Preconditioners for PDE-constrained optimisation problems

Sue Thorne

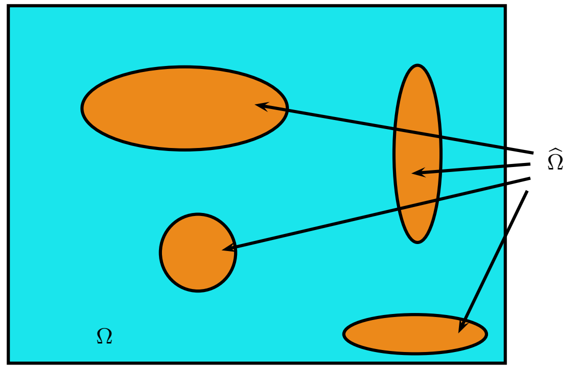
STFC Rutherford Appleton Laboratory

PDE-constrained optimisation

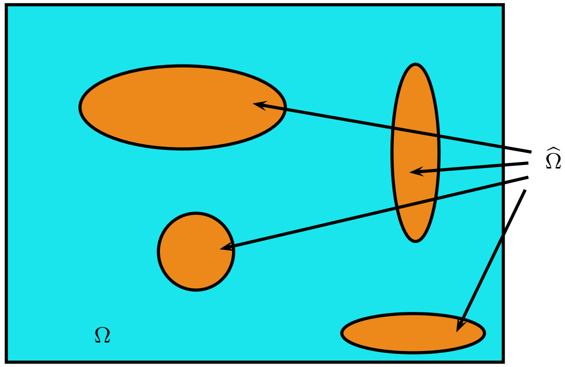
Different target temperatures

Calculate epicentre of earthquake

Distributed control



Distributed control



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

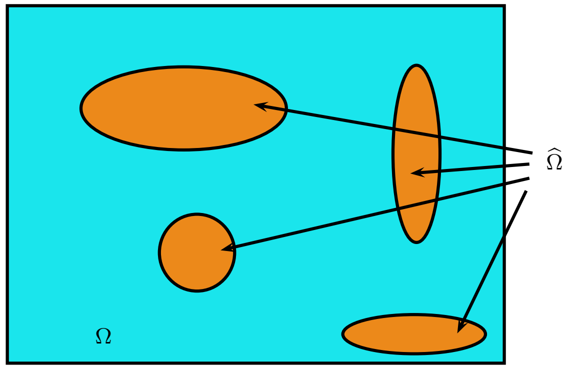
subject to

$$\begin{aligned} \mathcal{L}y &= u \text{ in } \Omega \\ \alpha_1 y + \alpha_2 \frac{\partial y}{\partial n} &= g \text{ on } \partial\Omega \end{aligned}$$

Here

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

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

$$\begin{aligned} \mathcal{L} y &= u \text{ in } \Omega \\ \alpha_1 y + \alpha_2 \frac{\partial y}{\partial n} &= g \text{ on } \partial\Omega \end{aligned}$$

Here

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

Distributed control

Discretize:

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

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

subject to

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

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Distributed control

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

subject to

$$Hy - Mu \quad = \quad d$$

where M is the mass matrix, H is matrix associated with $\mathcal{L}_{\mathtt{h}}$, $\widehat{M} = WMW$, $W = \mathsf{diag}(\omega_i)$,
$b = \widehat{M}\widehat{\mathtt{y}}_{\mathtt{h}}$ and $c = \widehat{\mathtt{y}}_{\mathtt{h}}^* \widehat{M}\widehat{\mathtt{y}}_{\mathtt{h}}$
H may be complex and indefinite but is always symmetric

Distributed control

$$\min_{y,u} \frac{1}{2} y^* \widehat{M} y - y^* b + c + \beta u^* M u + l^* \left(H y - M u - d \right)$$

Optimality conditions:

$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \widehat{M} & H^* \\ -M & H & 0 \end{bmatrix} \begin{bmatrix} u \\ y \\ l \end{bmatrix} = \begin{bmatrix} 0 \\ b \\ d \end{bmatrix}$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Distributed control

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

Optimality conditions:

$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \widehat{M} & H^* \\ -M & H & 0 \end{bmatrix} \begin{bmatrix} u \\ y \\ l \end{bmatrix} = \begin{bmatrix} 0 \\ b \\ d \end{bmatrix}$$

Simple reduction:

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

$$\begin{bmatrix} \widehat{M} & H^* \\ H & -\frac{1}{2\beta} M \end{bmatrix} \begin{bmatrix} y \\ l \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

Constraint preconditioners

$$\mathcal{A} = \begin{bmatrix} A & B^* \\ B & -C \end{bmatrix}, \mathcal{P}_c = \begin{bmatrix} G & B^* \\ B & -C \end{bmatrix}$$

$A = A^* \in \mathbb{C}^{n \times n}, C = C^* \in \mathbb{C}^{m \times m}$, rank$(B) = m$

Constraint preconditioners

$$\mathcal{A} = \begin{bmatrix} A & B^* \\ B & -C \end{bmatrix}, \mathcal{P}_c = \begin{bmatrix} G & B^* \\ B & -C \end{bmatrix}$$

$$A = A^* \in \mathbb{C}^{n\times n}, C = C^* \in \mathbb{C}^{m\times m}, \text{rank}(B) = m$$

