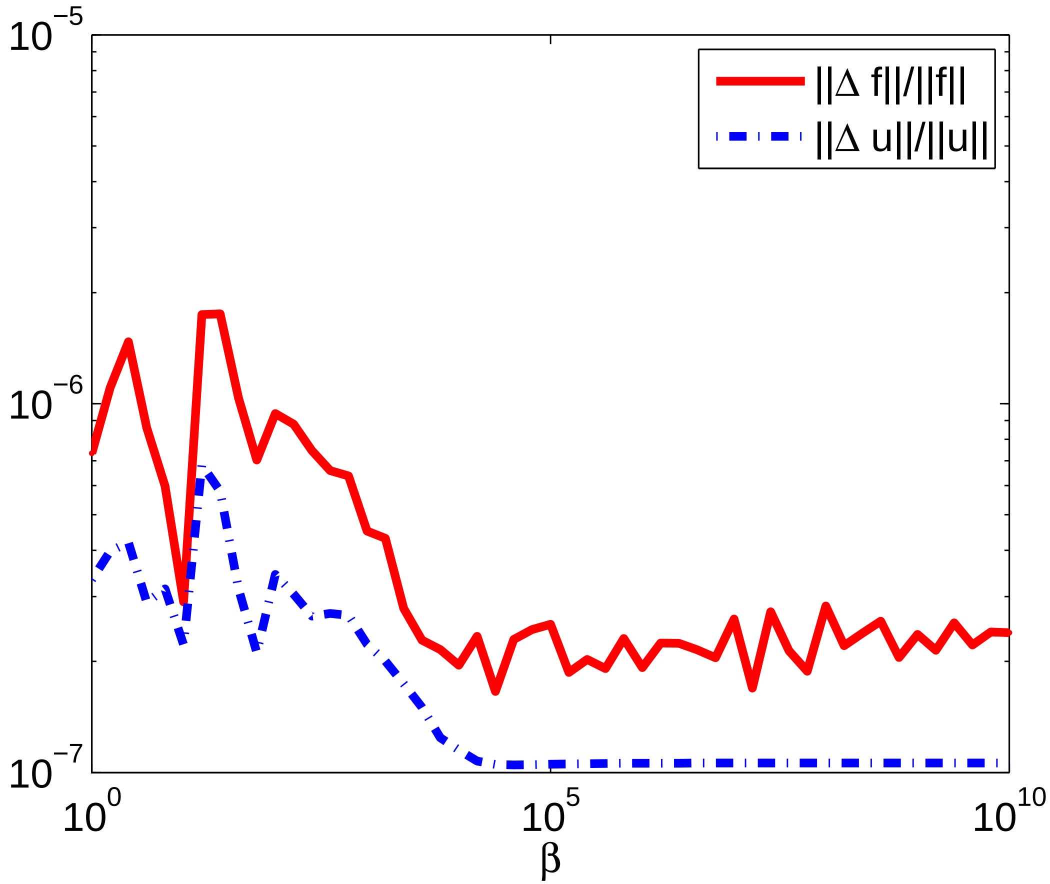
$$||\Delta s||_2 \leq \mathsf{u}\gamma_N \kappa(\mathcal{A})||s||_2$$



Effect of ill-conditioning on direct methods

Consider solving $\mathcal{A}s = b$ with backward-stable method, then

$$||\Delta s||_2 \leq \mathsf{u}\gamma_N \kappa(\mathcal{A})||s||_2$$



Effect of ill-conditioning on direct methods

Consider solving $\mathcal{A}s = b$ with backward-stable method, then

$$||\Delta s||_2 \leq \mathsf{u}\gamma_N \kappa(\mathcal{A})||s||_2$$

For large β, $\mathcal{A}$ has

- n eigenvalues that are $\mathcal{O}(\beta) \rightarrow$ correspond to $\mathbf{f}$
- $2n$ eigenvalues that are independent of $\beta \rightarrow$ correspond to $\mathbf{u}$ and l

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Distributed control - iterative methods

$$\min_{u,f} \frac{1}{2}u^T M u - u^T b + c + \beta f^T M f$$

subject to

$$Ku - Mf \quad = \quad d$$

$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \begin{bmatrix} f \\ u \\ l \end{bmatrix} = \begin{bmatrix} 0 \\ b \\ d \end{bmatrix}$$

Projected Preconditioned CG Method

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

Write

$$x = Yx_y + Zx_z,$$

where columns Z span nullspace of B and $[Y, Z]$ spans $\mathbb{R}^n$

$$\begin{aligned} BYx_y &= d, \\ Z^T AZx_z &= Z^T \left(b - AYx_y\right), \\ Y^T Bw &= Y^T (b - Ax). \end{aligned}$$

