PDE-constrained optimisation: from linear to nonlinear constraints

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)

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

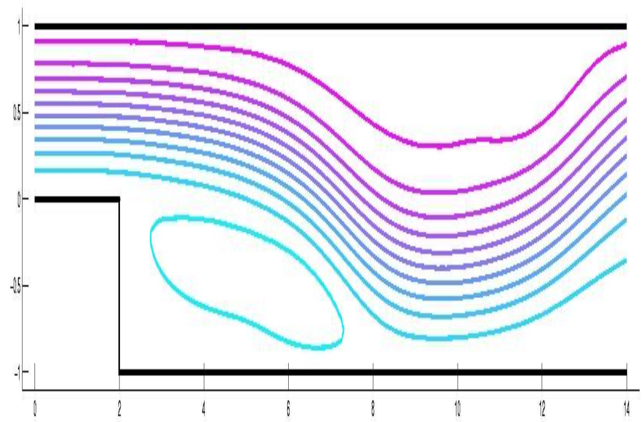
- Given u and boundary conditions g, calculate y, where

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

  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

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

Here

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

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Distributed control

Discretize:

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

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

subject to

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

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

Distributed control

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

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

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

Optimality conditions:

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

Distributed control

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

Optimality conditions:

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

Simple reduction:

$$u = \frac{1}{2\beta} y$$

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

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 original 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

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

[Rusten and Winther 1992]

Spectral properties of original linear system

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

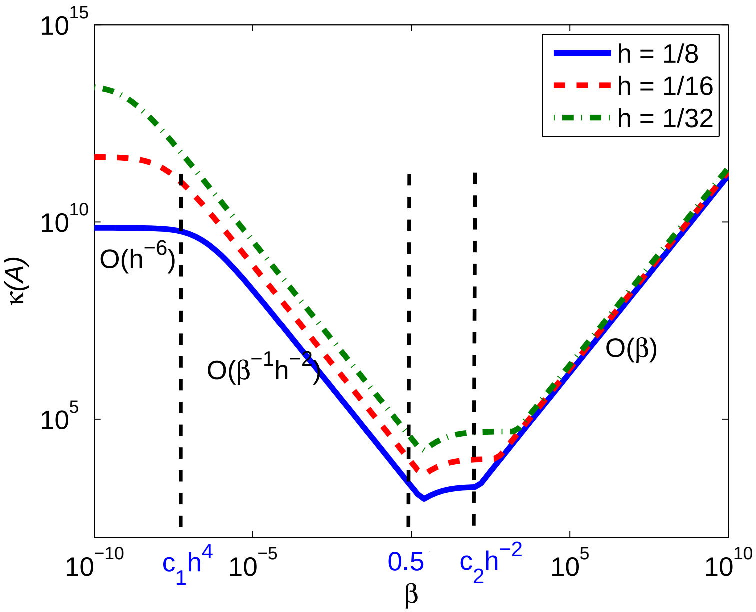
If A is symmetric and positive semi-definite, then $\lambda(H) \in I^- \cup I^+$, where

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

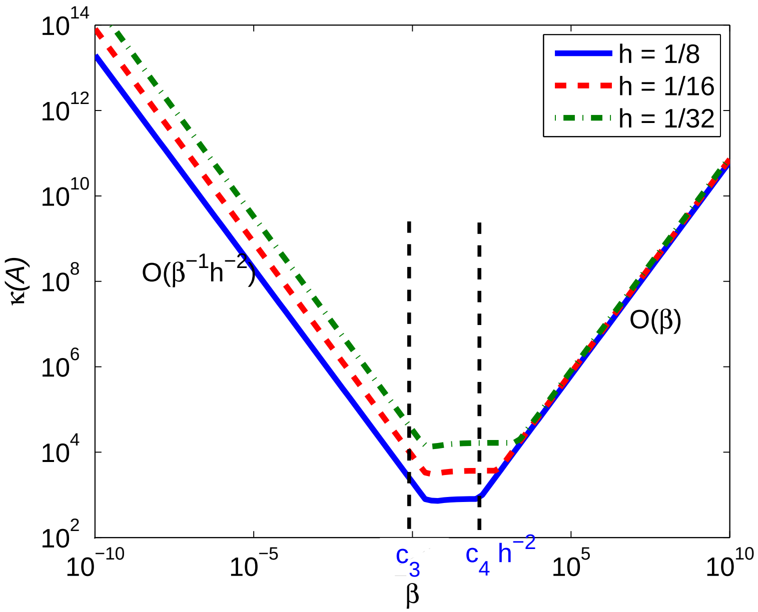
$l(A,B)$ defined in Dollar 2009 (revised)