Constraint preconditioner:

If $C = 0$, $\mathcal{P}_c^{-1}\mathcal{A}$ has

- $2m$ eigenvalues at 1
- remaining eigenvalues satisfy $Z^*AZx = \lambda Z^*GZx$ [Real case: Keller, Gould, Wathen (2000)]

Constraint preconditioners

$$\mathcal{A} = \begin{bmatrix} A & B^* \\ B & -C \end{bmatrix}, \mathcal{P}_c = \begin{bmatrix} G & B^* \\ B & -C \end{bmatrix}$$

$$A = A^* \in \mathbb{C}^{n\times n}, C = C^* \in \mathbb{C}^{m\times m}, \text{rank}(B) = m$$

Constraint preconditioner:

If $C = 0$, $\mathcal{P}_c^{-1}\mathcal{A}$ has

- $2m$ eigenvalues at 1
- remaining eigenvalues satisfy $Z^*AZx = \lambda Z^*GZx$ [Real case: Keller, Gould, Wathen (2000)]

If C is nonsingular, $\mathcal{P}_c^{-1}\mathcal{A}$ has

- m eigenvalues at 1
- remaining eigenvalues satisfy $\left(A + B^*C^{-1}B\right)x = \lambda\left(G + B^*C^{-1}B\right)x$ [Real case: Gould (1999)]

Projected Preconditioned CG Method

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

If $C = 0$, write $x = Yx_y + Zx_z$, where columns of 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 $Z^T GZ$.

Projected Preconditioned CG Method

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

If $C = 0$, write $x = Yx_y + Zx_z$, where columns of 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 $Z^T GZ$.
If C is nonsingular, $w = C^{-1} (Bx - d)$ and

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

If $A + B^T C^{-1} B$ is SPD, use PCG with preconditioner $G + B^T C^{-1} B$.

Projected Preconditioned CG Method

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

If $C = 0$, write $x = Yx_y + Zx_z$, where columns of Z span nullspace of B and $[Y, Z]$ spans $\mathbb{R}^n$

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

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

If C is nonsingular, $w = C^{-1} (Bx - d)$ and

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

If $A + B^T C^{-1} B$ is SPD, use PCG with preconditioner $G + B^T C^{-1} B$.

Use substitutions to remove $(Z, Y)/C^{-1}$ to obtain projected PCG (PPCG): require preconditioner

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

Can extend to complex case.

Example Problem 1

Forward problem:

$$\begin{aligned} -\nabla^2 y &= u \text{ in } \Omega = [0,1]^d, d = 2,3 \\ y &= \widehat{y} \text{ on } \partial\Omega \end{aligned}$$

where

$$\begin{aligned} \widehat{\Omega} &= \widehat{\Omega}_1 \cup \widehat{\Omega}_2 \\ \widehat{\Omega}_1 &= \begin{cases} \left\{(x_1, x_2) \,|\, \left(x_1 - \frac{5}{8}\right)^2 + \left(x_2 - \frac{3}{4}\right)^2 \le \frac{1}{25}\right\}, & d = 2, \\ \left\{(x_1, x_2, x_3) \,|\, \left(x_1 - \frac{5}{8}\right)^2 + \left(x_2 - \frac{3}{4}\right)^2 + \left(x_3 - \frac{7}{10}\right)^2 \le \frac{1}{16}\right\}, & d = 3 \end{cases} \\ \widehat{\Omega}_2 &= \partial\Omega \\ \hat{y}(x) &= \begin{cases} 2, & x \in \widehat{\Omega}_1 \\ 0, & x \in \widehat{\Omega}_2 \end{cases} \end{aligned}$$