If $Z^T AZ$ is SPD, then use PCG with preconditioner $Z^T GZ$.

$$\|e_k\|_{Z^T AZ} \leq 2 \|e_0\|_{Z^T AZ} \left(\frac{\sqrt{\kappa((Z^T GZ)^{-1} Z^T AZ)} - 1}{\sqrt{\kappa((Z^T GZ)^{-1} Z^T AZ)} + 1} \right)^k$$

Projected Preconditioned CG Method

Remove references to Z by making substitutions (Gould, Hribar, Nocedal, 2001):

Choose initial point x satisfying $Bx = d$
Compute $r = Ax - b$
Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g \\ v \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

Set $p = -g$
repeat
Set $\alpha = r^T g / p^T A p$
Set $x = x + \alpha p$ and $r^+ = r + \alpha A p$
Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g^+ \\ v^+ \end{bmatrix} = \begin{bmatrix} r^+ \\ 0 \end{bmatrix}$$

Set $\beta = (r^+)^T g^+ / r^T g$
Set $p = -g^+ + \beta p$, $r = r^+$ and $g = g^+$
until converged

Projected Preconditioned CG Method

Remove references to Z by making substitutions:

Choose initial point x satisfying $Bx = d$

Compute $r = Ax - b$

Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g \\ v \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

Set $p = -g$

repeat

Set $\alpha = r^T g / p^T A p$

Set $x = x + \alpha p$ and $r^+ = r + \alpha A p$

Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g^+ \\ v^+ \end{bmatrix} = \begin{bmatrix} r^+ \\ 0 \end{bmatrix}$$

Set $\beta = (r^+)^T g^+ / r^T g$

Set $p = -g^+ + \beta p$, $r = r^+$ and $g = g^+$

until converged

Projected Preconditioned CG Method

Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ d \end{bmatrix}$$

Compute $r = Ax - b$
Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g \\ v \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

Set $p = -g$
repeat
  Set $\alpha = r^T g / p^T Ap$
  Set $x = x + \alpha p$ and $r^+ = r + \alpha Ap$
  Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g^+ \\ v^+ \end{bmatrix} = \begin{bmatrix} r^+ \\ 0 \end{bmatrix}$$

  Set $\beta = (r^+)^T g^+ / r^T g$
  Set $p = -g^+ + \beta p$, $r = r^+$ and $g = g^+$
until converged

Projected Preconditioned CG Method

Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ d \end{bmatrix}$$

Compute $r = Ax - b$
Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g \\ v \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

Set $p = -g$ and $y = -v$
repeat
Set $\alpha = r^T g / p^T Ap$
Set $x = x + \alpha p$ and $r^+ = r + \alpha Ap$
Solve

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} g^+ \\ v^+ \end{bmatrix} = \begin{bmatrix} r^+ \\ 0 \end{bmatrix}$$

Set $\beta = (r^+)^T g^+ / r^T g$
Set $p = -g^+ + \beta p$, $r = r^+ - B^T v^+$, $y = y - v^+$ and $g = g^+$
until converged

Projected Preconditioned CG Method

(Dollar 2005) Can be generalised to

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Constraint preconditioners

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \qquad \mathcal{P} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix}$$

Constraint preconditioners

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \qquad \mathcal{P} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix}$$

Theorem (Keller, Gould, Wathen, 2000): If $A, G \in \mathbb{R}^{n \times n}$ are symmetric and $B \in \mathbb{R}^{m \times n}$ has full row rank, then $\mathcal{P}^{-1}\mathcal{A}$ has

- $2m$ eigenvalues at 1
- remaining $n - m$ are defined by

$$Z^T A Z x = \lambda Z^T G Z x,$$

where the columns of $Z \in \mathbb{R}^{n \times (n-m)}$ span nullspace of B. If G is nonsingular, then these eigenvalues interlace the eigenvalues of $G^{-1}A$. The Krylov subspace wrt $\mathcal{P}^{-1}\mathcal{A}$ has dimension at most $n - m + 2$