Spectral properties of original linear system

$\hat{\Omega} = \Omega$



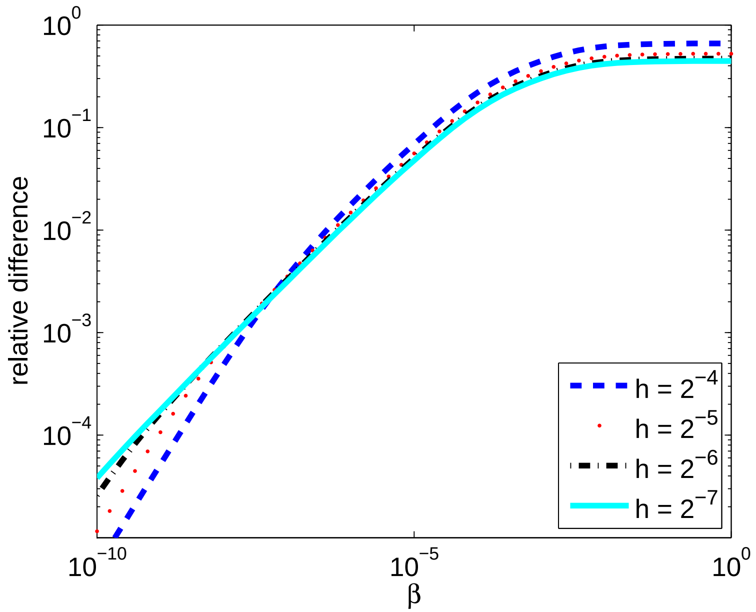
$\hat{\Omega} \neq \Omega$



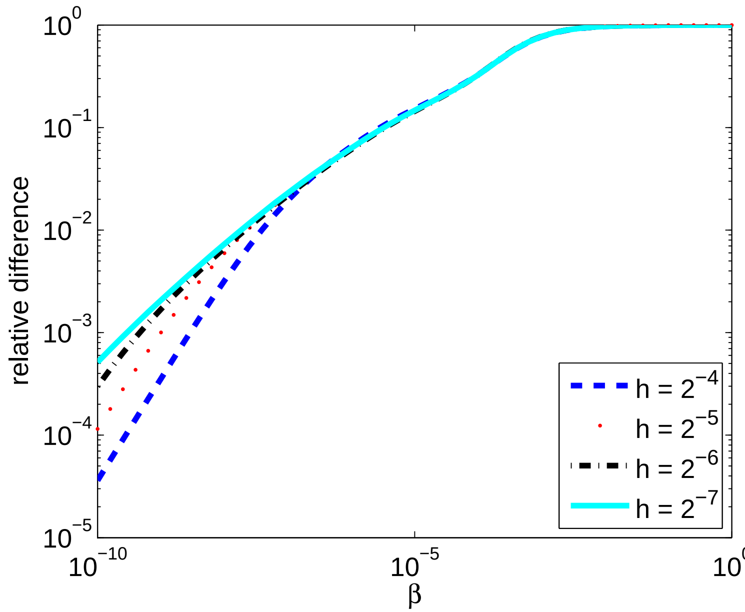
$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{y}(x_1,x_2)\|_{\hat{\Omega}_1}$	$\hat{y}(x_1,x_2)\|_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0,\frac{1}{2}\right]^2$	$\Omega/\hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x,y) : (x-\frac{5}{8})^2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

