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

Sue Thorne

STFC Rutherford Appleton Laboratory

In collaboration with

- Nick Gould (STFC Rutherford Appleton Laboratory)
- Dominique Orban (Ecole Polytechnique de Montreal)
- Tyrone Rees (University of British Columbia)
- Valeria Simoncini (Università di Bologna)
- Andy Wathen (University of Oxford)

Science & Technology Facilities Council
Rutherford Appleton Laboratory

PDE-constrained optimization

- Given u and boundary conditions g, calculate y, where

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

on some domain Ω

PDE-constrained optimization

- Given u and boundary conditions g, calculate y, where

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

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

PDE-constrained optimization

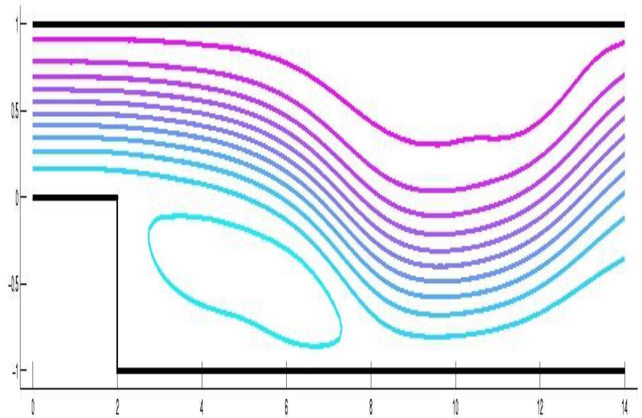
- Given u and boundary conditions g, calculate y, where

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

  on some domain Ω

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

Different target temperatures



Reduce recirculation

Distributed control

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

subject to

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

Here

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

Distributed control

Discretize:

$$\mathrm{y_h} = \sum y_j \phi_j, \quad \mathrm{u_h} = \sum u_j \phi_j$$

$$\min_{\mathrm{y_h},\mathrm{u_h}} \frac{1}{2} \left\| \omega(x) \left(\mathrm{y_h} - \widehat{y} \right) \right\|_2^2 + \beta \left\| \mathrm{u_h} \right\|_2^2$$

subject to

$$\begin{aligned} \mathcal{L}\mathrm{y_h} &= \mathrm{u_h} \text{ in } \Omega \\ \mathrm{y_h} &= g \text{ on } \delta\Omega \end{aligned}$$

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

Distributed control

$$
\begin{aligned}
\|\mathtt{u_h}\|_2^2 &= \int_\Omega \mathtt{u_h}^2 \\
&= \sum_i \sum_j u_i u_j \int_\Omega \phi_i \phi_j \\
&= u^T M u, \\
\frac{1}{2}\|\omega(x)\,(\mathtt{y_h} - \widehat{y})\|_2^2 &= \frac{1}{2}\int_\Omega \omega(x)\,(\mathtt{y_h} - \widehat{y})^2 \\
&= \frac{1}{2}\sum_i \sum_j y_i y_j \int_\Omega \omega_i \omega_j \phi_i \phi_j - 2\sum_j y_j \int_\Omega \omega_j \phi_j \widehat{y} + \frac{1}{2}\int_{\hat{\Omega}} \widehat{y}^2 \\
&= \frac{1}{2} y^T \bar{M} y - y^T b + c, \\
Ky &= Mu + d,
\end{aligned}
$$

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

Distributed control

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

subject to

$$Ky - Mu \quad = \quad d$$

Distributed control

$$\min_{y,u} \frac{1}{2} y^T \bar{M} y - y^T b + c + \beta u^T M u + l^T (Ky - Mu - d)$$

Optimality conditions:

