Preconditioning for PDE-constrained Optimization

Tyrone Rees, Martin Stoll and Andy Wathen
Oxford University

OCCAM
OXFORD CENTRE FOR COLLABORATIVE APPLIED MATHEMATICS

Supported by KAUST

OXFORD UNIVERSITY
COMPUTING LABORATORY

PDE-constrained Optimization

Given $\Omega \subset \mathbb{R}^2$ or $\mathbb{R}^3$, $\widehat{\mathbf{y}} \in X$ as some desired state and bounds $\underline{\mathbf{u}}, \overline{\mathbf{u}}$ then for some (regularisation) parameter β

$$\min_{\mathbf{y},\mathbf{u}} \ \frac{1}{2}\|\mathbf{y} - \widehat{\mathbf{y}}\|_X + \frac{\beta}{2}\|\mathbf{u}\|_Y$$

subject to

$$\mathcal{L}\mathbf{y} = \mathbf{u} \quad \text{in } \Omega, \quad \mathbf{y} = \widehat{\mathbf{y}} \quad \text{on } \partial\Omega_1 \text{ and } \frac{\partial \mathbf{y}}{\partial n} = \mathbf{g} \quad \text{on } \partial\Omega_2$$

$$\underline{\mathbf{u}} \leq \mathbf{u} \leq \overline{\mathbf{u}}$$

where $\mathcal{L}$ represents a partial differential operator

Typically $X = Y = L_2(\Omega)$ or some other Hilbert space of smoother functions

Also boundary control: given $\widehat{\mathbf{y}}$, $\mathbf{f}$

$$\min_{\mathbf{y},\mathbf{u}} \ \frac{1}{2}\|\mathbf{y}-\widehat{\mathbf{y}}\|_X + \frac{\beta}{2}\|\mathbf{u}\|_Y$$

subject to

$$\mathcal{L}\mathbf{y} = \mathbf{f} \quad \text{in} \ \ \Omega, \qquad \frac{\partial \mathbf{y}}{\partial n} = \mathbf{u} \quad \text{on} \ \ \partial\Omega$$

Simple sample problem:

desirable $\widehat{\mathbf{y}} \in L_2(\Omega)$, controlable body force $\mathbf{u}$

$$\min_{\mathbf{y},\mathbf{u}} \frac{1}{2}\|\mathbf{y} - \widehat{\mathbf{y}}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|\mathbf{u}\|^2_{L_2(\Omega)}$$

subject to

$$-\nabla^2 \mathbf{y} = \mathbf{u} \quad \text{in } \Omega, \qquad \mathbf{y} = \widehat{\mathbf{y}} \quad \text{on } \partial\Omega$$

with or without the bound constraints on the control

$$\underline{\mathbf{u}} \le \mathbf{u} \le \overline{\mathbf{u}}$$

Discretize then Optimize OR Optimize then Discretize

lead to similar algebraic problems for self-adjoint operator, but generally different for non-self-adjoint problems.

We use both approaches.

For the simple sample problem for Discretize then Optimize we have:

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

$$\text{subject to} \qquad -\nabla^2 \mathrm{y} = \mathrm{u} \quad \text{in } \Omega, \quad \mathrm{y} = \widehat{\mathrm{y}} \quad \text{on } \partial\Omega$$

Discretization: finite elements

$\mathrm{y}_h = \sum y_j \phi_j, \;\; y = (y_1, y_2, \ldots, y_n)^T$

$\mathrm{u}_h = \sum u_j \phi_j, \;\; u = (u_1, u_2, \ldots, u_n)^T$

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

$$\text{subject to} \qquad K y = M u + d$$

$M = \{m_{i,j}\}, m_{i,j} = \int_\Omega \phi_i \phi_j$ — mass matrix

$K = \{k_{i,j}\}, k_{i,j} = \int_\Omega \nabla \phi_i \cdot \nabla \phi_j$ — stiffness matrix so, without bound constraints Lagrangian:

$$\frac{1}{2}\boldsymbol{y}^T \boldsymbol{M}\boldsymbol{y} + \boldsymbol{y}^T\boldsymbol{b} + \frac{\boldsymbol{\beta}}{2}\boldsymbol{u}^T\boldsymbol{M}\boldsymbol{u} + \boldsymbol{\lambda}^T(\boldsymbol{K}\boldsymbol{y} - \boldsymbol{M}\boldsymbol{u} - \boldsymbol{d})$$

stationarity $\Rightarrow$ Saddle point system

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

so, without bound constraints Lagrangian:

$$\frac{1}{2}y^T M y + y^T b + \frac{\beta}{2}u^T M u + \lambda^T (Ky - Mu - d)$$

stationarity $\Rightarrow$ Saddle point system

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

$$\text{Note} \quad B = \begin{bmatrix} K & -M \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} M & 0 \\ 0 & \beta M \end{bmatrix}$$

in usual saddle point form

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

With bound constraints: Lagrangian:

$$\frac{1}{2}\boldsymbol{y}^T\boldsymbol{M}\boldsymbol{y} \quad + \quad \boldsymbol{y}^T\boldsymbol{b} + \frac{\boldsymbol{\beta}}{2}\boldsymbol{u}^T\boldsymbol{M}\boldsymbol{u} + \boldsymbol{\lambda}^T(\boldsymbol{K}\boldsymbol{y} - \boldsymbol{M}\boldsymbol{u} - \boldsymbol{d})$$
$$+ \quad \underline{\boldsymbol{\mu}}^T(\underline{\boldsymbol{u}} - \boldsymbol{u}) + \overline{\boldsymbol{\mu}}^T(\boldsymbol{u} - \overline{\boldsymbol{u}})$$

with $\underline{\boldsymbol{\mu}}, \overline{\boldsymbol{\mu}} \geq \boldsymbol{0}$ and the complimentarity conditions

$$\underline{\boldsymbol{\mu}}^T(\underline{\boldsymbol{u}} - \boldsymbol{u}) = \boldsymbol{0} = \overline{\boldsymbol{\mu}}^T(\boldsymbol{u} - \overline{\boldsymbol{u}})$$

Block diagonal preconditioners:

based on the observation (*Murphy, Golub & W (2000)*)

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

preconditioned by

$\begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix}$ has 3 distinct eigenvalues $\quad 1, \;\; \frac{1}{2} \pm \frac{\sqrt{5}}{2}$

where $S = BA^{-1}C^T$ (Schur Complement)

$\Rightarrow$ appropriate Krylov subspace iteration (MINRES when $B = C$, GMRES when $B \neq C$) terminates in 3 iterations

$\Rightarrow$ want approximations $\widehat{A}, \;\; \widehat{S} \Rightarrow$ 3 clusters

$\Rightarrow$ fast convergence

$$\textbf{Recall} \quad B = C = \begin{bmatrix} K & -M \end{bmatrix} \quad \textbf{and} \quad A = \begin{bmatrix} M & 0 \\ 0 & \beta M \end{bmatrix}$$

so $\widehat{S}$? $\quad S = BA^{-1}B^T$ (Schur Complement)