Spectral properties of original linear system

$\hat{\Omega} = \Omega$



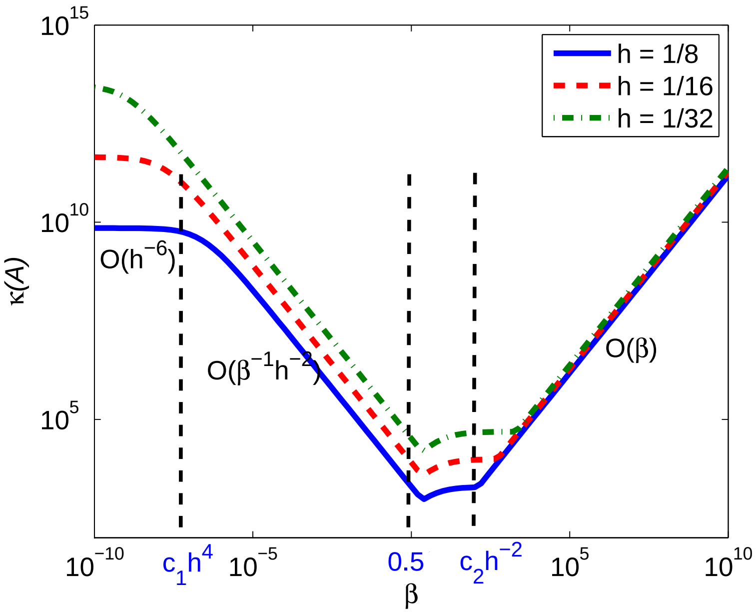
$\hat{\Omega} \neq \Omega$



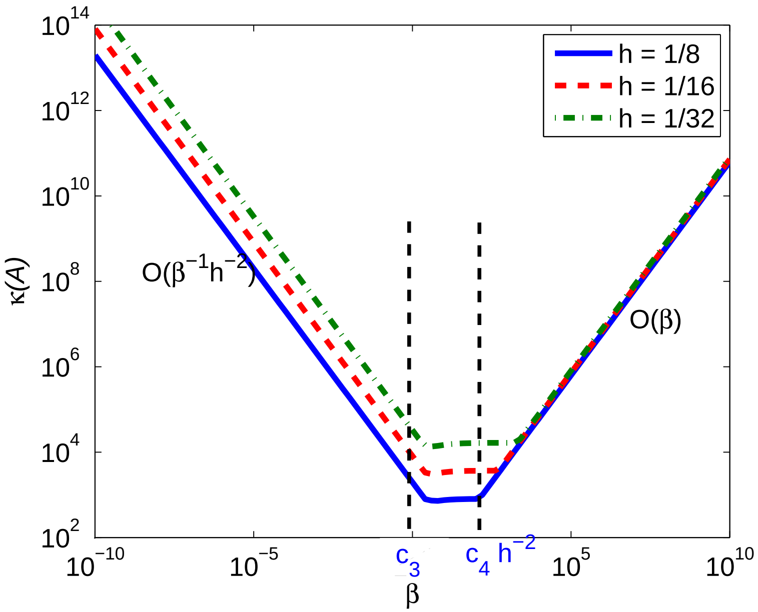
$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{y}(x_1, x_2)\vert_{\hat{\Omega}_1}$	$\hat{y}(x_1, x_2)\vert_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega / \hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x, y)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

Spectral properties of original linear system

$\hat{\Omega} = \Omega$



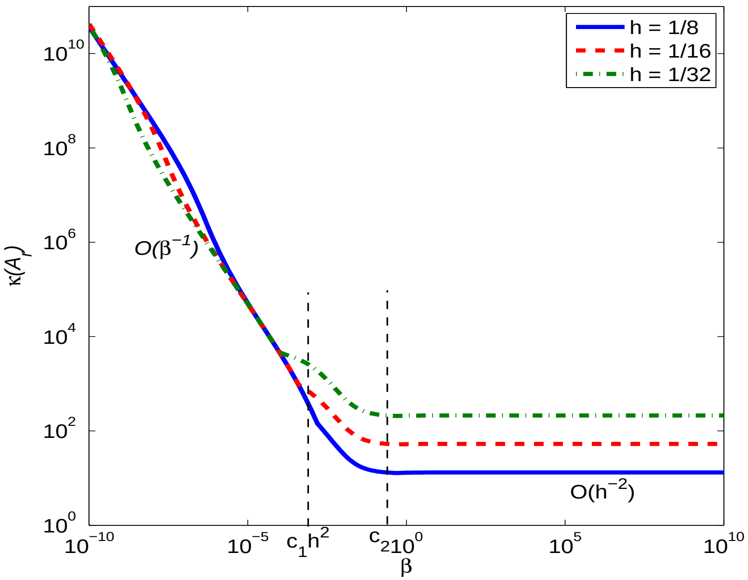
$\hat{\Omega} \neq \Omega$



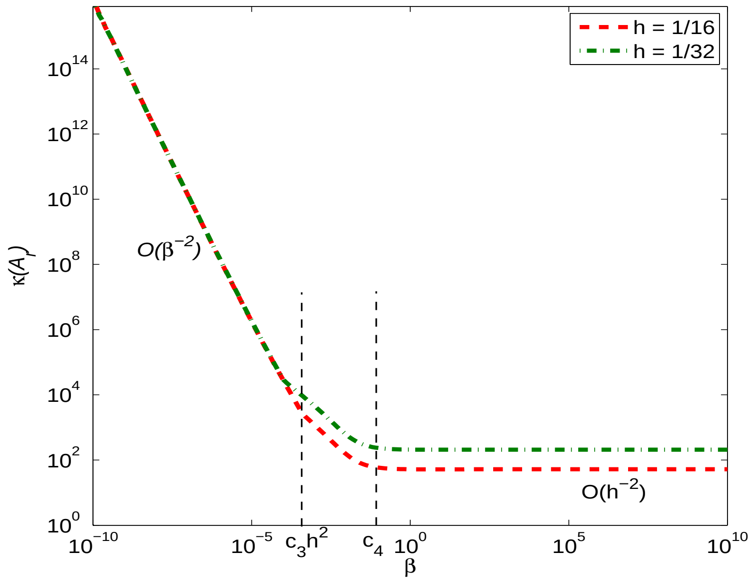
$\hat{\Omega}$	$\hat{\Omega}_1$	$\hat{\Omega}_2$	$\hat{y}(x_1,x_2)\vert_{\hat{\Omega}_1}$	$\hat{y}(x_1,x_2)\vert_{\hat{\Omega}_2}$
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left[0, \frac{1}{2}\right]^2$	$\Omega / \hat{\Omega}_1$	$(2x-1)^2\,(2y-1)^2$	0
$\hat{\Omega}_1 \cup \hat{\Omega}_2$	$\left\{(x,y) :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