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

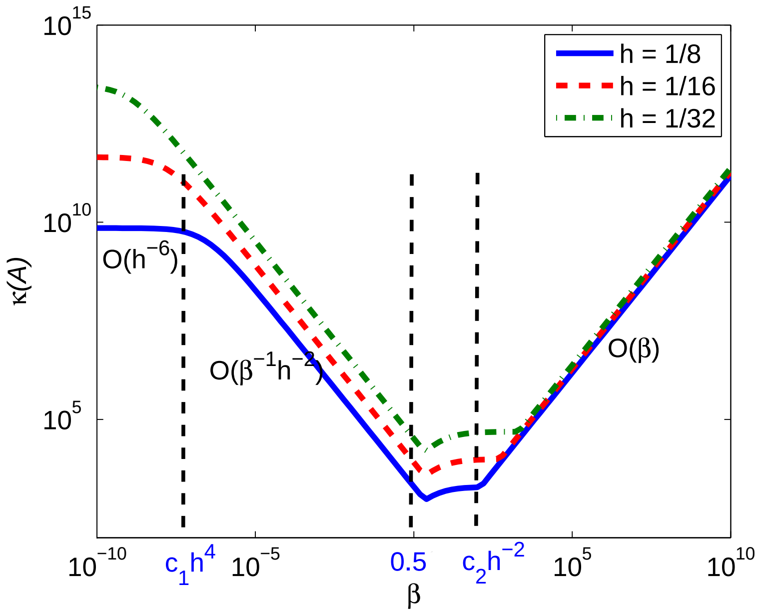
Direct vs Iterative Methods

	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?

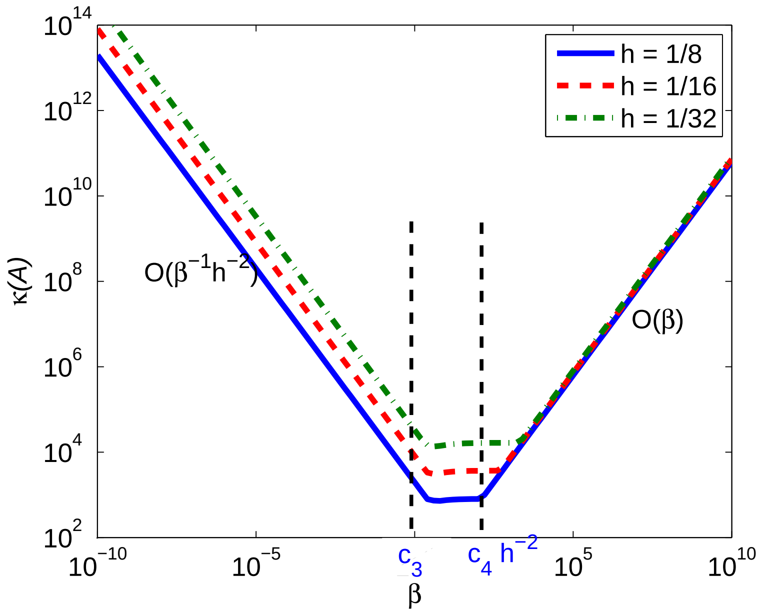
Definition: let $\kappa(\mathcal{A}) = ||\mathcal{A}||_2||\mathcal{A}^{-1}||_2$ be the condition number of $\mathcal{A}$