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

$$= \frac{1}{\beta} M + K M^{-1} K^T$$

For this problem unless approx $\beta < 10^{-6}$ dominant part is $\widehat{S} = KM^{-1}K^T$

(*Schöberl & Zulehner (2007)*)

Hence preconditioner for

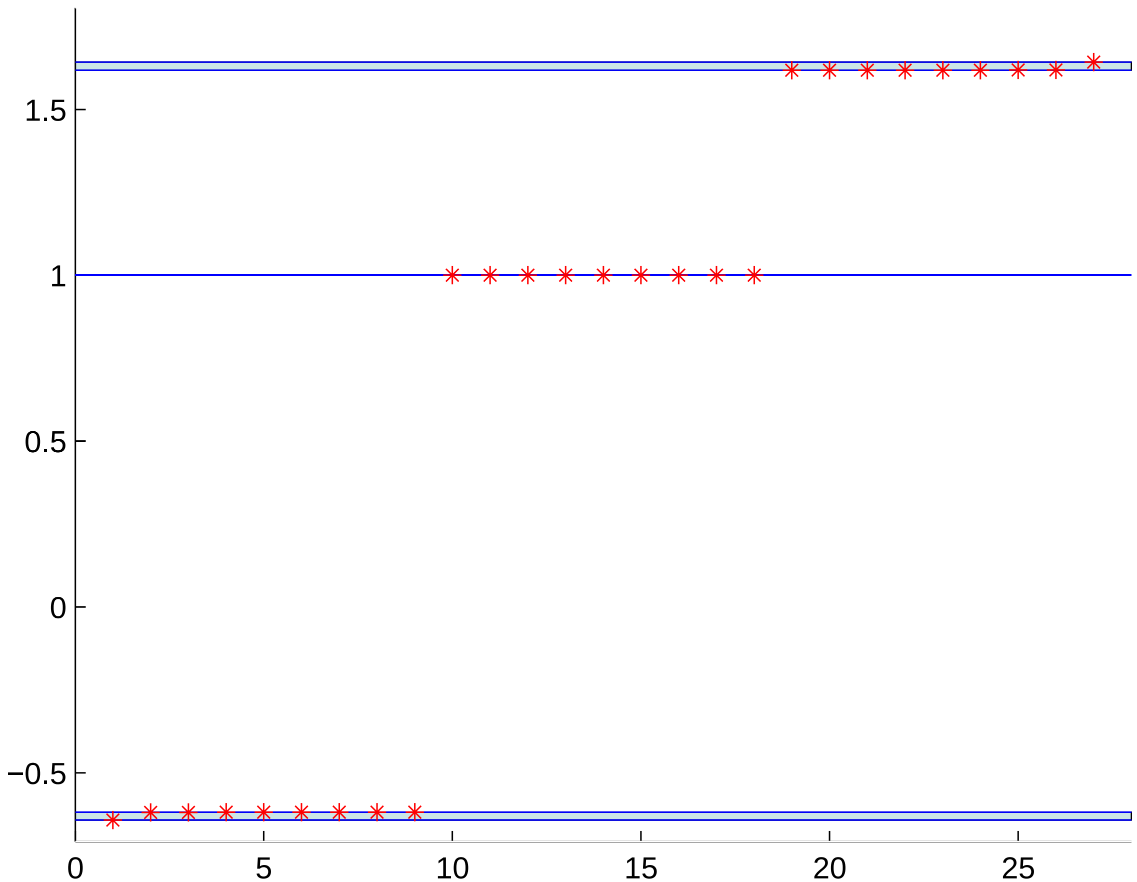
$$\mathcal{A} = \begin{bmatrix} M & 0 & K^T \\ 0 & \beta M & -M \\ K & -M & 0 \end{bmatrix} \text{ is } \mathcal{P} = \begin{bmatrix} M & 0 & 0 \\ 0 & \beta M & 0 \\ 0 & 0 & KM^{-1}K^T \end{bmatrix}$$

Eigenvalues ν of $\mathcal{P}^{-1}\mathcal{A}$

$$\begin{aligned} \nu &= 1, \\ \frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right) \le \; \nu \; &\le \frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_2}{\beta}}\right) \\ \text{or } \frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_2}{\beta}}\right) \le \; \nu \; &\le \frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right), \end{aligned}$$

where $\alpha_1, \ \alpha_2$ are positive constants independent of h.

$$\mathcal{P} = \begin{bmatrix} M & 0 & 0 \\ 0 & \beta M & 0 \\ 0 & 0 & KM^{-1}K^T \end{bmatrix}, \quad \beta = 10^{-2}$$



But

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

still expensive to use in practice so employ approximations

$$\widehat{M} \simeq M \quad \text{and} \quad \widehat{K} \simeq K$$

giving

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

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

so $\widehat{M}$?

Mass matrix is effectively preconditioned by its diagonal *(W (1987))*: M and D=diag(M)

eg. for Q1 (bilinear) finite elements in 2-dimensions (rectangles)

eigenvalues of $D^{-1}M$ all in $[1/4, 9/4]$

eg. for Q1 (trilinear) finite elements in 3-dimensions (bricks)

eigenvalues of $D^{-1}M$ all in $[1/8, 27/8]$

independently of mesh size h (ie. independently of discrete problem dimension)