Spectral properties of reduced 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

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

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

Solve

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

or

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

Possible preconditioners

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

Block diagonal preconditioner (talk by Andy Wathen)

$$\begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix}$$

Constraint preconditioner (this talk)

$$\begin{bmatrix} G & B^T \\ B & -C \end{bmatrix}$$

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

Science & Technology Facilities Council
Rutherford Appleton Laboratory

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 $n - m$ eigenvalues 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.

Preconditioner

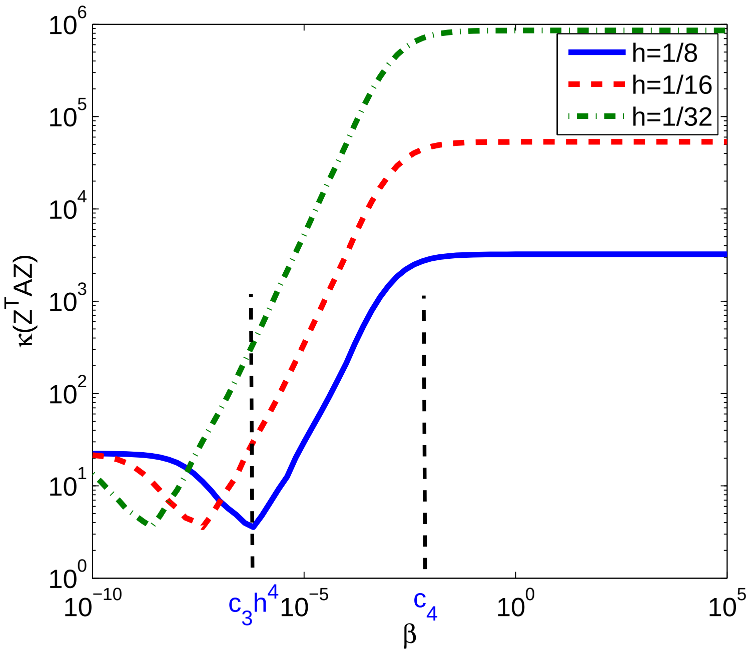
$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^T \\ -M & K & 0 \end{bmatrix} \quad Z = \begin{bmatrix} M^{-1}K \\ I \end{bmatrix}$$

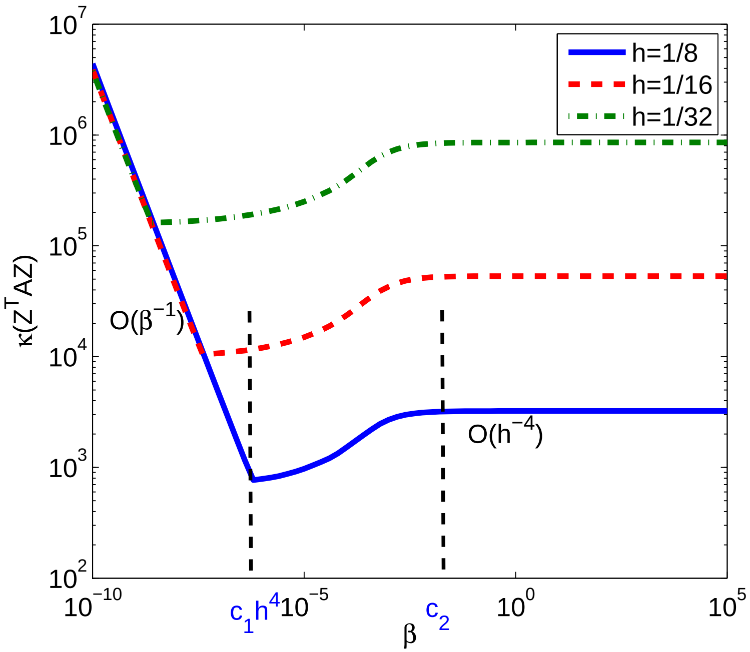
$$Z^T A Z = 2\beta K^T M^{-1} K + \bar{M}$$

Preconditioner

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

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





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

Constraint preconditioners cont.

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

Constraint preconditioners cont.