Preconditioner

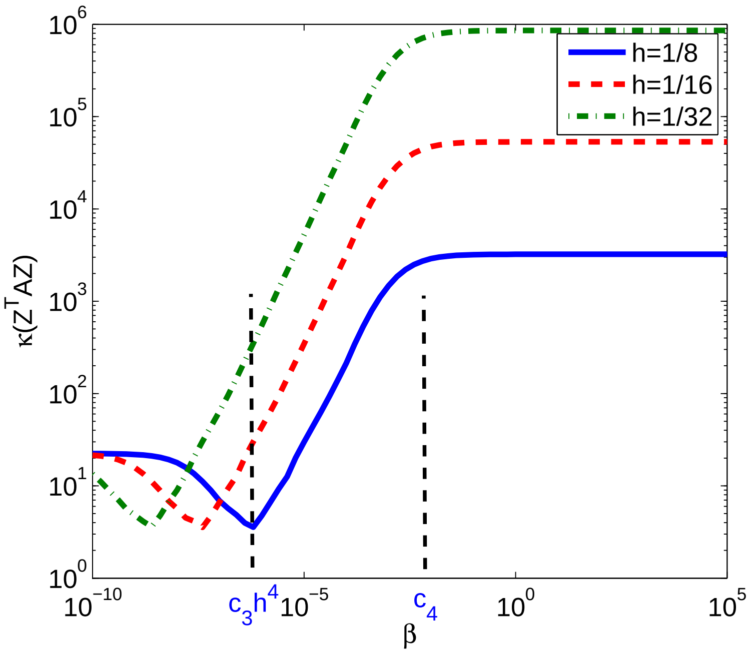
$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \quad Z = \begin{bmatrix} M^{-1}K \\ I \end{bmatrix}$$

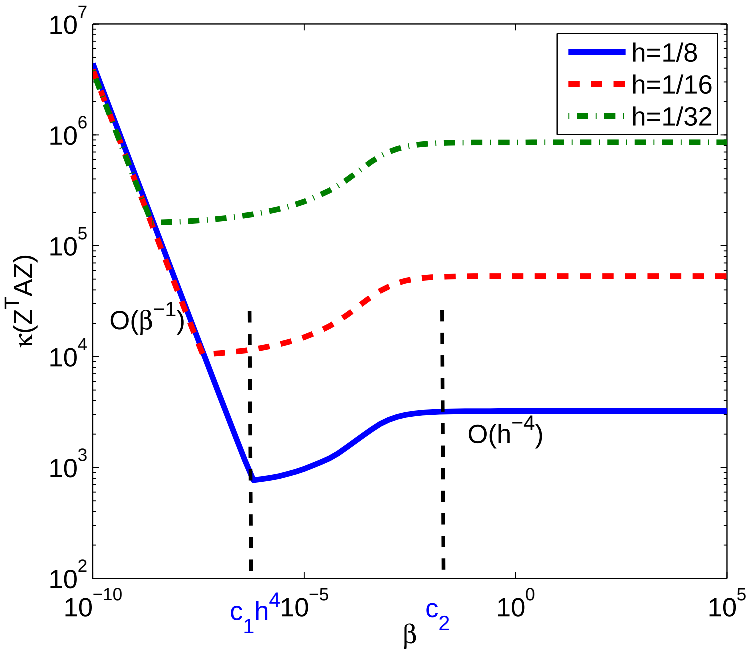
$$Z^T A Z = 2\beta K^T M^{-1} K + \bar{M}$$

Preconditioner

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \quad Z = \begin{bmatrix} M^{-1}K \\ I \end{bmatrix}$$

$$Z^T AZ = 2\beta K^T M^{-1} K + \bar{M}$$





Preconditioner

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \quad Z = \begin{bmatrix} M^{-1}K \\ I \end{bmatrix}$$

$$Z^T AZ = 2\beta K^T M^{-1} K + \bar{M}$$

$$\mathcal{P} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & 0 & K^T \\ -M & K & 0 \end{bmatrix} ?$$

$$Z^T GZ = 2\beta K^T M^{-1} K$$

$\bar{M} = M$	$\bar{M} \neq M$
$1 + \frac{ch^4}{2\beta} \leq \lambda \leq 1 + \frac{C}{2\beta}$ $c \leq \bar{c} \leq \bar{C} \leq C$	$1 + \frac{\bar{c}h^4}{2\beta} \leq \lambda \leq 1 + \frac{\bar{C}}{2\beta}$ $\lambda = 1$

Biros and Ghattas (2000)