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

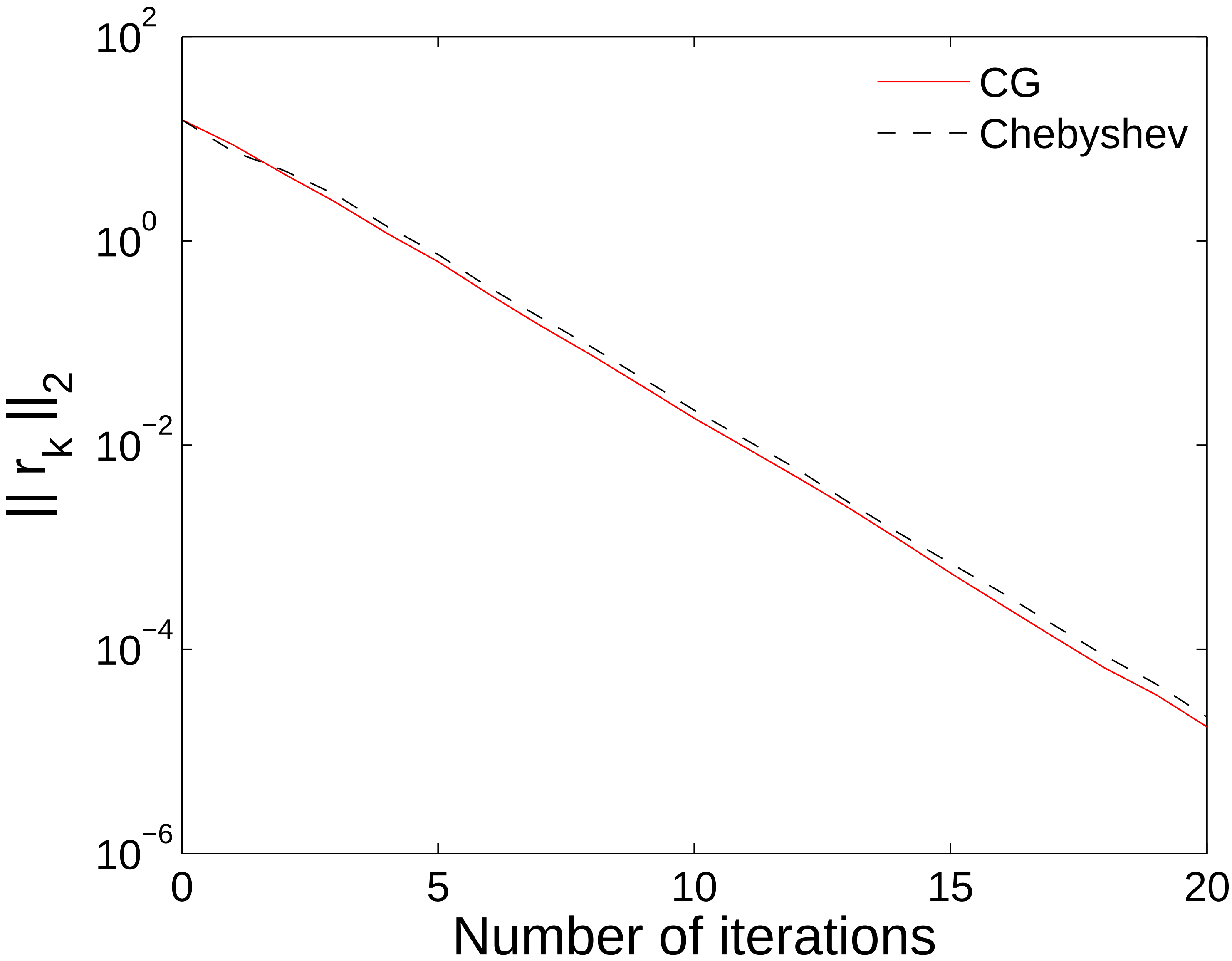
Could use

$$\widehat{A} = \begin{bmatrix} D & 0 \\ 0 & \beta D \end{bmatrix}$$

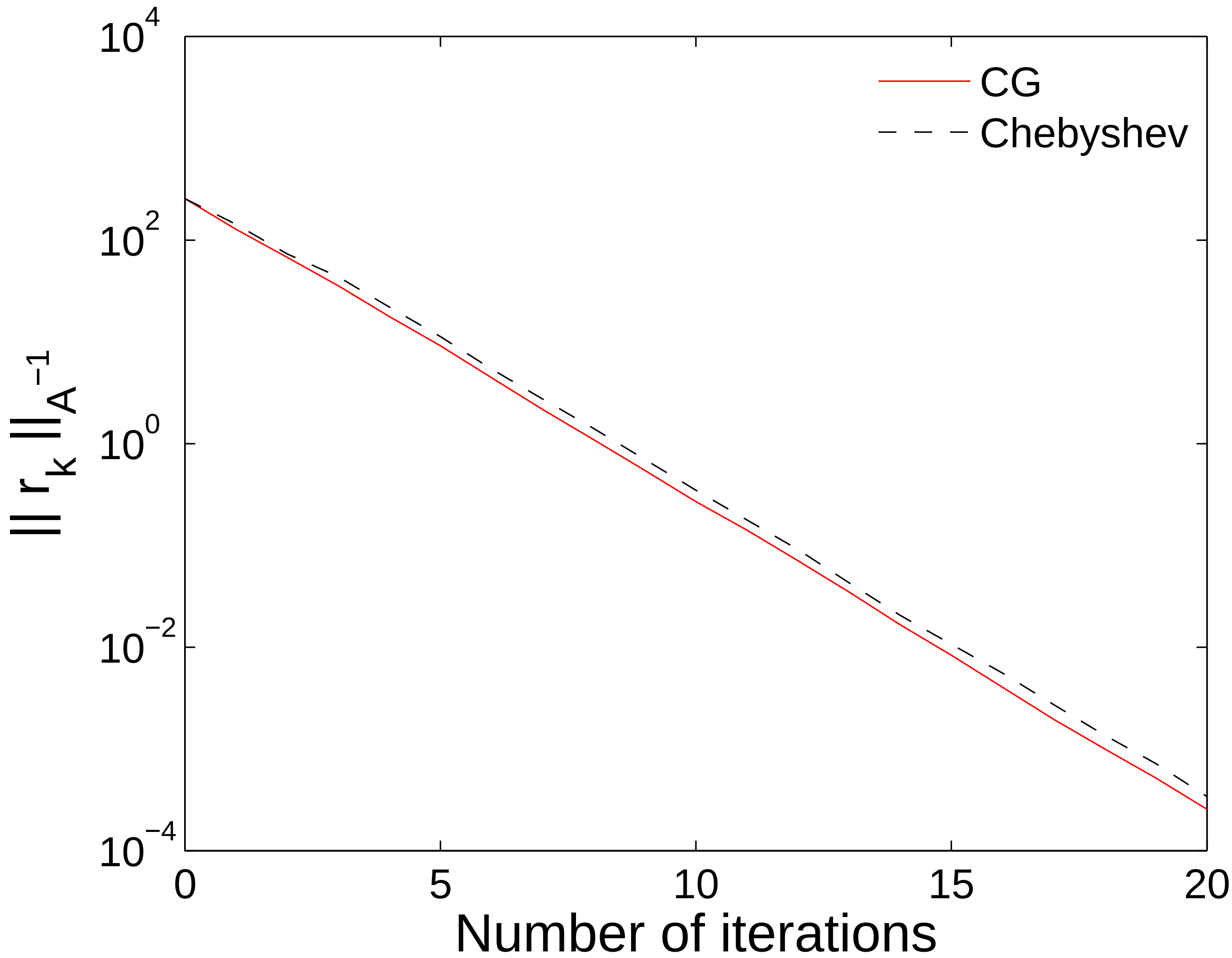
but better a few iterations of a diagonally preconditioned iteration for M:

- diagonally scaled Conjugate Gradients: leads to a nonlinear preconditioner
- diagonally scaled Chebyshev (semi-)iteration: is linear and we have precise eigenvalue inclusion intervals

(*W & Rees (2008)*)