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

Theorem: If $A, G \in \mathbb{R}^{n \times n}$ are symmetric, $B \in \mathbb{R}^{m \times n}$ has full row rank and C is symmetric and positive definite, then $\mathcal{P}^{-1}\mathcal{A}$ has

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

$$\left(A + B^T C^{-1} B\right) x = \lambda \left(G + B^T C^{-1} B\right) x.$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Preconditioner

$$\mathcal{A}_r = \begin{bmatrix} \bar{M} & K^T \\ K & -\frac{1}{2\beta} M \end{bmatrix}, \quad A + B^T C^{-1} B = 2\beta K^T M^{-1} K + \bar{M}$$

Preconditioner

$$\mathcal{A}_r = \begin{bmatrix} \bar{M} & K^T \\ K & -\frac{1}{2\beta}M \end{bmatrix}, \quad A + B^T C^{-1} B = 2\beta K^T M^{-1} K + \bar{M}$$

$$\mathcal{P}_r = \begin{bmatrix} G & K^T \\ K & -\frac{1}{2\beta}M \end{bmatrix}? \quad G + B^T C^{-1} B = 2\beta K^T M^{-1} K \quad \Rightarrow \quad G = 0$$

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

Preconditioner

$$\mathcal{A}_r = \begin{bmatrix} \bar{M} & K^T \\ K & -\frac{1}{2\beta}M \end{bmatrix}, \quad A + B^T C^{-1} B = 2\beta K^T M^{-1} K + \bar{M}$$

$$\begin{aligned} \mathcal{P}_r &= \begin{bmatrix} I & -K \\ 0 & \frac{1}{2\beta}M \end{bmatrix} \begin{bmatrix} 2\beta \tilde{K}^T M^{-1} \tilde{K} & 0 \\ 0 & -2\beta M^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -K^T & \frac{1}{2\beta}M \end{bmatrix} \\ &= \begin{bmatrix} 2\beta \tilde{K}^T M^{-1} \tilde{K} - 2\beta K^T M^{-1} K & K^T \\ K & -\frac{1}{2\beta}M \end{bmatrix}, \end{aligned}$$

where $\tilde{K}$ is an approximation to K.

$$G + B^T C^{-1} B = 2\beta \tilde{K}^T M^{-1} \tilde{K}$$

If $\tilde{K} = K$,

$\bar{M} = M$	$\bar{M} \neq M$
$1 + \frac{ch^4}{2\beta} \leq \lambda \leq 1 + \frac{C}{2\beta}$ $c \leq \bar{c}$ 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 $\beta = 5 \times 10^{-5}$:

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

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

- Solves with M : Direct method (`HSL_MA57`) or 20/30(30/50) 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, ifort compiler
- Hardware: Dell Precision T340, single Core2 Quad Q9550 processor (2.83GHz, 1333MHz FSB, 12MB L2 Cache), 4GB RAM

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Numerical Example

2D

N	n	Direct	PPCG(direct)	PPCG(approx)
8	147	0.004	0.00 (8)	0.004 (9)
16	675	0.004	0.004 (8)	0.004 (9)
32	2883	0.02	0.02 (8)	0.03 (9)
64	11907	0.11	0.06 (8)	0.10 (8)
128	48487	1.11	0.34 (7)	0.42 (8)
256	195075	6.24	2.32 (6)	2.04 (8)
512	783363	39.1	10.8 (6)	9.35 (8)

3D

N	n	Direct	PPCG(direct)	PPCG(approx)
4	81	0.004	0.004 (5)	0.000 (5)
8	1029	0.03	0.01 (7)	0.03 (7)
16	10125	0.75	0.21 (7)	0.34 (7)
32	89373	36.1	5.14 (7)	4.94 (7)
64	750141	1000+	192 (5)	41.9 (6)

2D reduced

N	n	Direct	PPCG(direct)	PPCG(approx)
8	98	0.000	0.00 (7)	0.00 (8)
16	450	0.004	0.004 (7)	0.008 (8)
32	1922	0.02	0.02 (7)	0.03 (8)
64	7938	0.06	0.06 (7)	0.11 (8)
128	32325	0.33	0.33 (7)	0.48 (8)
256	130050	1.97	2.36 (7)	2.42 (8)
512	522242	10.9	11.0 (7)	12.0 (9)

3D reduced

N	n	Direct	PPCG(direct)	PPCG(approx)
4	54	0.00	0.01 (4)	0.00 (4)
8	686	0.01	0.01 (7)	0.03 (7)
16	6750	0.34	0.22 (7)	0.45 (8)
32	59582	12.0	4.97 (7)	6.72 (8)
64	500094	797	187 (5)	51.3 (6)

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

Numerical Example

2D

N	n	Direct	PPCG(direct)	PPCG(approx)
8	147	0.00	0.00 (4)	0.00 (4)
16	675	0.01	0.00 (4)	0.00 (4)
32	2883	0.04	0.01 (4)	0.02 (4)
64	11907	0.24	0.06 (4)	0.07 (5)
128	48487	1.74	0.30 (5)	0.30 (5)
256	195075	11.0	2.16 (5)	1.45 (5)
512	783363	93.5	10.1 (5)	6.50 (5)