Bilinear **Q1** elements

Linear system properties

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

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

Linear system properties

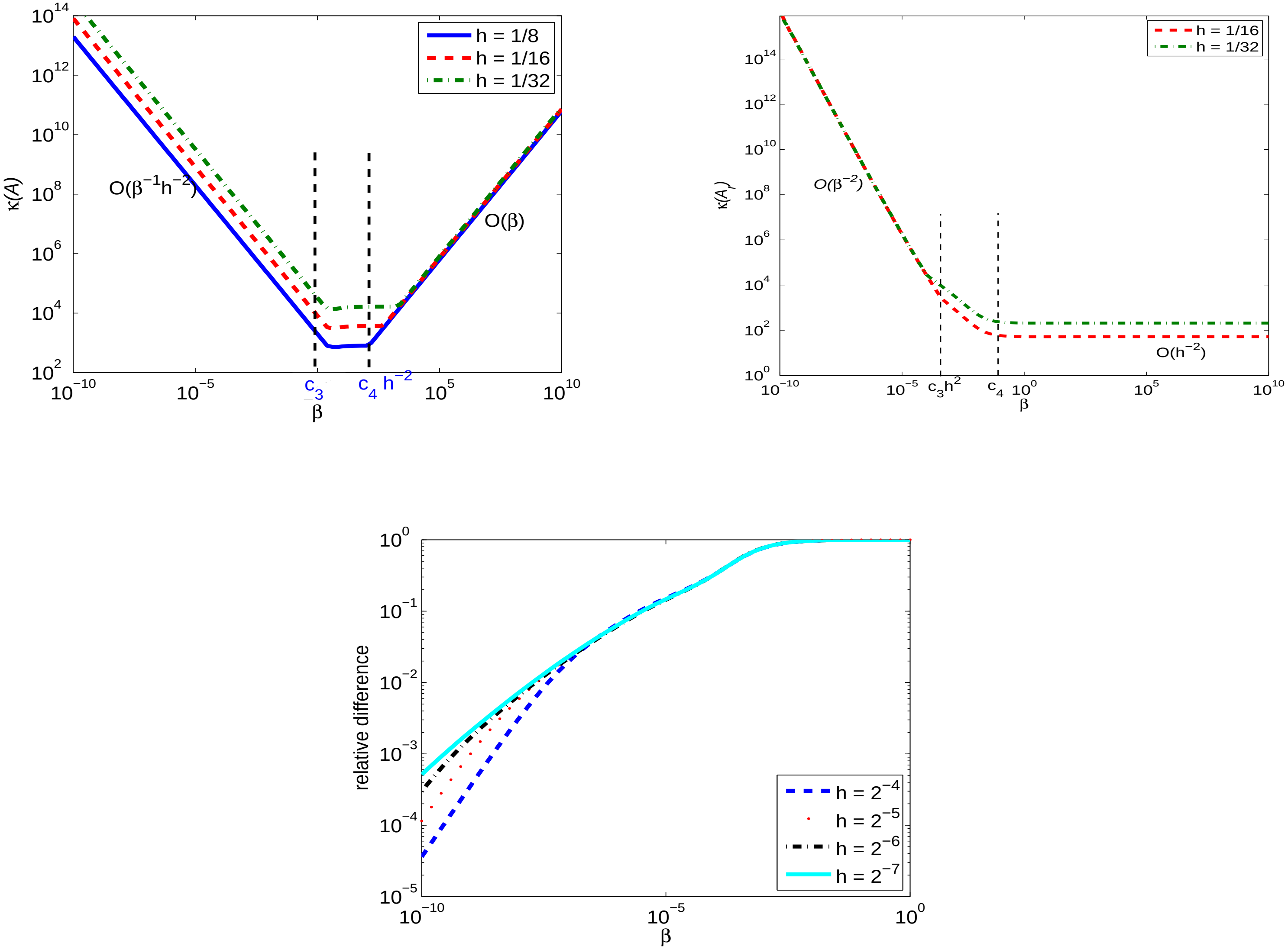
$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}$$

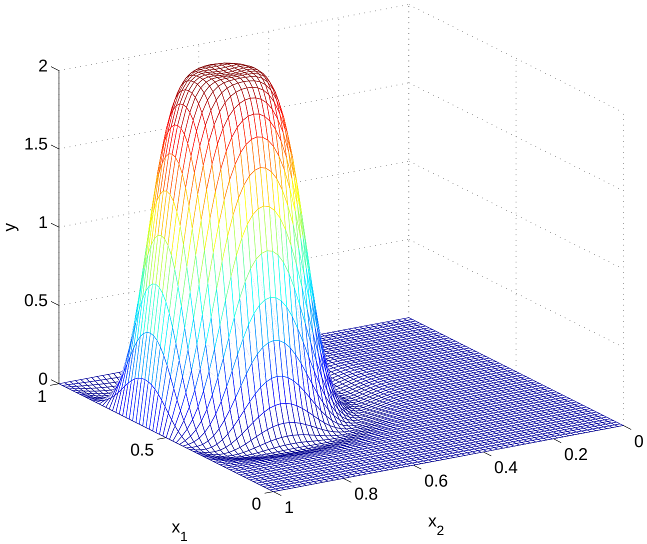
If A is symmetric and positive semi-definite, then $\lambda(\mathcal{A}) \in I^- \cup I^+$, where

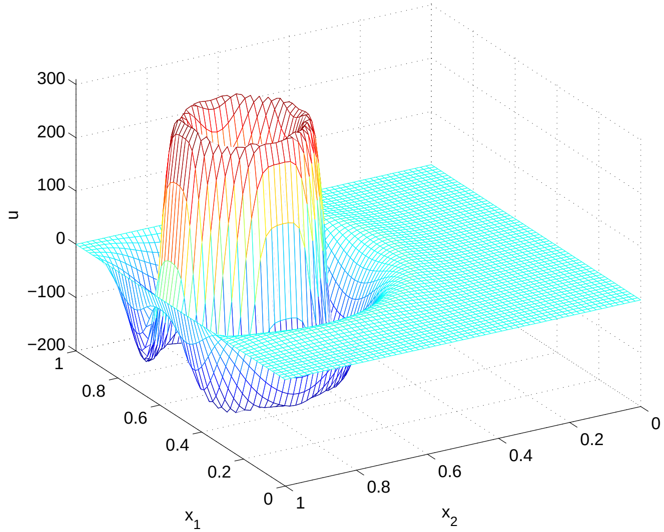
$$\begin{aligned} I^- &= \left[\frac{1}{2}\left(\lambda_{\min}(A) - \sqrt{\lambda_{\min}^2(A) + 4\|B\|^2}\right), \frac{1}{2}\left(\|A\| - \sqrt{\|A\|^2 + 4\sigma_{\min}^2(B)}\right)\right], \\ I^+ &= \left[l(A,B), \frac{1}{2}\left(\|A\| + \sqrt{\|A\|^2 + 4\|B\|^2}\right)\right], \end{aligned}$$