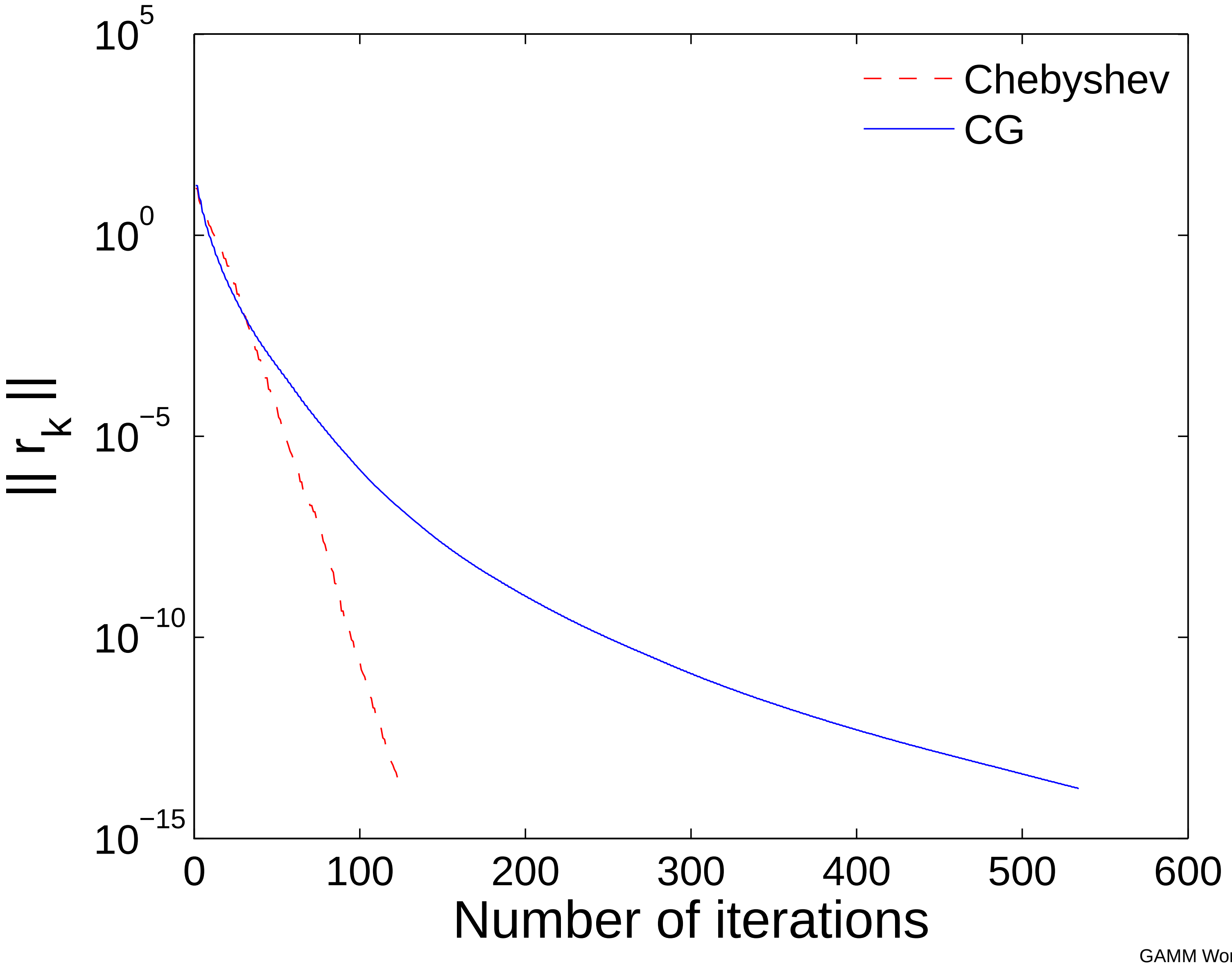


CG
Chebyshev
10^{2}
10^{0}
10^{-2}
10^{-4}
10^{-6}
0
5
10
15
20
|| r_k ||$_2$
Number of iterations



10^4
10^2
10^0
10^-2
10^-4
|| r_k ||_A^-1
CG
Chebyshev
0
5
10
15
20
Number of iterations

and what can happen when CG used as inner iteration:



$$\text{Recall} \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{K} & -\boldsymbol{M} \end{bmatrix} \quad \text{and} \quad \boldsymbol{A} = \begin{bmatrix} \boldsymbol{M} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\beta M} \end{bmatrix}$$

so $\widehat{\boldsymbol{S}}$? $\quad \boldsymbol{S} = \boldsymbol{B A^{-1} B^T}$ (Schur Complement)

$$= \begin{bmatrix} \boldsymbol{K} & -\boldsymbol{M} \end{bmatrix} \begin{bmatrix} \boldsymbol{M}^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\boldsymbol{\beta}} \boldsymbol{M}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{K}^T \\ -\boldsymbol{M} \end{bmatrix}$$

$$= \frac{1}{\boldsymbol{\beta}} \boldsymbol{M} + \boldsymbol{K M^{-1} K^T}$$

Unless approx $\beta < 10^{-6}$ dominant part is $\widehat{\boldsymbol{S}} = \boldsymbol{K M^{-1} K^T}$

But want efficient approximations $\widehat{\boldsymbol{K}} \simeq \boldsymbol{K}$
ie. a good preconditioner for the PDE and adjoint
(Braess & Peisker (1986))

For $\mathcal{L} = -\nabla^2$, K is a discrete Laplacian: use multigrid cycles

- geometric mutigrid: relaxed Jacobi smoothing, standard grid transfers
  $\rightarrow \mathcal{P}_{MG}$
- algebraic multigrid: HSL routine HSL_MI20
  (Boyle, Mihajlovic & Scott (2007))
  $\rightarrow \mathcal{P}_{AMG}$

In our examples:

$\widehat{K}$ is the action of 2 V-cycles

$\widehat{M}$ is the action of 20 Chebyshev semi-iterative steps

Example problem: $\Omega = [0,1]^d, d = 2, 3$

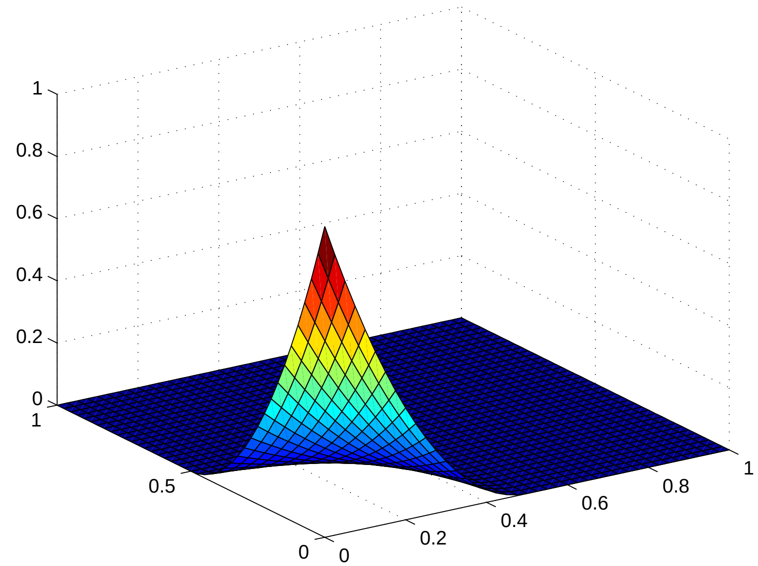
$$\min_{\mathbf{y},\mathbf{u}} \frac{1}{2}\|\mathbf{y} - \widehat{\mathbf{y}}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|\mathbf{u}\|^2_{L_2(\Omega)}$$

subject to

$$-\nabla^2 \mathbf{y} = \mathbf{u} \quad \text{in } \Omega, \qquad \mathbf{y} = \widehat{\mathbf{y}} \quad \text{on } \partial\Omega$$