3D

N	n	Direct	PPCG(direct)	PPCG(approx)
4	81	0.00	0.00 (3)	0.00 (3)
8	1029	0.03	0.01 (4)	0.02 (4)
16	10125	0.84	0.21 (5)	0.30 (5)
32	89373	41.0	4.79 (5)	4.46 (5)
64	750141	1000+	187 (5)	45.6 (5)

2D reduced

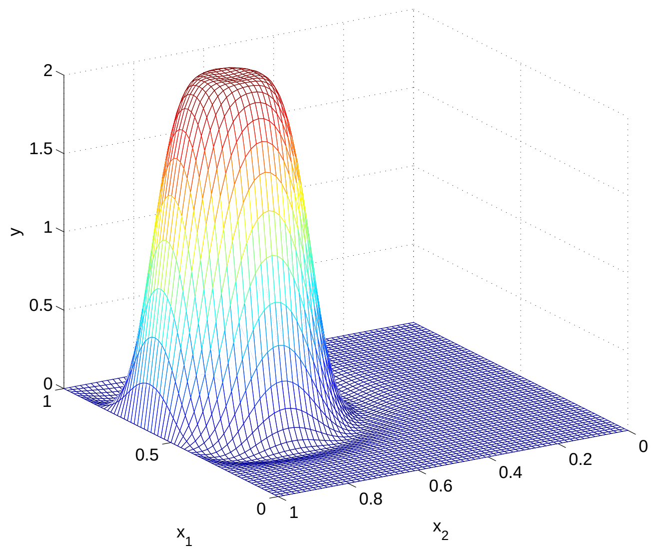
N	n	Direct	PPCG(direct)	PPCG(approx)
8	98	0.00	0.00 (4)	0.00 (4)
16	450	0.00	0.00 (4)	0.004 (4)
32	1922	0.02	0.01 (4)	0.02 (4)
64	7938	0.14	0.06 (4)	0.07 (5)
128	32325	0.79	0.30 (4)	0.41 (5)
256	130050	4.10	2.16 (4)	1.83 (5)
512	522242	24.6	10.1 (5)	7.86 (5)

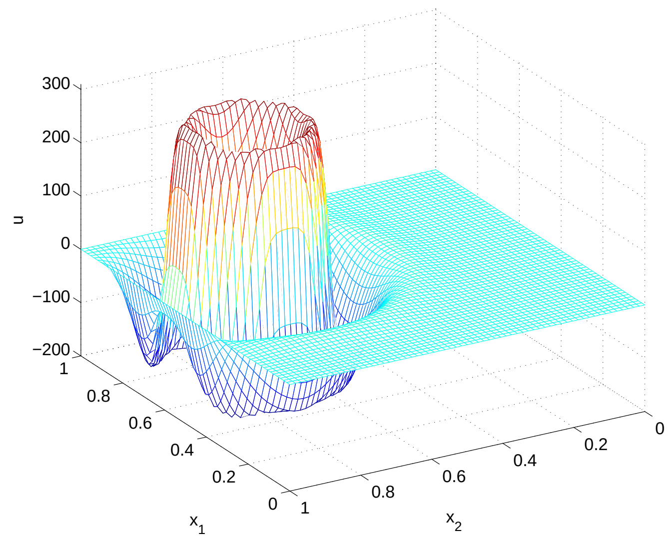
3D reduced

N	n	Direct	PPCG(direct)	PPCG(approx)
4	54	0.00	0.00 (3)	0.00 (3)
8	686	0.01	0.004 (4)	0.02 (4)
16	6750	0.41	0.21 (4)	0.30 (4)
32	59582	20.9	4.83 (4)	4.64 (4)
64	500094	1000+	192 (5)	52.1 (5)