Spectral properties of linear system

$\hat{\Omega} = \Omega$



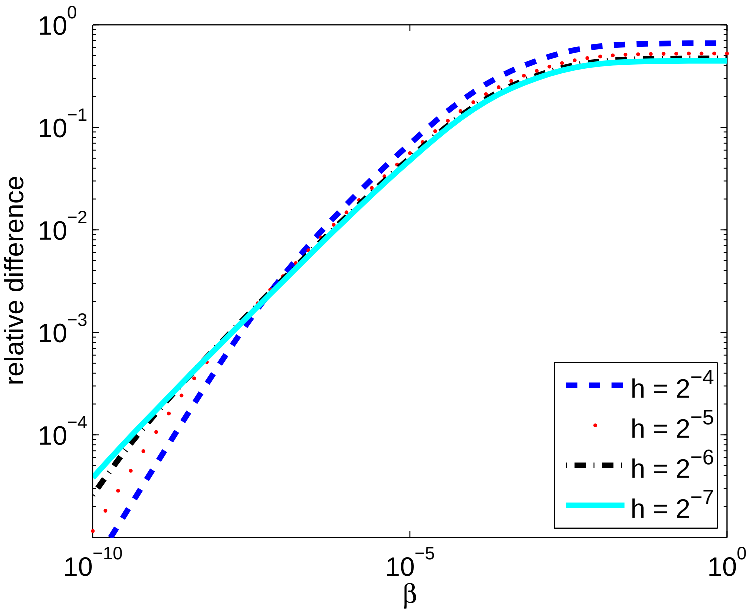
$\hat{\Omega} \neq \Omega$



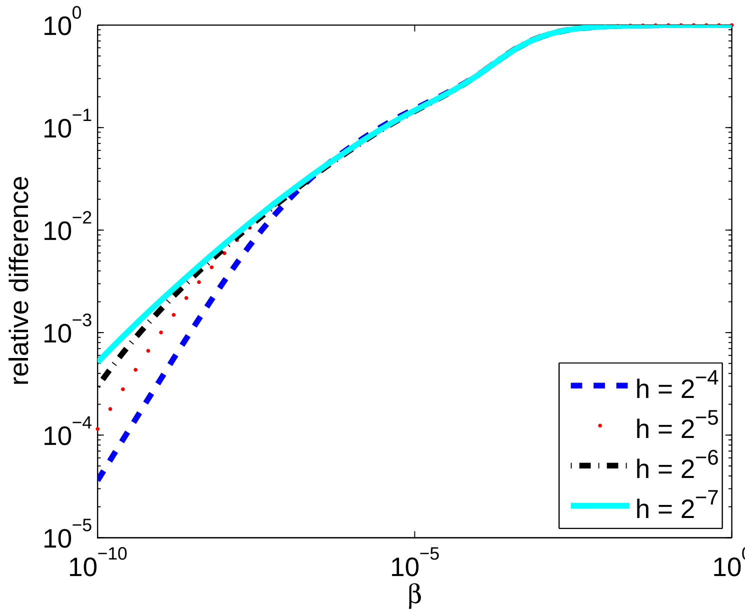
$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{y}(x_1, x_2)\|_{\hat{\Omega}_1}$	$\hat{y}(x_1, x_2)\|_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega / \hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x,y) : (x-\frac{5}{8})^2 + (y-\frac{3}{4})^2 \leq \frac{1}{25}\right\}$	$\partial\Omega$	2	0