Q1 (bilinear/trilinear) finite elements, $\beta = 10^{-2}$

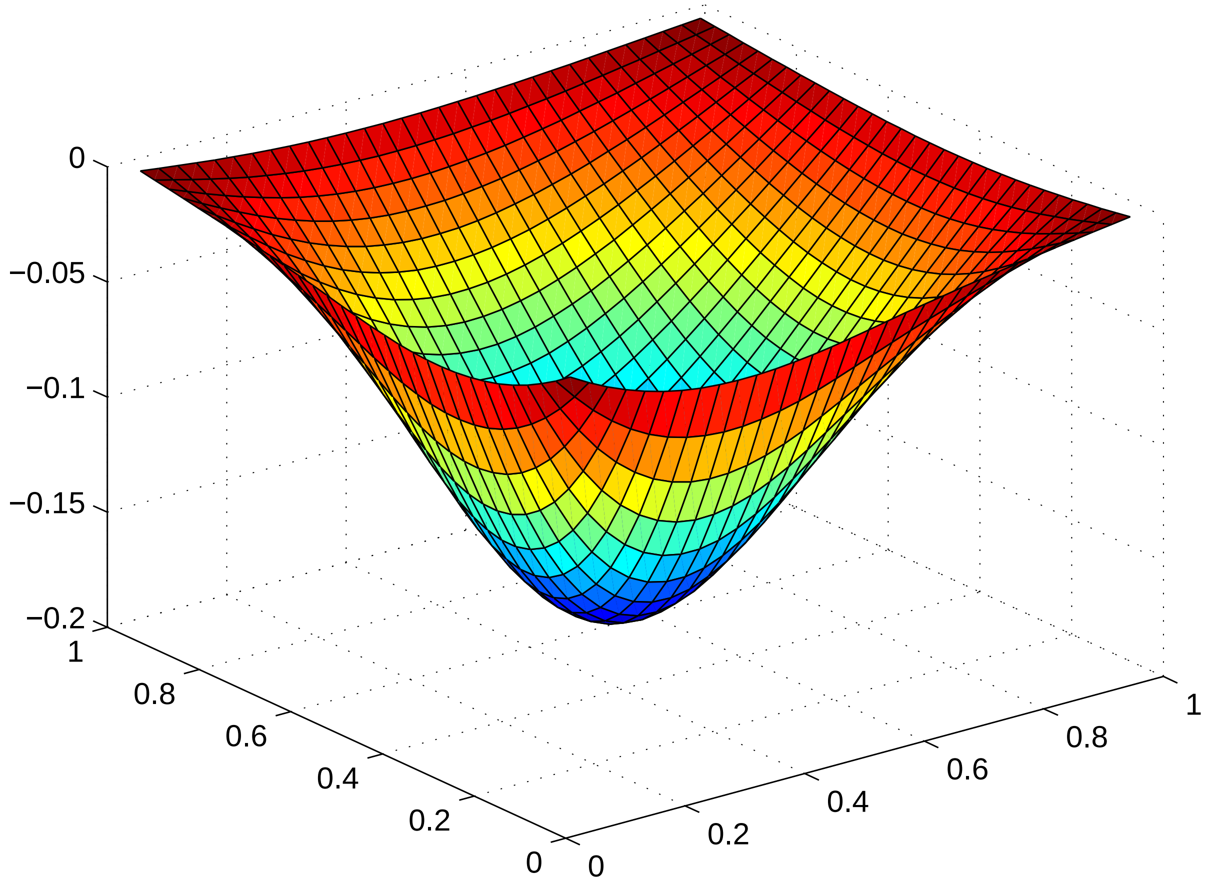
$\widehat{\mathbf{y}}$:



CPU times (MINRES iterations) in 2D, tol 10^{-8}

h	3n	backslash	MINRES ($\mathcal{P}_{MG}$)	MINRES ($\mathcal{P}_{AMG}$)
2^{-2}	27	0.0002	0.15 (10)	0.032 (10)
2^{-3}	147	0.002	0.17 (10)	0.038 (10)
2^{-4}	675	0.009	0.26 (12)	0.071 (12)
2^{-5}	2883	0.062	0.50 (12)	0.18 (12)
2^{-6}	11907	0.37	1.67 (12)	0.80 (12)
2^{-7}	48387	2.22	6.60 (12)	3.95 (14)
2^{-8}	195075	15.7	30.9 (12)	19.0 (14)
2^{-9}	783363	—	134 (11)	88.9 (13)

Control:



CPU times (MINRES iterations) in 3D, tol 10^{-8}

h	3n	backslash	MINRES ($\mathcal{P}_{MG}$)	MINRES ($\mathcal{P}_{AMG}$)
2^{-2}	81	0.001	0.14 (8)	0.031 (8)
2^{-3}	1029	0.013	0.28 (10)	0.14 (10)
2^{-4}	10125	25.5	2.04 (10)	2.30 (10)
2^{-5}	89373	—	19.2 (10)	26.7 (10)
2^{-6}	750141	—	230 (10)	—

With bound constraints: an active set strategy (or projected gradient) $\Rightarrow$ an outer nonlinear loop