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

Linear system properties




2
1.5
1
0.5
0
y
1
0.5
0
x1
1
0.8
0.6
0.4
0.2
0
x2


300
200
100
0
−100
−200
u
1
0.8
0.6
0.4
0.2
0
x1
1
0.8
0.6
0.4
0.2
0
x2

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Preconditioner

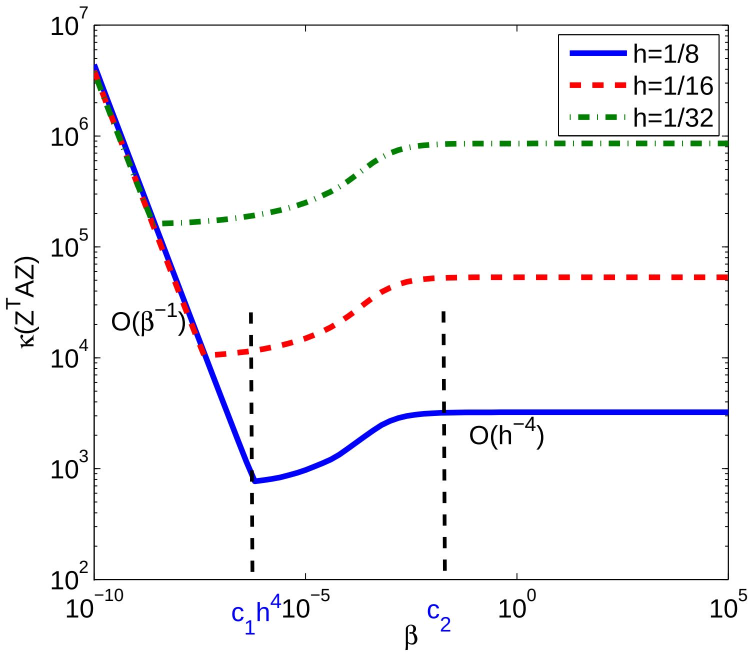
$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \widehat{M} & K^T \\ -M & K & 0 \end{bmatrix} \quad Z = \begin{bmatrix} M^{-1}K \\ I \end{bmatrix}$$

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

Preconditioner

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

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



Preconditioner

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

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

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

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

$\widehat{M} = M$	$\widehat{M} \neq M$
$1 + \frac{ch^4}{2\beta} \leq \lambda \leq 1 + \frac{C}{2\beta}$ $c \leq \bar{c}$ and $\bar{C} \leq C$	$1 + \frac{\bar{c}h^4}{2\beta} \leq \lambda \leq 1 + \frac{\bar{C}}{2\beta}$ $\lambda = 1$
Rees, Dollar, Wathen (2010)	Thorne (2011)

Biros and Ghattas (2000)

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Preconditioner

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

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

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

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

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

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

Preconditioner

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

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

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

Preconditioner

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

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

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

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

If $\tilde{K} = K$,

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

Numerical Example

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

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

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

- Solves with M : Direct method (`HSL_MA57`) or 20(30) 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: Single Quad core processor (2.83GHz, 1333MHz FSB, 12MB L2 Cache), 4GB RAM

Numerical Example

2D

h	N	Direct	PPCG(direct)	PPCG(approx)
2^{-3}	147	0.00	0.00 (4)	0.00 (4)
2^{-4}	675	0.01	0.00 (4)	0.00 (4)
2^{-5}	2883	0.04	0.01 (4)	0.02 (4)
2^{-6}	11907	0.24	0.06 (4)	0.07 (5)
2^{-7}	48487	1.74	0.30 (5)	0.30 (5)
2^{-8}	195075	11.0	2.16 (5)	1.45 (5)
2^{-9}	783363	93.5	10.1 (5)	6.50 (5)

3D

h	N	Direct	PPCG(direct)	PPCG(approx)
2^{-2}	81	0.00	0.00 (3)	0.00 (3)
2^{-3}	1029	0.03	0.01 (4)	0.02 (4)
2^{-4}	10125	0.84	0.21 (5)	0.30 (5)
2^{-5}	89373	41.0	4.79 (5)	4.46 (5)
2^{-6}	750141	1000+	187 (5)	45.6 (5)

2D reduced

h	N	Direct	PPCG(direct)	PPCG(approx)
2^{-3}	98	0.00	0.00 (4)	0.00 (4)
2^{-4}	450	0.00	0.00 (4)	0.004 (4)
2^{-5}	1922	0.02	0.01 (4)	0.02 (4)
2^{-6}	7938	0.14	0.06 (4)	0.07 (5)
2^{-7}	32325	0.79	0.30 (4)	0.41 (5)
2^{-8}	130050	4.10	2.16 (4)	1.83 (5)
2^{-9}	522242	24.6	10.1 (5)	7.86 (5)