Preconditioner

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \quad Z = \begin{bmatrix} M^{-1}K \\ I \end{bmatrix}$$

$$Z^T A Z = 2\beta K^T M^{-1} K + \bar{M}$$

$$\mathcal{P} = \begin{bmatrix} 0 & 0 & -M \\ 0 & 2\beta K^T M^{-1} K & K^T \\ -M & K & 0 \end{bmatrix}?$$

$$Z^T G Z = 2\beta K^T M^{-1} K$$

$\bar{M} = M$	$\bar{M} \neq M$
$1 + \frac{ch^4}{2\beta} \leq \lambda \leq 1 + \frac{C}{2\beta}$ $c \leq \bar{c} \leq \bar{C} \leq C$	$1 + \frac{\bar{c}h^4}{2\beta} \leq \lambda \leq 1 + \frac{\bar{C}}{2\beta}$ $\lambda = 1$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Numerical Example

Using bilinear **Q1** elements and setting $\beta = 5 \times 10^{-5}$:

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} 0 & 0 & -M \\ 0 & 2\beta K^T M^{-1} K & K^T \\ -M & K & 0 \end{bmatrix}$$

- Solves with M : Direct method (`HSL_MA57`) or 20 Chebyshev semi-iterations
- Solves with K : Direct method (`HSL_MA57`) or two(three) V-cycles of AMG (`HSL_MI20`)
- PPCG: relative tolerance 10^{-9} for $r^T Z(Z^T G Z)^{-1} Z^T r$, `HSL_MI27` (soon to be released)
- Fortran 95, NAG f95 compiler
- Hardware: Dell Precision T340, single Core2 Quad Q9550 processor (2.83GHz, 1333MHz FSB, 12MB L2 Cache), 4GB RAM

Numerical Example

Using bilinear **Q1** elements and setting $\beta = 5 \times 10^{-5}$:

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} 0 & 0 & -M \\ 0 & 2\beta K^T M^{-1} K & K^T \\ -M & K & 0 \end{bmatrix}$$

- Solves with M : Direct method (`HSL_MA57`) or 20 Chebyshev semi-iterations
- Solves with K : Direct method (`HSL_MA57`) or two(three) V-cycles of AMG (`HSL_MI20`)
- PPCG: relative tolerance 10^{-9} for $r^T Z(Z^T G Z)^{-1} Z^T r$, `HSL_MI27` (soon to be released)
- Fortran 95, NAG f95 compiler
- Hardware: Dell Precision T340, single Core2 Quad Q9550 processor (2.83GHz, 1333MHz FSB, 12MB L2 Cache), 4GB RAM

2D

N	n	Direct	PPCG(direct)	PPCG(approx)
8	147	0.002	0.001 (8)	0.003 (9)
16	675	0.01	0.006 (8)	0.011 (9)
32	2883	0.04	0.025 (8)	0.044 (9)
64	11907	0.19	0.12 (8)	0.17 (8)
128	48487	1.59	0.55 (7)	0.72 (8)
256	195075	8.82	3.27 (6)	3.18 (8)
512	783363	53.5	21.5 (6)	14.2 (8)

3D

N	n	Direct	PPCG(direct)	PPCG(approx)
4	81	0.001	0.002 (7)	0.002 (7)
8	1029	0.04	0.02 (8)	0.05 (8)
16	10125	1.25	0.33 (8)	0.64 (8)
32	89373	38.0	6.61 (7)	7.32 (7)
64	750141	1000+	217 (5)	59.0 (6)

Numerical Example

Using bilinear **Q1** elements and setting $\beta = 5 \times 10^{-5}$:

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} 0 & 0 & -M \\ 0 & 2\beta K^T M^{-1} K & K^T \\ -M & K & 0 \end{bmatrix}$$