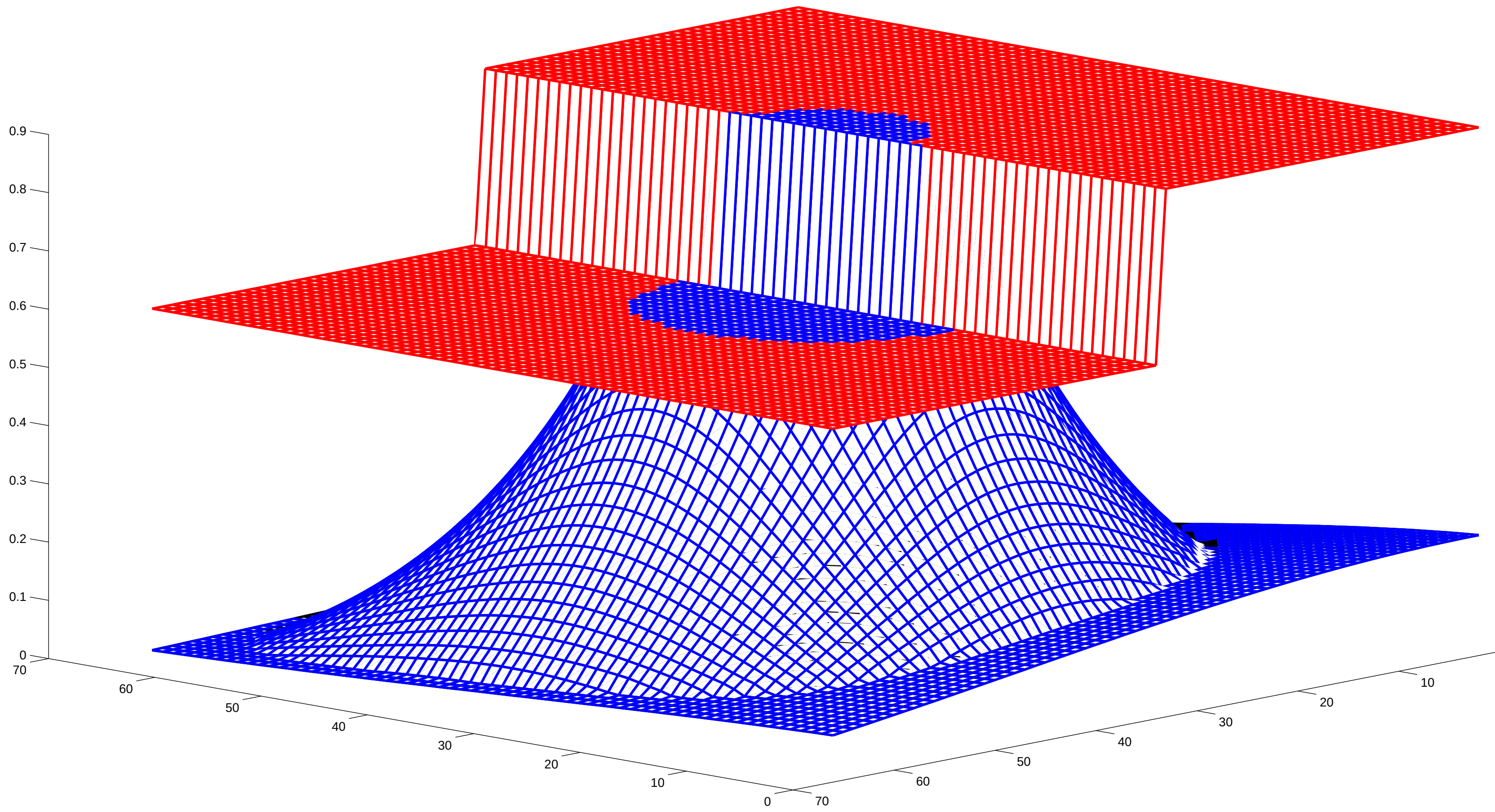
MINRES solution: Number of AMG V-cycles for Laplacian in 2D, tol 10^{-6}

h	3n	PDE solve	Unconstrained Control prob	Bound-constrained Control prob
2^{-2}	27	4	40	68(2)
2^{-3}	147	4	40	104(3)
2^{-4}	675	4	40	160(4)
2^{-5}	2883	6	40	160(4)
2^{-6}	11907	6	44	180(4)
2^{-7}	48387	6	48	240(5)
2^{-8}	195075	6	48	280(5)
2^{-9}	783363	6	56	320(5)

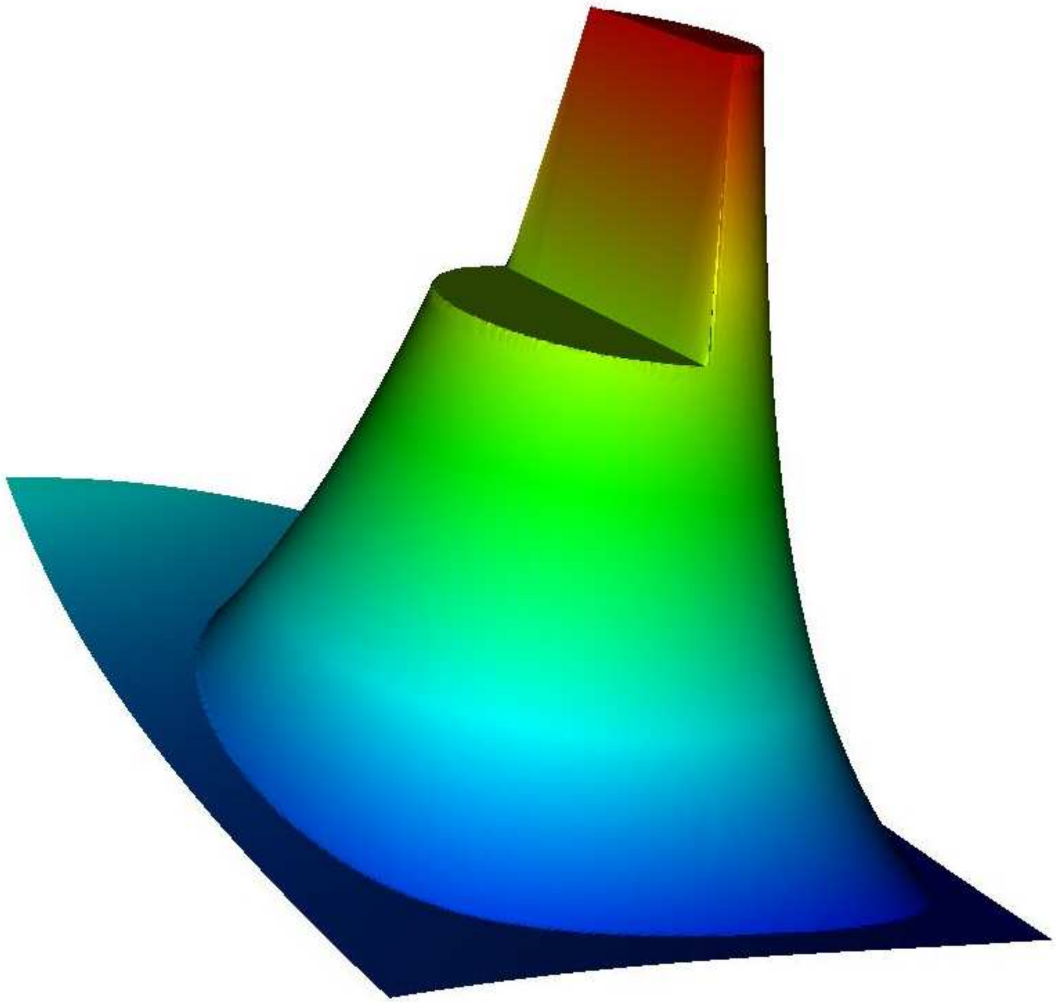
energy: unconstrained $\searrow 1.483 \times 10^{-3}$

bound constrained $\searrow 1.732 \times 10^{-3}$

Control and bounds:



Control using deal.II and VISIT:



Boundary control:

CPU times (MINRES iterations) in 2D, tol 10^{-4}

h	3n	backslash	MINRES ($\mathcal{P}_{MG}$)	MINRES ($\mathcal{P}_{AMG}$)
2^{-2}	66	0.0006	0.13 (15)	0.044 (7)
2^{-3}	194	0.005	0.15 (15)	0.034 (7)
2^{-4}	642	0.011	0.24 (13)	0.085 (13)
2^{-5}	2306	0.078	0.48 (13)	0.18 (13)
2^{-6}	8706	0.70	1.62 (13)	1.12 (21)
2^{-7}	33794	6.73	6.54 (13)	6.99 (31)
2^{-8}	133122	69.1	27.2 (13)	13.7 (13)
2^{-9}	528386	—	109 (13)	103 (27)