Spectral properties of linear system

$\hat{\Omega} = \Omega$



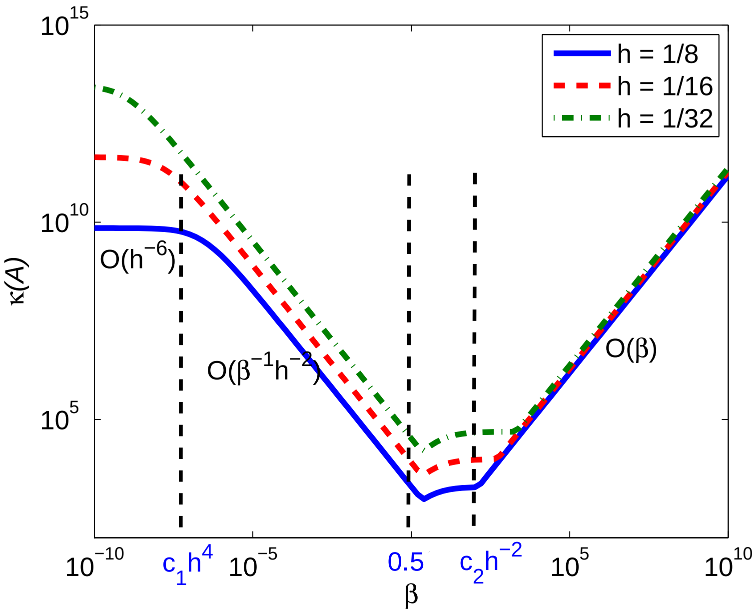
$\hat{\Omega} \neq \Omega$



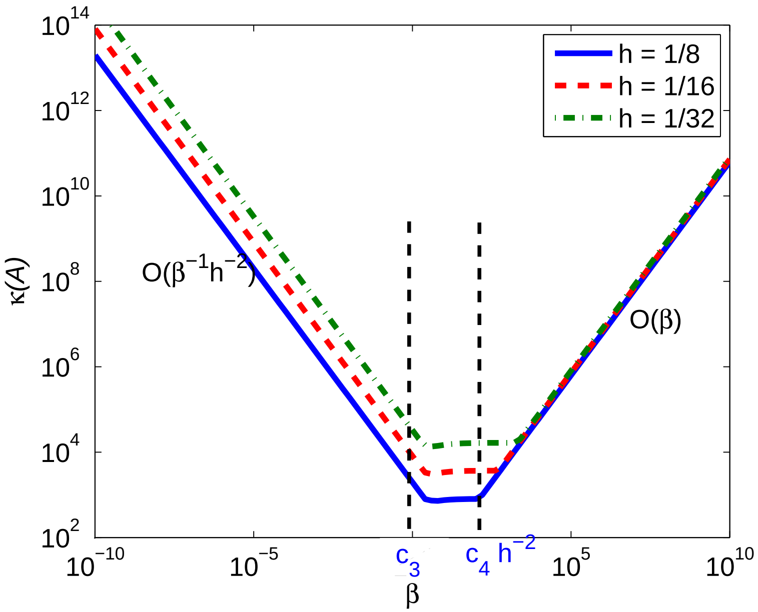
$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{y}(x_1, x_2)\|_{\hat{\Omega}_1}$	$\hat{y}(x_1, x_2)\|_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega / \hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x,y) : (x-\frac{5}{8})^2 + (y-\frac{3}{4})^2 \leq \frac{1}{25}\right\}$	$\partial\Omega$	2	0

Spectral properties of linear system

$\hat{\Omega} = \Omega$



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



$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{y}(x_1, x_2)\vert_{\hat{\Omega}_1}$	$\hat{y}(x_1, x_2)\vert_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega / \hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	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

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Distributed control - iterative methods

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

subject to

$$Ky - Mu \quad = \quad d$$

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

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 $Z^T GZ$.

$$\|e_k\|_{Z^TAZ} \le 2\,\|e_0\|_{Z^TAZ}\left(\frac{\sqrt{\kappa((Z^TGZ)^{-1}Z^TAZ)}-1}{\sqrt{\kappa((Z^TGZ)^{-1}Z^TAZ)}+1}\right)^k$$

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

(Dollar 2005) Can be generalised to

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

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

$$Z^T AZx = \lambda Z^T GZx,$$

where the columns of $Z \in \mathbb{R}^{n\times(n-m)}$ span nullspace of B.

Preconditioner

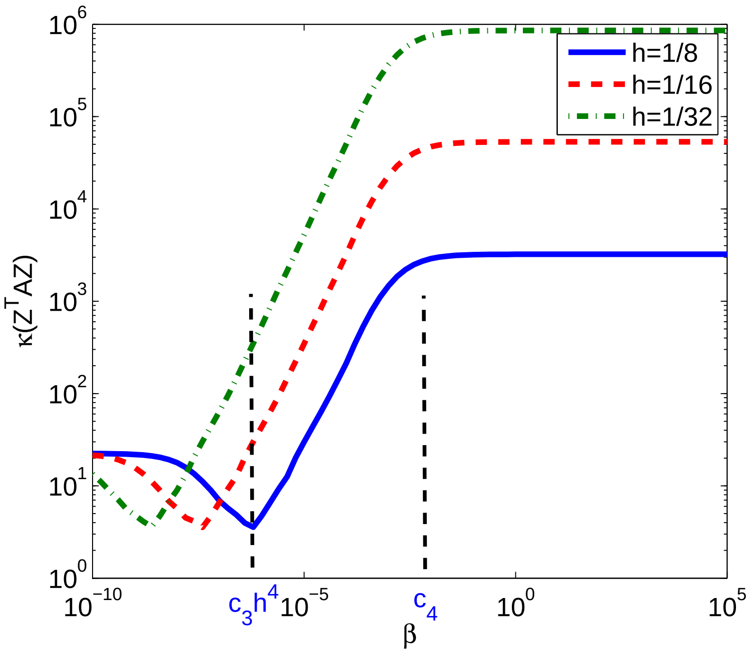
$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \quad Z = \begin{bmatrix} M^{-1}K \\ I \end{bmatrix}$$