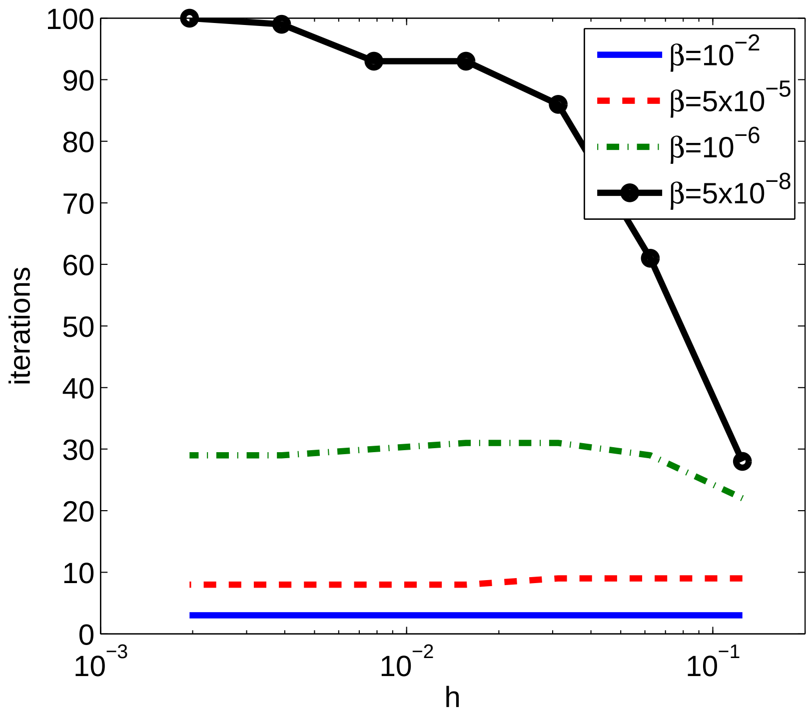
- Solves with M : Direct method (`HSL_MA57`) or 20 Chebyshev semi-iterations
- Solves with K : Direct method (`HSL_MA57`) or two(three) V-cycles of AMG (`HSL_MI20`)
- PPCG: relative tolerance 10^{-9} for $r^T Z(Z^T G Z)^{-1} Z^T r$, `HSL_MI27` (soon to be released)
- Fortran 95, NAG f95 compiler
- Hardware: Dell Precision T340, single Core2 Quad Q9550 processor (2.83GHz, 1333MHz FSB, 12MB L2 Cache), 4GB RAM

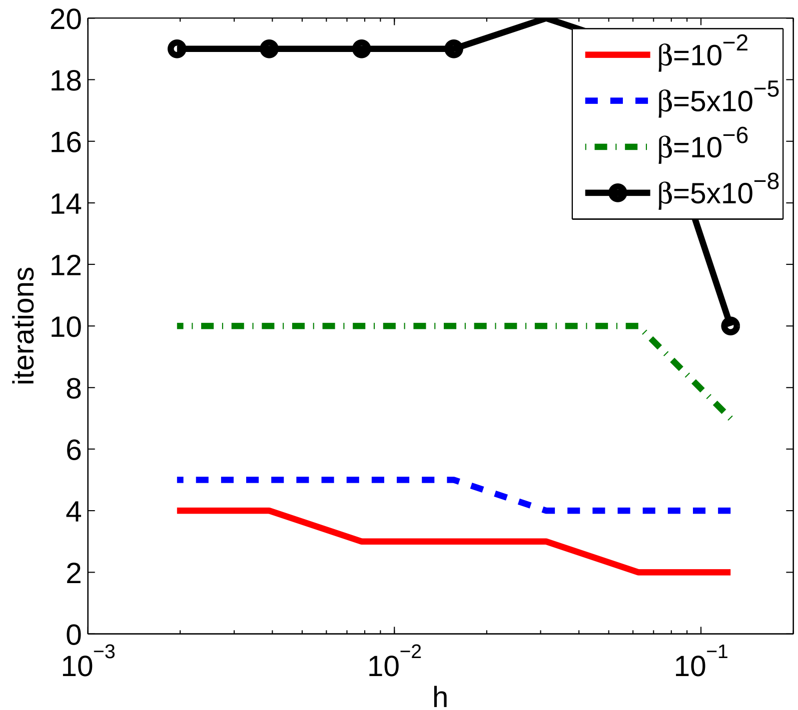
N	n	Direct	PPCG(direct)	PPCG(approx)
8	147	0.002	0.001 (4)	0.002 (4)
16	675	0.01	0.01 (4)	0.007 (4)
32	2883	0.10	0.02 (4)	0.03 (4)
64	11907	0.35	0.10 (4)	0.13 (5)
128	48487	2.78	0.50 (5)	0.53 (5)
256	195075	16.8	3.11 (5)	2.36 (5)
512	783363	147	20.5 (5)	10.3 (5)

N	n	Direct	PPCG(direct)	PPCG(approx)
4	81	0.001	0.001 (3)	0.001 (3)
8	1029	0.05	0.02 (4)	0.03 (4)
16	10125	1.19	0.31 (5)	0.49 (5)
32	89373	59.2	6.32 (5)	6.00 (5)
64	750141	1000+	219 (5)	58.9 (5)

Behaviour of preconditioner with β

$\bar{M} = M$	$\bar{M} \neq M$
$1 + \frac{ch^4}{2\beta} \leq \lambda \leq 1 + \frac{C}{2\beta}$ $c \leq \bar{c} \leq \bar{C} \leq C$	$1 + \frac{\bar{c}h^4}{2\beta} \leq \lambda \leq 1 + \frac{\bar{C}}{2\beta}$ $\lambda = 1$