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

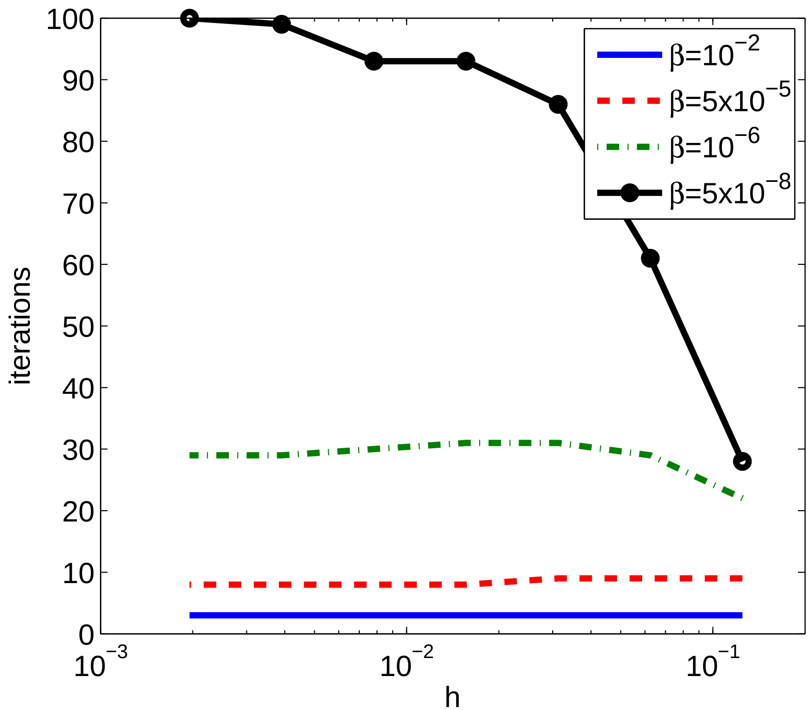
Numerical Example

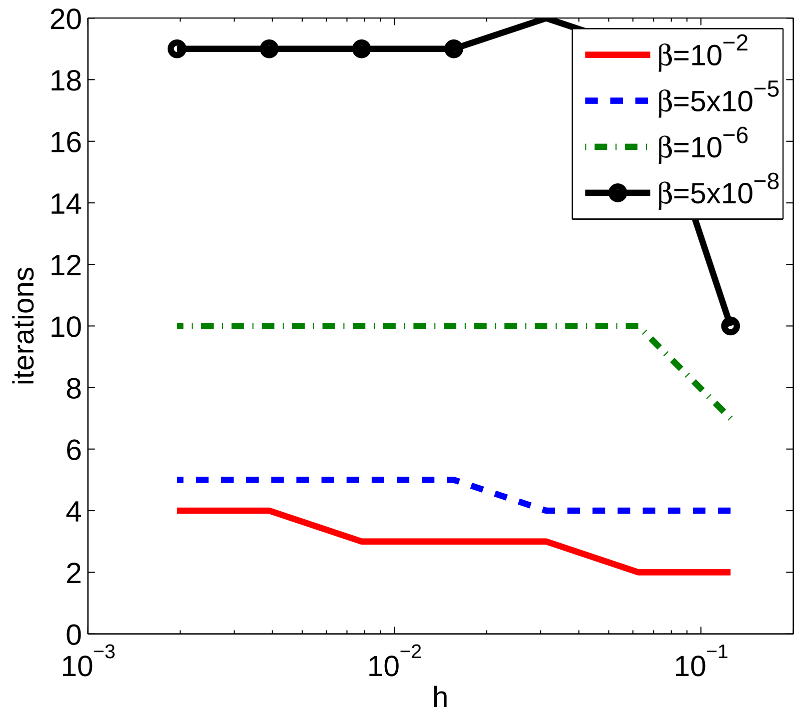




Behaviour of preconditioner with β

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





Distributed control with nonlinear PDEs

$$\min_{y,u} \frac{1}{2} \|y - \widehat{y}\|_2^2 + \beta \|u\|_2^2$$

subject to

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

Optimality conditions:

$$\begin{aligned} My + J(y)^T l &= b, \\ 2\beta Mu - Ml &= 0, \\ F(y) - Mu &= d. \end{aligned}$$

Trust-funnel method (Gould and Toint)

$$\min_x f(x) \quad \text{subject to} \quad c(x) = 0$$

Attempts to consider the objective function and constraints as independently as possible

Trust-funnel method (Gould and Toint)

$$\min_x f(x) \quad \text{subject to} \quad c(x) = 0$$

Attempts to consider the objective function and constraints as independently as possible

$$\text{Find } n \text{ to reduce } \quad \|c_k + J_k n\| \quad \text{subject to} \quad \|n\| \leq \Delta_1$$

$$\text{Find } l \text{ to reduce } \quad \left\|g_k + J_k^T l\right\|$$

$$\text{Find } t \text{ to reduce } \quad g_k^T t + \frac{1}{2} t^T H_k t \quad \text{subject to} \quad J_k t = 0 \quad \text{and} \quad \|t\| \leq \Delta_2$$

$$x_{k+1} = x_k + n + t$$

Trust-funnel method (Gould and Toint)

$$\min_x f(x) \quad \text{subject to} \quad c(x) = 0$$

Attempts to consider the objective function and constraints as independently as possible

$$\text{Find } n \text{ to reduce} \quad \|c_k + J_k n\| \quad \text{subject to} \quad \|n\| \le \Delta_1$$

$$\text{Find } l \text{ to reduce} \quad \left\|g_k + J_k^T l\right\|$$

$$\text{Find } t \text{ to reduce} \quad g_k^T t + \frac{1}{2} t^T H_k t \quad \text{subject to} \quad J_k t = 0 \quad \text{and} \quad \|t\| \le \Delta_2$$

$$x_{k+1} = x_k + n + t$$

- Adjust Δ_1 and Δ_2 for convergence
- Only require matrix-vector multiplications (preconditioning?)
- Alternative matrix-free method by Curtis, Nocedal and Wächter