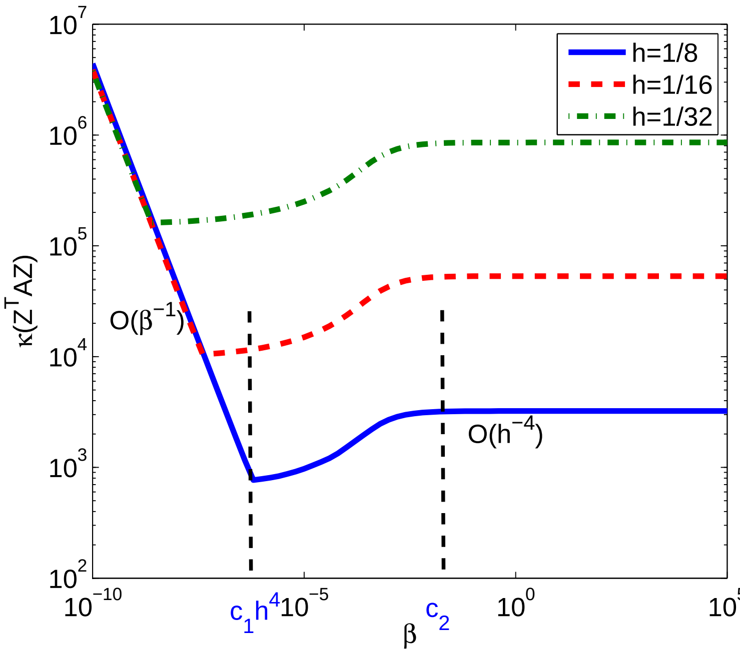
$$Z^T AZ = 2\beta K^T M^{-1} K + \bar{M}$$

Preconditioner

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

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





Preconditioner

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

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

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

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

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

Biros and Ghattas (2000)

Preconditioner

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

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

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

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

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

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`
- 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`
- 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`
- Fortran 95, NAG f95 compiler
- Hardware: Dell Precision T340, single Core2 Quad Q9550 processor (2.83GHz, 1333MHz FSB, 12MB L2 Cache), 4GB RAM

2D

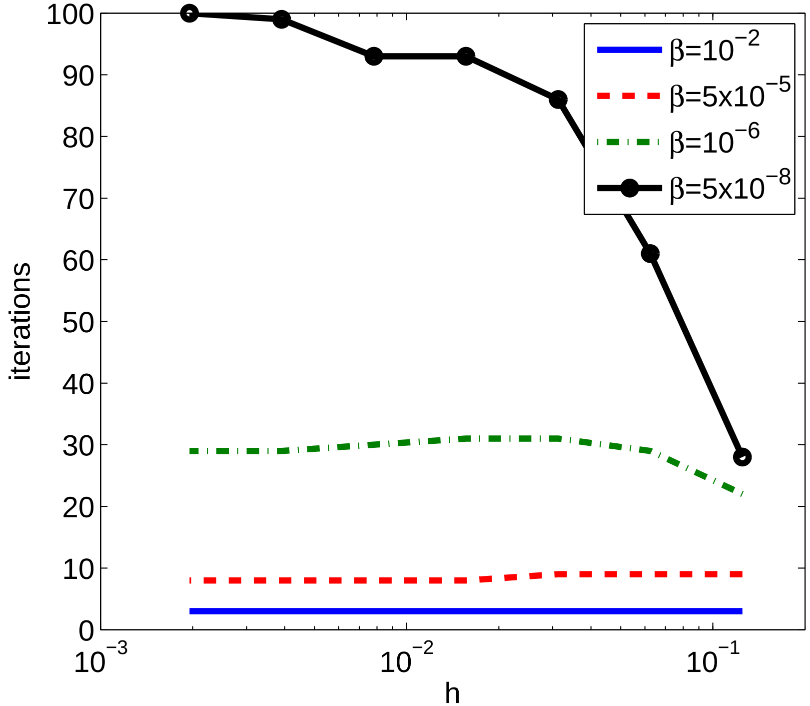
N	n	Direct	PPCG(direct)	PPCG(approx)
8	147	0.002	0.001 (4)	0.002 (4)
16	675	0.01	0.01 (4)	0.007 (4)
32	2883	0.10	0.02 (4)	0.03 (4)
64	11907	0.35	0.10 (4)	0.13 (5)
128	48487	2.78	0.50 (5)	0.53 (5)
256	195075	16.8	3.11 (5)	2.36 (5)
512	783363	147	20.5 (5)	10.3 (5)