Neumann boundary control

$$\min_{\mathtt{u},\mathtt{g}} \frac{1}{2} \left\| \omega(x) \left(\mathsf{u} - \hat{\mathsf{u}} \right) \right\|_2^2 + \beta \left\| \mathsf{g} \right\|_2^2$$

subject to

$$\begin{aligned} \mathcal{L}\mathsf{u} &= \mathtt{f} \text{ in } \Omega \\ \frac{\partial \mathsf{u}}{\partial n} &= \mathsf{g} \text{ on } \delta\Omega \end{aligned}$$

Here

$$\omega(x) = \begin{cases} 1 & x \in \hat{\Omega} \\ 0 & \text{otherwise} \end{cases}$$

Neumann boundary control

Discretize:

$$\mathtt{u_h} = \sum u_j \phi_j + \sum \hat{u}_j \hat{\phi}_j, \quad \mathtt{g_h} = \sum g_j \hat{\phi}_j, \quad \mathtt{f_h} = \sum f_j \phi_j + \sum \hat{f}_j \hat{\phi}_j$$

$$\min_{\mathtt{u_h},\mathtt{g_h}} \frac{1}{2} \left\| \omega(x) \left(\mathtt{u_h} - \widehat{u} \right) \right\|_2^2 + \beta \left\| \mathtt{g_h} \right\|_2^2$$

subject to

$$\begin{aligned} \mathcal{L}\mathtt{u_h} &= \mathtt{f_h} \text{ in } \Omega \\ \frac{\partial \mathtt{u_h}}{\partial \mathtt{n}} &= \mathtt{g_h} \text{ on } \partial\Omega \end{aligned}$$

Neumann boundary control

$$
\begin{aligned}
\|\omega(x)\,(\mathtt{u_h} - \widehat{u})\|_2^2 &= \int_\Omega \omega(x)\,(\mathtt{u_h} - \widehat{u})^2 \\
&= \begin{bmatrix} u^T & \hat{u}^T \end{bmatrix} \begin{bmatrix} \bar{M}_{II} & \bar{M}_{IB} \\ \bar{M}_{BI} & \bar{M}_{BB} \end{bmatrix} \begin{bmatrix} u \\ \hat{u} \end{bmatrix} - u^T b - \hat{u}^T \hat{b} + c \\
\|\mathtt{g_h}\|_2^2 &= g^T M_g g \\
\begin{bmatrix} d \\ \hat{d} \end{bmatrix} &= \begin{bmatrix} K_{II} & K_{IB} \\ K_{BI} & K_{BB} \end{bmatrix} \begin{bmatrix} u \\ \hat{u} \end{bmatrix} - \begin{bmatrix} 0 \\ M_g \end{bmatrix} g
\end{aligned}
$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Neumann boundary control

$$\min_{u,\hat{u},f} \begin{bmatrix} u^T & \hat{u}^T \end{bmatrix} \begin{bmatrix} \bar{M}_{II} & \bar{M}_{IB} \\ \bar{M}_{BI} & \bar{M}_{BB} \end{bmatrix} \begin{bmatrix} u \\ \hat{u} \end{bmatrix} - u^T b - \hat{u}^T \hat{b} + c + \beta g^T M_g g$$

subject to

$$\begin{bmatrix} K_{II} & K_{IB} \\ K_{BI} & K_{BB} \end{bmatrix} \begin{bmatrix} u \\ \hat{u} \end{bmatrix} - \begin{bmatrix} 0 \\ M_g \end{bmatrix} g = \begin{bmatrix} d \\ \hat{d} \end{bmatrix}$$

$$\left[\begin{array}{ccc|cc} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ \hline 0 & K_{II} & K_{IB} & 0 & 0 \\ -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right] \begin{bmatrix} g \\ u \\ \hat{u} \\ l \\ \hat{l} \end{bmatrix} = \begin{bmatrix} 0 \\ b \\ \hat{b} \\ d \\ \hat{d} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & B_1^T \\ 0 & G_1 & B_2^T \\ B_1 & B_2 & 0 \end{bmatrix}$$

$$\left[\begin{array}{ccc|cc} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ \hline 0 & K_{II} & K_{IB} & 0 & 0 \\ -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right] \Leftrightarrow \left[\begin{array}{cccc|c} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ 0 & K_{II} & K_{IB} & 0 & 0 \\ \hline -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|cc} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ \hline 0 & K_{II} & K_{IB} & 0 & 0 \\ -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right] \Leftrightarrow \left[\begin{array}{cccc|c} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ 0 & K_{II} & K_{IB} & 0 & 0 \\ \hline -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right]$$