3D reduced

h	N	Direct	PPCG(direct)	PPCG(approx)
2^{-2}	54	0.00	0.00 (3)	0.00 (3)
2^{-3}	686	0.01	0.004 (4)	0.02 (4)
2^{-4}	6750	0.41	0.21 (4)	0.30 (4)
2^{-5}	59582	20.9	4.83 (4)	4.64 (4)
2^{-6}	500094	1000+	192 (5)	52.1 (5)

Recent work

Pearson and Wathen: If $\widehat{\Omega} = \Omega$, K is symmetric and the eigenvalues of $M^{-1}K$ are real and positive, then the eigenvalues of

$$\left(M + 2\beta K M^{-1} K\right) x = \lambda (M + \sqrt{2\beta} K) M^{-1} (M + \sqrt{2\beta} K) x$$

lie in $\left[\frac{1}{2}, 1\right]$.

Simoncini: Block diagonal and indefinite (approximate constraint) preconditioners for reduced systems.

Example Problem 2

Forward problem (geophysical migration problem from seismic imaging):

$$\begin{aligned} -\nabla^2 y - k^2 y &= u \text{ in } \Omega = [0,800] \times [0,800] \times [0,160], \\ iky + \frac{\partial y}{\partial n} &= g \text{ on } \partial\Omega, \end{aligned}$$

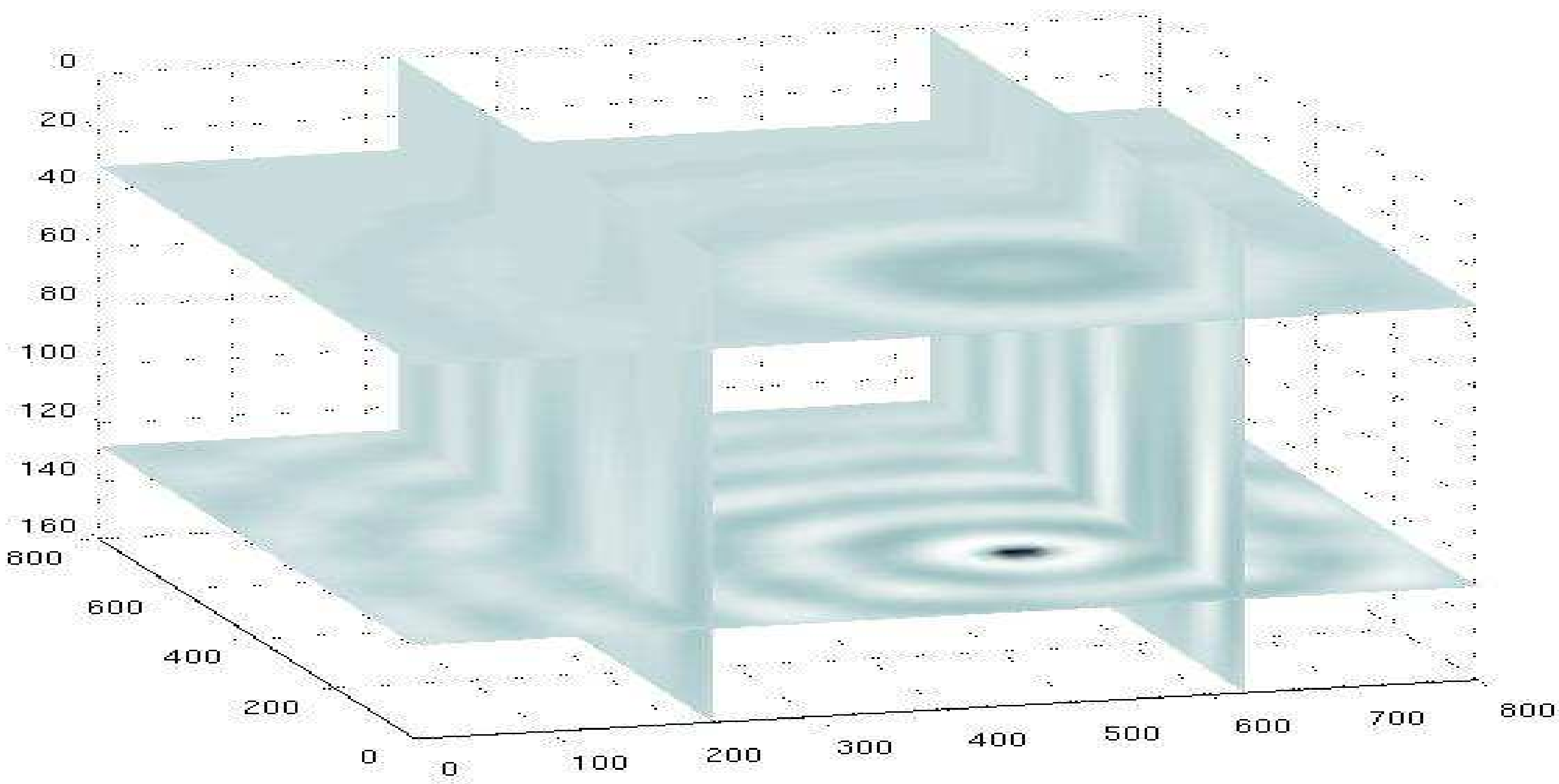
where

$$\begin{aligned} k(x_1, x_2) &= \begin{cases} 1.2k_0, & x_3 < 30 + 0.01x_1 + 0.005x_2, \\ 1.5k_0, & x_3 > 80 + 0.005x_1 + 0.002x_2, \\ k_0, & \text{otherwise}, \end{cases} \\ k_0 &= \frac{2\pi}{10h} \end{aligned}$$