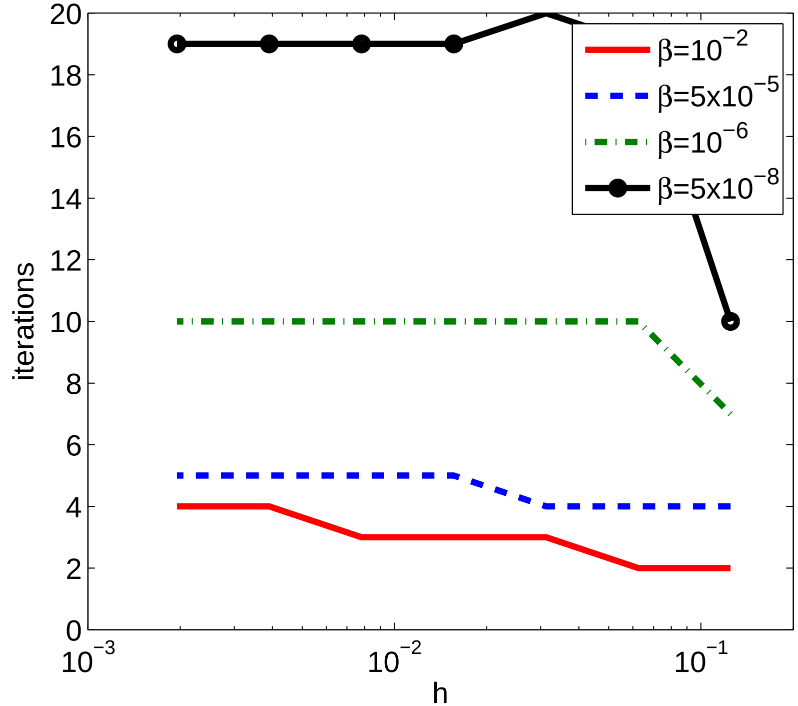
3D

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

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





Alternative methods

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

$$\Rightarrow l = 2\beta u$$

$$\begin{bmatrix} \bar{M} & K^T \\ K & -\frac{1}{2\beta} M \end{bmatrix} \begin{bmatrix} y \\ l \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

Use PPCG with preconditioner of the form

$$\begin{bmatrix} G & K^T \\ K & -\frac{1}{2\beta} M \end{bmatrix}$$

Reduced System and PPCG

$$\begin{bmatrix} \bar{M} & K \\ K & -\frac{1}{2\beta}M \end{bmatrix} \begin{bmatrix} y \\ l \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

Preconditioner

$$\begin{aligned} P &= \begin{bmatrix} I & -K \\ 0 & -\frac{1}{2\beta}M \end{bmatrix} \begin{bmatrix} 2\beta\tilde{K}M^{-1}\tilde{K} & 0 \\ 0 & -2\beta M^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -K & -\frac{1}{2\beta}M \end{bmatrix} \\ &= \begin{bmatrix} 2\beta\left(\tilde{K}M^{-1}\tilde{K} - KM^{-1}K\right) & K \\ K & -\frac{1}{2\beta}M \end{bmatrix} \end{aligned}$$

Non-unit eigenvalues identical to those of original system with corresponding constraint preconditioner
However, solve with preconditioner for reduced system is more expensive (one extra matrix-vector multiplication with M)