$$Z^T A Z = \begin{bmatrix} \hat{A} & \hat{B}^T \\ \hat{B} & 0 \end{bmatrix}, \quad \hat{A} = 2\beta \begin{bmatrix} K_{IB} \\ K_{BB} \end{bmatrix} M_g^{-1} \begin{bmatrix} K_{BI} & K_{BB} \end{bmatrix} + \bar{M}, \quad \hat{B} = \begin{bmatrix} K_{II} & K_{IB} \end{bmatrix}$$

$$\left[\begin{array}{ccc|cc} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ \hline 0 & K_{II} & K_{IB} & 0 & 0 \\ -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right] \Leftrightarrow \left[\begin{array}{cccc|c} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ 0 & K_{II} & K_{IB} & 0 & 0 \\ \hline -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right]$$

$$Z^T A Z = \begin{bmatrix} \hat{A} & \hat{B}^T \\ \hat{B} & 0 \end{bmatrix}, \quad \hat{A} = 2\beta \begin{bmatrix} K_{IB} \\ K_{BB} \end{bmatrix} M_g^{-1} \begin{bmatrix} K_{BI} & K_{BB} \end{bmatrix} + \bar{M}, \quad \hat{B} = \begin{bmatrix} K_{II} & K_{IB} \end{bmatrix}$$

$Z^T AZ$ indefinite $\Rightarrow$ avoid PPCG
Solve $Z^T AZx_z = Z^T (b - AYx_y)$ with BICGSTAB, MINRES, SQMR or...

$$\left[\begin{array}{ccc|cc} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ \hline 0 & K_{II} & K_{IB} & 0 & 0 \\ -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right] \Leftrightarrow \left[\begin{array}{cccc|c} 2\beta M_g & 0 & 0 & 0 & -M_g \\ 0 & \bar{M}_{II} & \bar{M}_{IB} & K_{II} & K_{IB} \\ 0 & \bar{M}_{BI} & \bar{M}_{BB} & K_{BI} & K_{BB} \\ 0 & K_{II} & K_{IB} & 0 & 0 \\ \hline -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right]$$

$$Z^T AZ = \begin{bmatrix} \hat{A} & \hat{B}^T \\ \hat{B} & 0 \end{bmatrix}, \quad \hat{A} = 2\beta \begin{bmatrix} K_{IB} \\ K_{BB} \end{bmatrix} M_g^{-1} \begin{bmatrix} K_{BI} & K_{BB} \end{bmatrix} + \bar{M}, \quad \hat{B} = \begin{bmatrix} K_{II} & K_{IB} \end{bmatrix}$$