Reduce $g_k^T t + \frac{1}{2} t^T H_k t$ subject to $J_k t = 0$ and $\|t\| \leq \Delta_2,$

where

$$\begin{aligned} g_k &= \nabla f(x_k) + H_k n_k, \\ H_k &= \nabla^2 f(x_k) + \sum_{i=1}^{m} [l_{k-1}]_i C_{ik}, \\ C_{ik} &= C_{ik}^T \approx \nabla_{xx} c_i(x_k) \end{aligned}$$

$\min_x f(x)$ **subject to** $c(x) = 0$

Reduce $g_k^T t + \frac{1}{2} t^T H_k t$ subject to $J_k t = 0$ and $\|t\| \leq \Delta_2$,

where

$$\begin{aligned} g_k &= \nabla f(x_k) + H_k n_k, \\ H_k &= \nabla^2 f(x_k) + \sum_{i=1}^{m} [l_{k-1}]_i \, C_{ik}, \\ C_{ik} &= C_{ik}^T \approx \nabla_{xx} c_i(x_k) \end{aligned}$$

Apply PPCG to

$$\begin{bmatrix} H_k & J_k^T \\ J_k & 0 \end{bmatrix} \begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} g_k \\ 0 \end{bmatrix}$$

Initialise $t = 0$. Iterate until convergence or $\|t\| \geq \Delta_2$. If $\|t\| \geq \Delta_2$, back-track to boundary.

Distributed control with nonlinear PDEs

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

subject to

$$\begin{aligned} -\nabla \cdot \left[\left(1+y^2\right)\nabla y\right] &= u \text{ in } \Omega \\ y &= \widehat{y} \text{ on } \delta\Omega \end{aligned}$$

Distributed control with nonlinear PDEs

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

subject to

$$\begin{aligned} -\nabla \cdot \left[\left(1 + y^2\right) \nabla y \right] &= u \text{ in } \Omega \\ y &= \widehat{y} \text{ on } \delta\Omega \end{aligned}$$

$$\left[\begin{array}{cc|c} 2\beta M & 0 & -M \\ 0 & M + \sum_{i=1}^{m} \left[l_{k-1} \right]_i \nabla^2 F_j(y_k) & J(y_k)^T \\ \hline -M & J(y_k) & 0 \end{array} \right]$$

Distributed control with nonlinear PDEs

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

subject to

$$\begin{aligned} -\nabla \cdot \left[\left(1+y^2\right)\nabla y\right] &= u \text{ in } \Omega \\ y &= \widehat{y} \text{ on } \delta\Omega \end{aligned}$$

$$\left[\begin{array}{cc|c} 2\beta M & 0 & -M \\ 0 & M + \sum_{i=1}^{m} \left[l_{k-1}\right]_i \nabla^2 F_j(y_k) & K^T + L(y_k)^T \\ \hline -M & K + L(y_k) & 0 \end{array}\right]$$

Distributed control with nonlinear PDEs

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

subject to

$$\begin{aligned} -\nabla \cdot \left[\left(1+y^2\right)\nabla y\right] &= u \text{ in } \Omega \\ y &= \widehat{y} \text{ on } \delta\Omega \end{aligned}$$

$$\left[\begin{array}{cc|c} 2\beta M & 0 & -M \\ 0 & M + \sum_{i=1}^{m} [l_{k-1}]_i \nabla^2 F_j(y_k) & K^T + L(y_k)^T \\ \hline -M & K + L(y_k) & 0 \end{array}\right]$$

$$P_1 = \left[\begin{array}{cc|c} I & 0 & -M \\ 0 & I & K^T + L(y_k)^T \\ \hline -M & K + L(y_k) & 0 \end{array}\right], P_2 = \left[\begin{array}{cc|c} 0 & 0 & -M \\ 0 & 2\beta K^T M^{-1} K & K^T + L(y_k)^T \\ \hline -M & K + L(y_k) & 0 \end{array}\right]$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Preliminary results

N		T-F iterations	PPCG calls	Total PPCG its	Max PPCG its	Average PPCG its
2^3	P_1	19	14	490	$50^*(3)$	35
	P_2	5	3	34	12	11
2^4	P_1	24	15	1388	$226^*(5)$	93
	P_2	5	3	39	20	13

In progress - Python package that will discretise a given problem and use trust-funnel method to solve the problems (uses FEniCS). Idea: plug in different linear solvers/preconditioners

Conclusions and Future Work

- PDE-constrained problems difficult to solve
- Avoid accurate solves with discretized PDE
- Use block structure
- Mesh size independent convergence
- Nonlinear PDEs - sequence of more challenging saddle-point systems
- 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`
- ‘Nonlinear programming without a penalty function or filter’ Gould, Toint, Math. Prog 2010
- ‘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