	Target 1		Target 2	
h^{-1}	PPCG(orig)	PPCG(red)	PPCG(orig)	PPCG(red)
2^4	0.10 (10)	0.11 (9)	0.07 (6)	0.08 (6)
2^5	0.24 (10)	0.28 (9)	0.16 (6)	0.20 (6)
2^6	0.86 (10)	0.98 (9)	0.50 (5)	0.61 (5)
2^7	3.99 (10)	4.50 (9)	2.31 (5)	2.73 (5)
2^8	17.3 (10)	21.8 (10)	10.1 (5)	12.5 (5)
2^9	69.8 (10)	87.9 (10)	45.8 (6)	57.2 (6)

$$\beta = 5 \times 10^{-5}$$

Solving original saddle-point problem: MINRES/SYMMBK

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

Use MINRES/SYMMBK with preconditioner of the form

$$\begin{bmatrix} 2\beta M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & \tilde{K} M^{-1} \tilde{K} \end{bmatrix}$$

$\bar{M} = M$	$\lambda = 1$ $\frac{1}{2}\left(1+\sqrt{5+\frac{ch^4}{2\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1+\sqrt{5+\frac{C}{2\beta}}\right)$ $\frac{1}{2}\left(1-\sqrt{5+\frac{C}{2\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1-\sqrt{5+\frac{ch^4}{2\beta}}\right)$
$\bar{M} \neq M$	$\lambda \in [0,1]$ $\frac{1}{2}\left(\sqrt{5+\frac{ch^4}{2\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1+\sqrt{5+\frac{C}{2\beta}}\right)$ $-\frac{1}{2}\left(\sqrt{5+\frac{C}{2\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1-\sqrt{5+\frac{ch^4}{2\beta}}\right)$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Solving reduced saddle-point problem: MINRES/SYMMBK

$$\begin{bmatrix} \bar{M} & K^T \\ K & \frac{1}{2\beta}M \end{bmatrix} \begin{bmatrix} y \\ l \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

Use MINRES/SYMMBK with preconditioner of the form

$$\begin{bmatrix} \tilde{K}M^{-1}\tilde{K} & 0 \\ 0 & \frac{1}{2\beta}M \end{bmatrix}$$

$\bar{M} = M$	$-\frac{1}{2}\left(1+\sqrt{5}\right) \le \lambda \le -1$ $\frac{1}{2}\left(\frac{ch^4}{2\beta} - 1 + \sqrt{4 + \left(1 + \frac{ch^4}{2\beta}\right)}\right) \le \lambda \le \frac{1}{2}\left(\frac{C}{2\beta} - 1 + \sqrt{4 + \left(1 + \frac{C}{2\beta}\right)}\right)$
$\bar{M} \ne M$	$-\frac{1}{2}\left(1+\sqrt{5}\right) \le \lambda \le -1$ $\frac{1}{2}\left(\frac{\bar{c}h^4}{2\beta} - 1 + \sqrt{4 + \left(1 + \frac{\bar{c}h^4}{2\beta}\right)}\right) \le \lambda \le \frac{1}{2}\left(\frac{\bar{C}}{2\beta} - 1 + \sqrt{4 + \left(1 + \frac{\bar{C}}{2\beta}\right)}\right)$ $\lambda = -\frac{1}{2}\left(1+\sqrt{5}\right)$ $\lambda = \frac{1}{2}\left(-1+\sqrt{5}\right)$
	$c \le \bar{c}$ and $\bar{C} \le C$

	Target 1		Target 2	
h^{-1}	SYMMBK(orig)	SYMMBK(red)	SYMMBK(orig)	SYMMBK(red)
2^4	0.20 (25)	0.13 (26)	0.57 (75)	0.08 (14)
2^5	0.50 (25)	0.37 (27)	1.47 (78)	0.20 (14)
2^6	1.79 (25)	1.14 (23)	5.00 (75)	0.70 (14)
2^7	8.28 (25)	5.42 (23)	24.9 (80)	3.57 (15)
2^8	36.2 (25)	21.4 (21)	119 (84)	15.3 (15)
2^9	145 (25)	97.1 (24)	473 (86)	57.9 (14)

$$\beta = 5 \times 10^{-5}$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Conclusions and Future Work

- PDE-constrained problems difficult to solve
- Avoid accurate solves with discretized PDE
- Use block structure
- 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 `HSL2011`
- ‘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
- ‘PDE-constrained optimization and constraint preconditioners’, Thorne, RAL TR-2010-016