Source at $[519, 220, 130]$

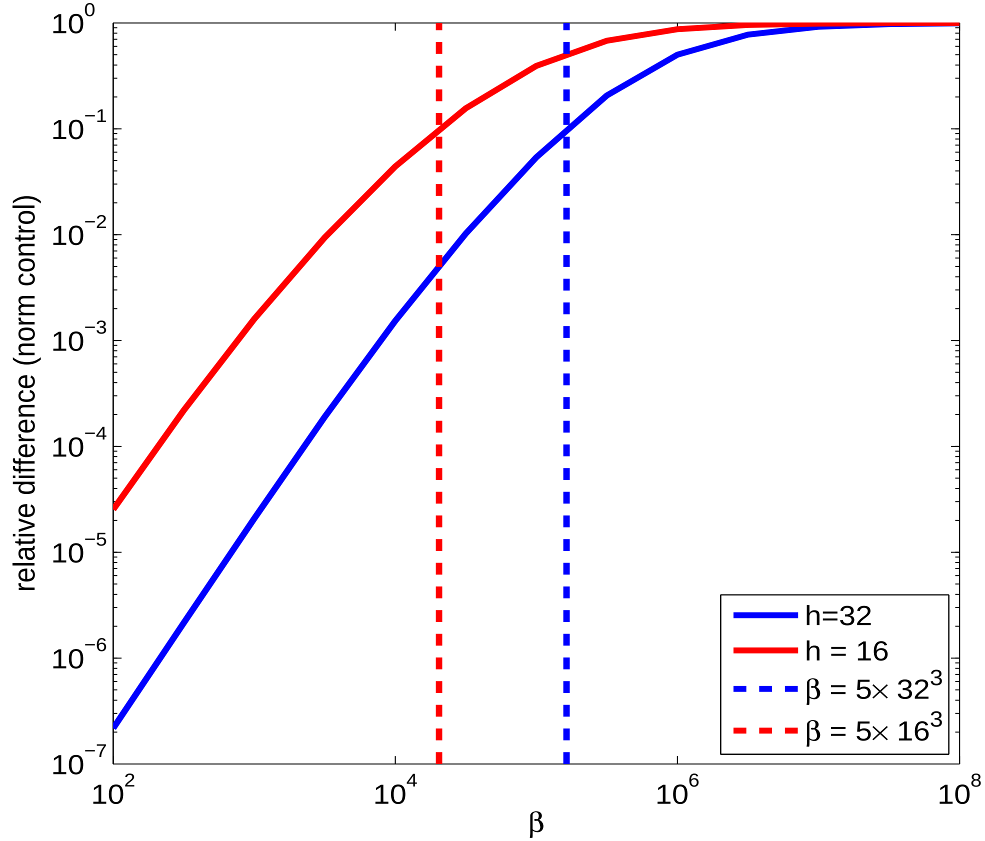
Finite difference discretisation [Huber (Basel)]

Example Problem ($h = 16$)



Given measurements of y at half-spheres (radius 10) equally distributed with centers on boundary with $x_3 = 0$, find the source.