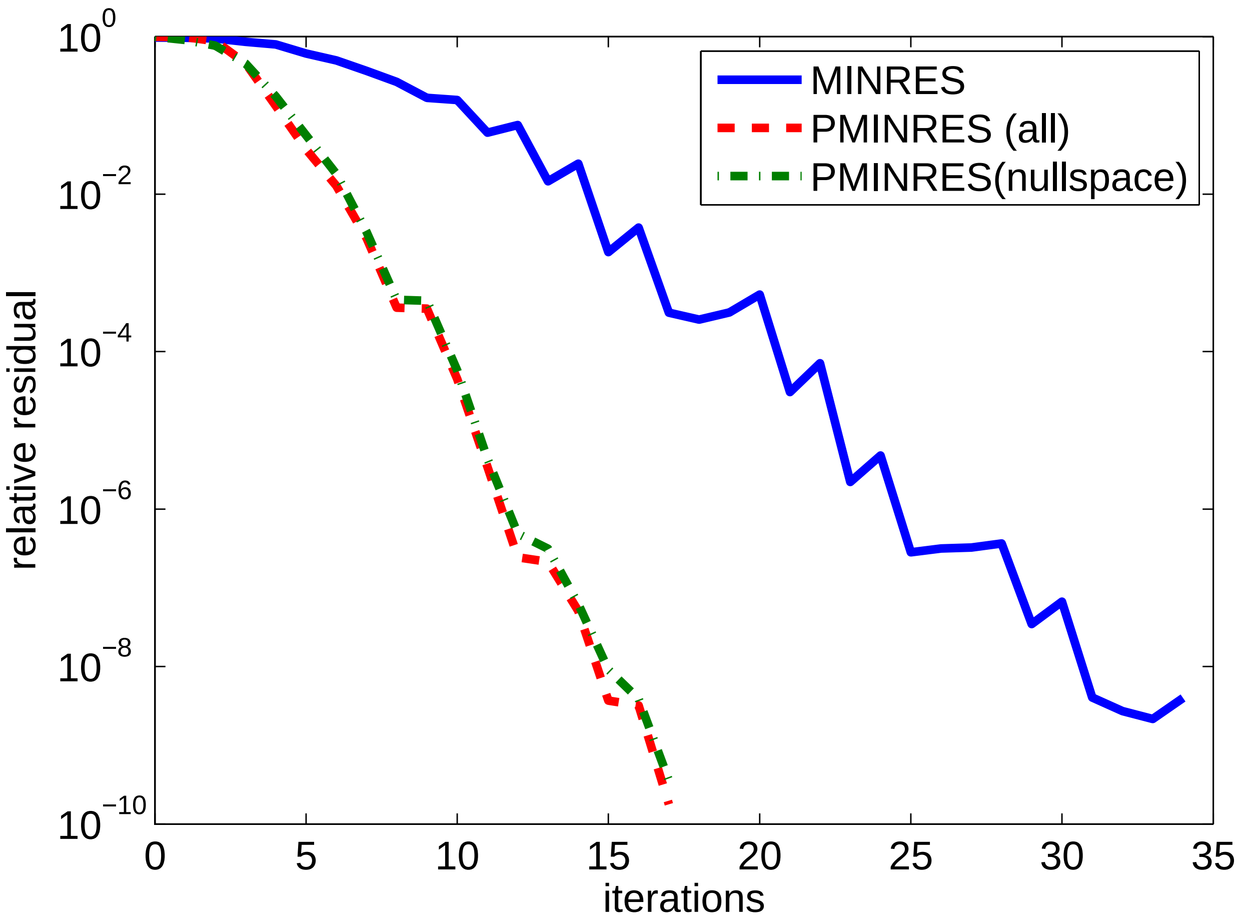
$Z^T AZ$ indefinite $\Rightarrow$ avoid PPCG
Solve $Z^T AZx_z = Z^T (b - AYx_y)$ with BICGSTAB, MINRES, SQMR or...
Solves with

$$\left[\begin{array}{cccc|c} & & & & -M_g \\ & G & & & K_{IB} \\ & & & & K_{BB} \\ & & & & 0 \\ \hline -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right] \text{ and } \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & -M_f^T \\ 0 & I & 0 & 0 & K_{IB} \\ 0 & 0 & I & 0 & K_{BB} \\ 0 & 0 & 0 & I & 0 \\ \hline -M_g & K_{BI} & K_{BB} & 0 & 0 \end{array}\right]$$

PMINRES and Distributed Control

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & M & K^T \\ -M & K & 0 \end{bmatrix}$$

MINRES	PMINRES
$P = \begin{bmatrix} 2\beta M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & KM^{-1}K \end{bmatrix}$	$P = \begin{bmatrix} 0 & 0 & -M \\ 0 & 2\beta K^T M^{-1} K & K^T \\ -M & K & 0 \end{bmatrix}$
$\lambda = 1$ $\frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_2}{\beta}}\right)$ $\frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_2}{\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right)$	$\lambda = 1$ $1+\frac{ch^4}{2\beta} \leq \lambda \leq 1+\frac{C}{2\beta}$
2 solves with M 2 solves with K 5 matrix-vector multiplications with M 2 matrix-vector multiplications with K	2 solves with M 2 solves with K 3 matrix-vector multiplications with M 3 matrix-vector multiplications with K


relative residual
iterations
MINRES
PMINRES (all)
PMINRES(nullspace)
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
0
5
10
15
20
25
30
35

Conclusions and Future Work

- PDE-constrained problems difficult to solve
- Avoid any solves with discretized PDE
- Use block structure
- Constraint preconditioners lead to projected iterative methods
- Mesh size independent convergence
- Regularization parameter independent convergence?
- Nonlinear PDEs, time-dependent PDEs, different regularization terms
- `HSL_MA57` and `HSL_MI20` are part of `HSL2007`, which is free for all academics
- `HSL_MI27` will be part of `HSL2007`
- 'Optimal solvers for PDE-constrained optimization' Rees, Dollar, Wathen, SISC 2010
- 'Properties of linear systems in PDE-constrained optimization. Part I: Distributed control', Dollar, RAL TR-2009-017
- 'Properties of linear systems in PDE-constrained optimization. Part II: Boundary control', Thorne, RAL TR-2009-018
- 'PDE-constrained optimization and constraint preconditioners', Thorne, RAL TR-2010-016