Control of Convection-Diffusion

$$\min_{\mathbf{y},\mathbf{u}} \frac{1}{2}||\mathbf{y} - \widehat{\mathbf{y}}||^2_{L_2(\Omega)} + \frac{\beta}{2}||\mathbf{u}||^2_{L_2(\Omega)}$$

such that
$$\begin{aligned} -\epsilon\nabla^2\mathbf{y} + \vec{\mathbf{w}}\cdot\nabla\mathbf{y} &= \mathbf{u} \ \text{ in } \ \Omega \\ \text{with} \quad \mathbf{y} &= \mathbf{g} \ \text{ on } \ \partial\Omega, \end{aligned}$$

where $\widehat{\mathbf{y}}$, $\mathbf{g}$, $\vec{\mathbf{w}}$, ϵ and β are given ($|\vec{\mathbf{w}}| = 1, \epsilon = 1/200$ here)

Finite elements: SUPG (*Hughes & Brookes (1982)*) $\Rightarrow$

$$(K_{sd})_{i,j} = \epsilon\int_\Omega \nabla\phi_i\cdot\nabla\phi_j + \int_\Omega (\vec{w}\cdot\nabla\phi_i)\phi_j + \delta\int_\Omega (\vec{w}\cdot\nabla\phi_i)(\vec{w}\cdot\nabla\phi_j)$$

and $(M_{sd})_{i,j} = \int_\Omega \phi_i\phi_j + \delta\int_\Omega \phi_i(\vec{w}\cdot\nabla\phi_j)$

Discretize then optimize

$$\begin{bmatrix} M & 0 & K_{sd}^T \\ 0 & \beta M & -M_{sd}^T \\ K_{sd} & -M_{sd} & 0 \end{bmatrix} \begin{bmatrix} y \\ u \\ p \end{bmatrix} = \begin{bmatrix} M\widehat{y} \\ 0 \\ d \end{bmatrix}$$

Schur complement: $S = \frac{1}{\beta} M_{sd} M^{-1} M_{sd}^T + K_{sd} M^{-1} K_{sd}^T$

Preconditioner: $\mathcal{P} := \begin{bmatrix} M & 0 & 0 \\ 0 & \beta M & 0 \\ 0 & 0 & K_{sd} M^{-1} K_{sd}^T \end{bmatrix}$

Schur complement: $$S = \frac{1}{\beta} M_{sd} M^{-1} M_{sd}^T + K_{sd} M^{-1} K_{sd}^T$$

Preconditioner: $$\mathcal{P} := \begin{bmatrix} M & 0 & 0 \\ 0 & \beta M & 0 \\ 0 & 0 & K_{sd} M^{-1} K_{sd}^T \end{bmatrix}$$

h	δ	δ / h_k	MINRES iterations			
			$\beta = 1$	$\beta = 10^{-2}$	$\beta = 10^{-4}$	$\beta = 10^{-6}$
2^{-2}	.17	0.49	9	18	21	21
2^{-3}	.08	0.48	11	33	97	117
2^{-4}	.04	0.46	12	43	203	353
2^{-5}	.02	0.42	12	47	281	(> 500)
2^{-6}	.01	0.34	12	49	324	(> 500)

Schur complement: $$S = \frac{1}{\beta} M_{sd} M^{-1} M_{sd}^T + K_{sd} M^{-1} K_{sd}^T$$

Preconditioner: $$\mathcal{P}_{mass} := \begin{bmatrix} M & 0 & 0 \\ 0 & \beta M & 0 \\ 0 & 0 & \frac{1}{\beta} M_{sd} M^{-1} M_{sd}^T \end{bmatrix}$$

h	δ	δ/h	MINRES iterations			
			$\beta = 10^{-2}$	$\beta = 10^{-4}$	$\beta = 10^{-6}$	$\beta = 10^{-8}$
2^{-2}	.17	.49	18	7	3	3
2^{-3}	.08	.48	54	11	5	3
2^{-4}	.04	.46	147	19	5	3
2^{-5}	.02	.42	400	43	7	3
2^{-6}	.01	.34	(> 500)	111	11	4

Optimize then discretize

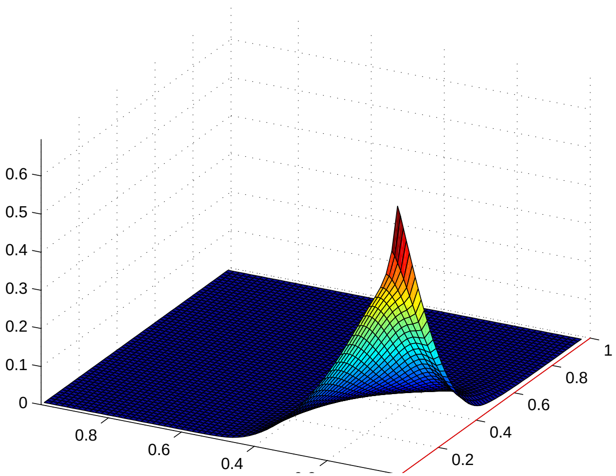
$$\begin{bmatrix} N_{sd} & 0 & L_{sd} \\ 0 & \beta M & -M \\ K_{sd} & -M_{sd} & 0 \end{bmatrix} \begin{bmatrix} y \\ u \\ p \end{bmatrix} = \begin{bmatrix} M\widehat{y} \\ 0 \\ d \end{bmatrix}$$

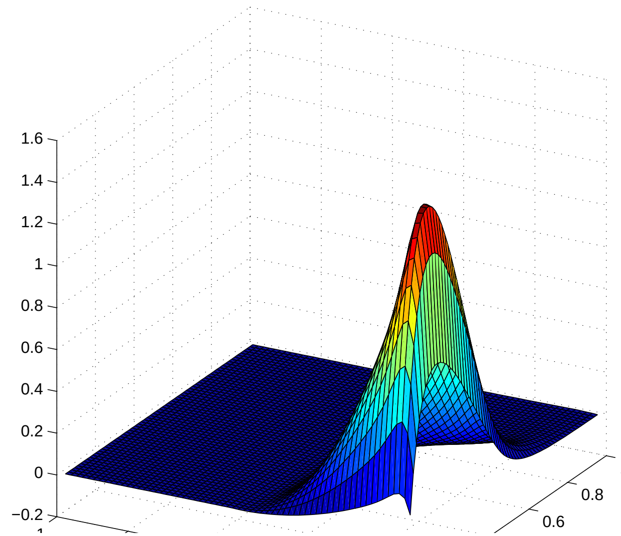
$$\mathcal{P} := \begin{bmatrix} N_{sd} & 0 & 0 \\ 0 & \beta M & 0 \\ 0 & 0 & K_{sd} N_{sd}^{-1} L_{sd} \end{bmatrix}$$

or $\mathcal{P}_{mass} := \begin{bmatrix} N_{sd} & 0 & 0 \\ 0 & \beta M & 0 \\ 0 & 0 & \frac{1}{\beta} M_{sd} \end{bmatrix}$