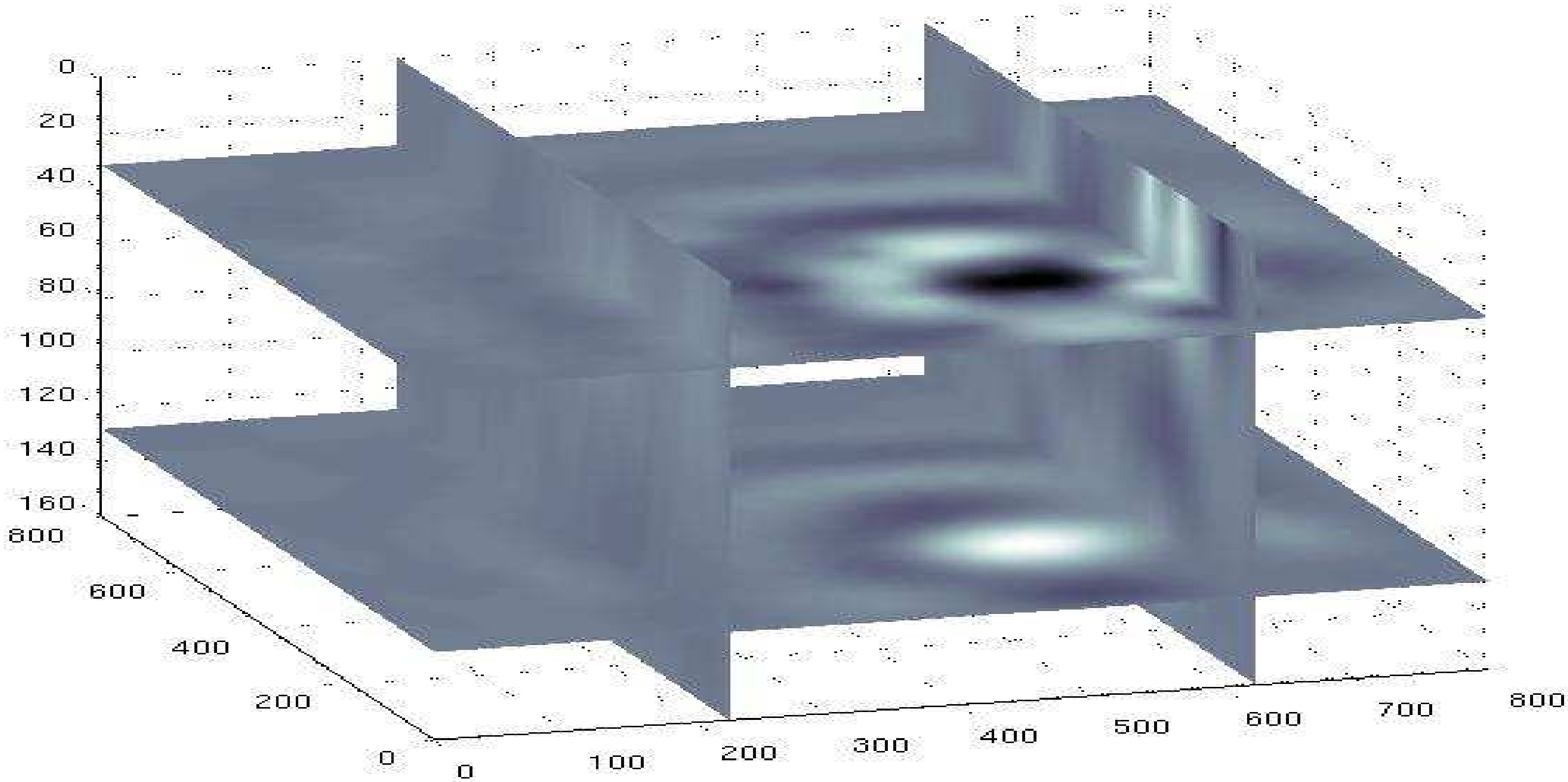
Choice of β



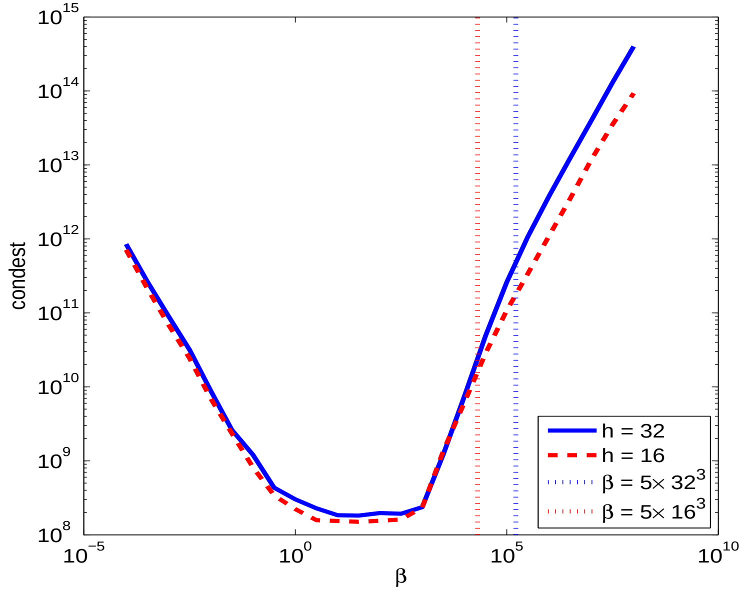
$$\beta = 5h^3$$

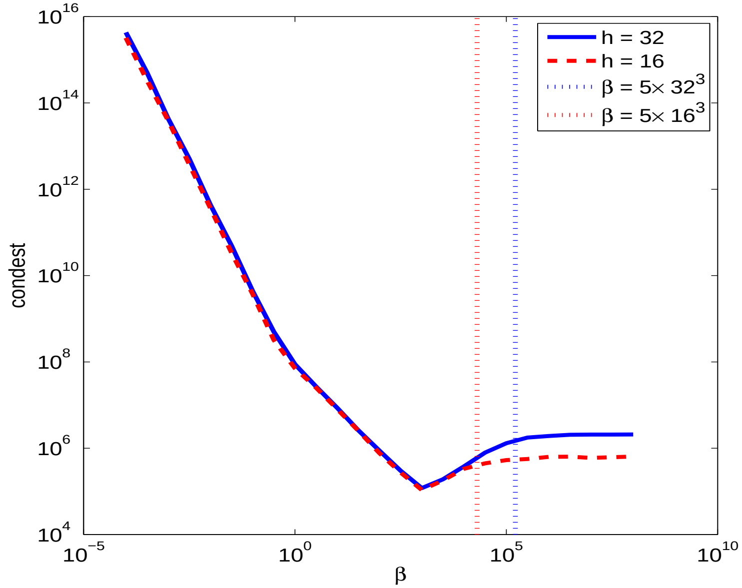
Example Problem ($h = 16, \beta = 5h^3$)

Control returned from optimisation problem (original source at [519,220,130])



Spectral properties of linear systems





Very ill-conditioned: need good preconditioner

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Choice of preconditioner

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \widehat{M} & H^* \\ -M & H & 0 \end{bmatrix} \qquad Z = \begin{bmatrix} M^{-1}H \\ I \end{bmatrix}$$

$$Z^* AZ = 2\beta H^* M^{-1} H + \widehat{M}$$

$$\mathcal{P} = \begin{bmatrix} 0 & 0 & -M \\ 0 & 2\beta H^* M^{-1} H & H^* \\ -M & H & 0 \end{bmatrix}$$

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Numerical Example

$$-\nabla^2 y - k^2 y = u,$$

$$\begin{aligned} -\nabla^2 y - k^2 y &= u \text{ in } \Omega = [0,600] \times [0,600] \times [0,160], \\ iky + \frac{\partial y}{\partial n} &= g \text{ on } \partial\Omega, \end{aligned}$$

Using finite differences and setting $\beta = 5h^3$:

$$\mathcal{A} = \begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \widehat{M} & H^* \\ -M & H & 0 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} 0 & 0 & -M \\ 0 & 2\beta H^* M^{-1} H & H^* \\ -M & H & 0 \end{bmatrix}$$

Let $\tilde{H}$ be the matrix formed from discretising the shifted problem $-\nabla^2 y - (1 - 0.01i)k^2 y = u$,

- Solves with M : Use fact that $M = h^3 I$
- Solves with H : Direct method (`HSL_MA86`); SQMR with multilevel preconditioner, or one application of multilevel preconditioner
- Multilevel preconditioner: ILUPACK (condest=20, droptol=0.005) applied to $\tilde{H}$
- PPCG : residual decreased by 10^{-6} (`HSL_MI27`)
- Fortran 95, gfortran compiler
- Hardware: Two Quad core processors (2.5GHz, 1333MHz FSB, 12MB L2 Cache), 32GB RAM

Numerical experiments

Reduced System

h	N		Direct (1 core)	Direct (8 cores)
32	8112	Setup time	2.30	1.21
		Solve time	0.04	0.03
		Total time	2.34	1.24
16	57222	Setup time	238	156
		Solve time	0.88	0.70
		Total time	239	157

Numerical experiments

Original System

h	N		Direct (1 core)	Direct (8)	PPCG (direct,1)	PPCG (direct,8)	PPCG (SQMR)	PPCG (approx)
32	12168	Forward solve	0.09	0.04	0.09	0.04	0.16 (10)	0.16 (10)
		Setup time	1.84	1.07	0.21	0.07	0.14	0.14
		Solve time	0.05	0.03	0.09	0.04	0.22	0.06
		Total time	1.89	1.10	0.29	0.12	0.36	0.20
		PPCG its	-	-	5	5	5 (143)	15
16	85833	Forward solve	2.31	0.68	2.31	0.68	2.99 (12)	2.99 (12)
		Setup time	80.6	34.8	4.49	1.29	2.68	2.69
		Solve time	0.70	0.40	1.73	1.03	4.44	1.60
		Total time	81.3	35.2	6.22	2.32	7.09	4.28
		PPCG its	-	-	8	8	7 (223)	30
8	642663	Forward solve	144	28.0	144	28.0	123 (15)	123 (15)
		Setup time	1439	311	286	54.3	114	115
		Solve time	7.67	4.22	68.7	39.8	211	90.4
		Total time	1447	315	355	94.1	325	205
		PPCG its	-	-	22	22	13 (459)	84 (1184)
4	4969323	Forward solve	-	-	-	-	1996 (12)	1996 (12)
		Setup time	-	-	-	-	1928	1930
		Solve time	-	-	-	-	3518	*
		Total time	-	-	-	-	5446	*
		PPCG its	-	-	-	-	22 (590)	*

Science & Technology Facilities Council
Rutherford Appleton Laboratory

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:

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

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

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

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

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

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

- 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

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

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.

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Distributed control with nonlinear PDEs

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

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

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

subject to

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

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

Distributed control with nonlinear PDEs

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

subject to

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

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

Distributed control with nonlinear PDEs

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

subject to

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

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

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

Preliminary results

h		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 (with Gould and Orban) - 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

- Even the simplest PDE-constrained optimzation problems give linear systems that are highly ill-conditioned
- As PDE becomes more involved, the linear system becomes even more challenging
- Poisson distributed control: optimal preconditioners available
- Helmholtz distributed control: shifted multilevel ilu preconditioner for forward problem not optimal so overall preconditioner not optimal. However, slow increase in PPCG iterations
- Distributed control with non-linear PDEs: sequence of linear systems; reuse preconditioner; Python package in development