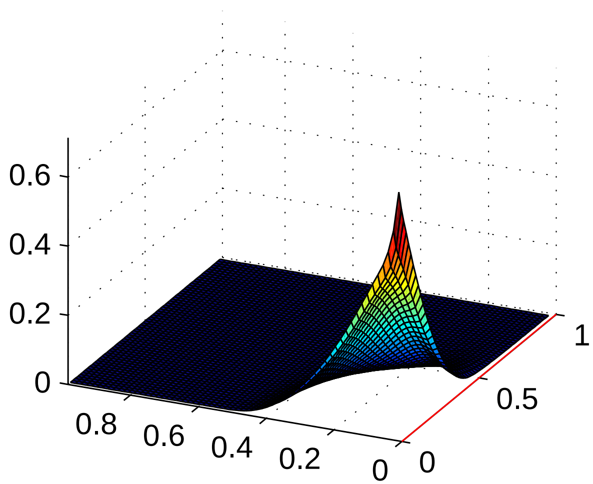
Now nonsymmetric $\Rightarrow$ GMRES , but results similar to Discretize then Optimize

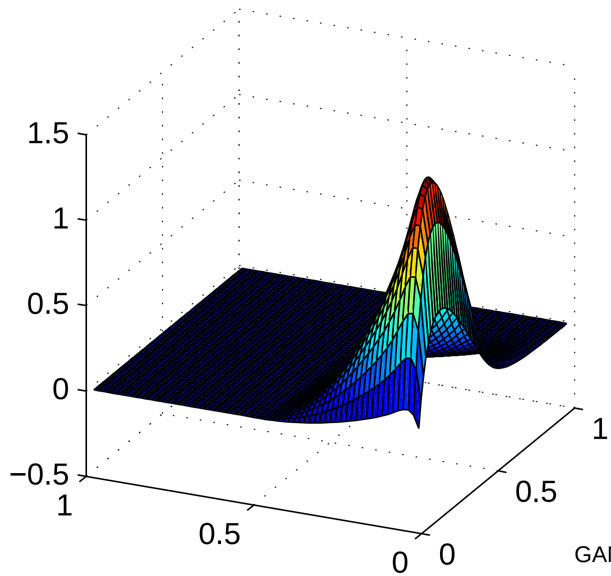
D
I
t
I
O


0.6
0.5
0.4
0.3
0.2
0.1
0
0.8
0.6
0.4
1
0.8
0.6
0.4
0.2


1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
−0.2
1
0.8
0.6

O
I
t
I
D


0.6
0.4
0.2
0
0.8
0.6
0.4
0.2
0
0
0.5
1


1.5
1
0.5
0
−0.5
1
0.5
0
0
0.5
1

Stokes Control

$$\min_{\mathbf{y},\mathbf{p},\mathbf{u}} \frac{1}{2}\|\vec{\mathbf{y}} - \hat{\mathbf{y}}\|^2_{L^2(\Omega)} + \frac{1}{2}\|\mathbf{p} - \hat{\mathbf{p}}\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|\mathbf{u}\|^2_{L^2(\Omega)}$$

$$\text{subject to} \qquad \begin{aligned} -\nabla^2\vec{\mathbf{y}} + \nabla\mathbf{p} &= \mathbf{u} \\ \nabla\cdot\vec{\mathbf{y}} &= 0 \end{aligned}$$

$\vec{\mathbf{y}}$: velocity, $\mathbf{p}$: pressure.

Mixed finite elements for (forward) Stokes problem:

$$\begin{bmatrix} \underline{K} & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} y \\ p \end{bmatrix} = \begin{bmatrix} u \\ g \end{bmatrix}, \; \underline{K} = \begin{bmatrix} K & 0 \\ 0 & K \end{bmatrix} \quad \text{in } \mathbb{R}^2$$

Cost functional

$$\frac{1}{2}y^T M_y y - y^T b + \frac{1}{2}p^T M_p p - p^T d + \frac{\beta}{2}u^T M_u u$$

combined with constraint via the Lagrangian $\Rightarrow$

$$\begin{bmatrix} M_y & 0 & 0 & \frac{K}{B} & B^T \\ 0 & M_p & 0 & \frac{K}{B} & 0 \\ 0 & 0 & \beta M_u & -M_u & 0 \\ \frac{K}{B} & B^T & -M_u & 0 & 0 \\ \frac{K}{B} & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ p \\ u \\ \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} b \\ d \\ 0 \\ h \\ k \end{bmatrix}.$$

Block diagonal preconditioner:

Mass matrices approximated by Chebyshev as before,

Stokes preconditioner:

$$\left[\begin{array}{cc} \widehat{\underline{\boldsymbol{K}} \quad \boldsymbol{B}^T} \\ \boldsymbol{B} & \mathbf{0} \end{array}\right] = \left[\begin{array}{cc} \widehat{\underline{\boldsymbol{K}}} & \mathbf{0} \\ \boldsymbol{B} & \widehat{\boldsymbol{M}_p} \end{array}\right]$$

where $\widehat{\underline{\boldsymbol{K}}}$ is multigrid cycles for each discrete scalar Laplacian as before
(*Silvester & W (1993), Klawonn (1998)*)

Gives symmetric Schur complement approximation

Control problem:

h	# elts	$\begin{bmatrix} \underline{K} & 0 \\ B & M_p \end{bmatrix}$ #its	time	$\begin{bmatrix} \underline{\widehat{K}} & 0 \\ B & \widehat{M_p} \end{bmatrix}$ #its	time	#Vcyc
2^{-2}	344	36	0.4	35	0.6	140
2^{-3}	1512	47	1.1	47	1.2	188
2^{-4}	6344	53	5.2	57	3.9	228
2^{-5}	25992	57	37	70	22	280
2^{-6}	105224	57	168	115	153	460

(Forward) Stokes solve:

h	# elts	#its	time	# Vcyc
2^{-2}	187	25	0.04	25
2^{-3}	659	27	0.1	27
2^{-4}	2467	30	0.3	30
2^{-5}	9539	30	1.3	30
2^{-6}	37507	28	5.7	28
2^{-7}	148739	28	23	28

References

- *Rees, T., Dollar , H.S. & Wathen, A.J., 2008, 'Optimal solvers for PDE-constrained optimization', to appear in SIAM J Sci. Comput., Oxford University NA report 08/10.*

- *Wathen, A.J. & Rees, T., 2008, 'Chebyshev semi-iteration in preconditioning', to appear in ETNA., Oxford University NA report 08/14*

- *Murphy, M.F., Golub, G.H. & Wathen, A.J., 2000, 'A note on preconditioning for indefinite linear systems', SIAM J. Sci. Comput.* **21***(6), pp. 1969-1972.*

Acknowledgements

This work is partially supported by Award No. KUK-C1-013-04 made by King Abdullah University of Science and Technology (KAUST) and by the Engineering and Physical Sciences Research Council via studentship OUCL/TR/14.