Efficient Block Preconditioning for a C^1 Finite Element Discretisation of the Dirichlet biharmonic problem

Jennifer Pestana, Richard Muddle, Matthias Heil, Françoise Tisseur & Milan Mihalović

July 31, 2017

University of Strathclyde Glasgow

The biharmonic problem

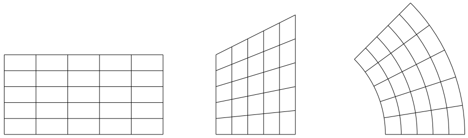
The problem

The biharmonic problem

Find $u \in C^4(\Omega)$ such that

$$\nabla^4 u = f, \qquad u = \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega,$$

where $\Omega \subset \mathbb{R}^2$ has piecewise smooth boundary $\partial\Omega$, $f \in L^2(\Omega)$.



Two approaches

Original form

$$\nabla^4 u = f, \qquad u = \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega$$

Mixed form

$$-\nabla^2 w = f$$
$$-\nabla^2 u = w$$

$$u = \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega$$

Two approaches

Original form

$$\nabla^4 u = f, \qquad u = \frac{\partial u}{\partial n} = 0 \textbf{ on } \partial\Omega$$

Mixed form

$$-\nabla^2 w = f$$

$$-\nabla^2 u = w$$

$$u = \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega$$

The finite element formulation

Finite element formulation

$$a(u, v) := \int_\Omega \nabla^2 u \, \nabla^2 v \; \mathrm{d}\Omega = \int_\Omega fv \; \mathrm{d}\Omega,$$

$$u = \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega$$

$u, v \in S_0 \subset H_0^2(\Omega), \quad S_0 = \mathrm{span}\{\phi_1, \phi_2, \ldots, \phi_N\}$

The finite element formulation

$$\sum_{i=1}^{N} u_i \int_\Omega \nabla^2\phi_i \, \nabla^2\phi_j \, \mathrm{d}\Omega = \int_\Omega f \phi_j \, \mathrm{d}\Omega$$

Linear system

$\mathcal{A}\boldsymbol{u} = \boldsymbol{b}$,

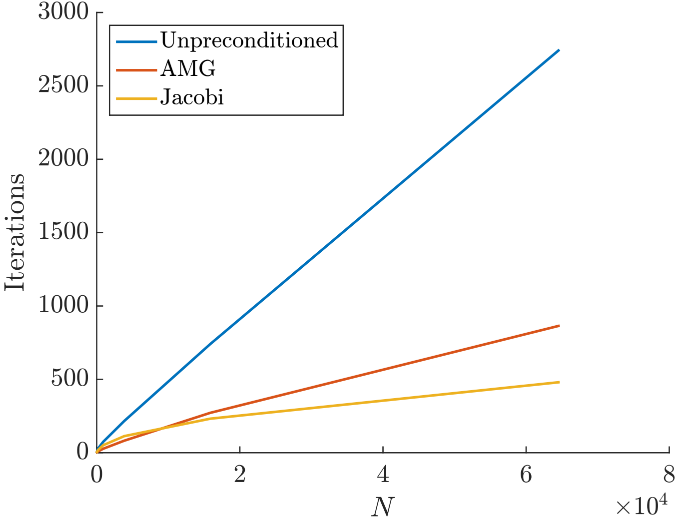
$$(\mathcal{A})_{ij} = \underbrace{\int_\Omega \nabla^2\phi_i \, \nabla^2\phi_j \, \mathrm{d}\Omega,}_{a(\phi_i,\phi_j)}$$

$$(\boldsymbol{u})_i = u_i, \qquad (\boldsymbol{b})_j = \int_\Omega f \phi_j \, \mathrm{d}\Omega$$

$\mathcal{A}$ is symmetric positive definite and sparse

Preconditioned conjugate gradients

PCG



Unpreconditioned
AMG
Jacobi
3000
2500
2000
1500
1000
500
0
Iterations
0
2
4
6
8
N
$\times 10^4$

PCG

- Additive Schwarz methods
  X. Zhang (1994,1996)
- BPX preconditioning
  Oswald (1995)
- Problem-specific multigrid methods
  S. Zhang (1989), Bramble & X. Zhang (1995)
- Fast auxiliary space (FASP) preconditioning
  S. Zhang and Xu (2014)

Block preconditioners

Block preconditioners

$$\mathcal{A}_{ij} = \int_{\Omega} \nabla^2 \phi_i \nabla^2 \phi_j \, \mathrm{d}\Omega$$

$$\mathcal{A} = \begin{array}{cl} \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{12}^T & A_{22} & A_{23} & A_{24} \\ A_{13}^T & A_{23}^T & A_{33} & A_{34} \\ A_{14}^T & A_{24}^T & A_{34}^T & A_{44} \end{bmatrix} & \begin{array}{l} \} \, u \\ \} \, \partial u / \partial x \\ \} \, \partial u / \partial y \\ \} \, \partial^2 u / \partial x \partial y \end{array} \end{array}$$

$A_{ij} \in \mathbb{R}^{n \times n}, \; i, j = 1, \ldots, 4$

First preconditioner

$$\mathcal{P}_1 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & A_{23} & \\ A_{13}^T & A_{23}^T & A_{33} & \\ & & & A_{44} \end{bmatrix}$$

	4×4	8×8	16×16	32×32
$\lambda_{min}(\mathcal{P}_1^{-1}\mathcal{A})$	0.72	0.64	0.61	0.60
$\lambda_{max}(\mathcal{P}_1^{-1}\mathcal{A})$	1.28	1.36	1.39	1.40

A first analysis

$$\mathcal{A} = \left[\begin{array}{ccc|c} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{12}^T & A_{22} & A_{23} & A_{24} \\ A_{13}^T & A_{23}^T & A_{33} & A_{34} \\ \hline A_{14}^T & A_{24}^T & A_{34}^T & A_{44} \end{array}\right] \qquad \mathcal{P}_1 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & A_{23} & \\ A_{13}^T & A_{23}^T & A_{33} & \\ & & & A_{44} \end{bmatrix}$$

A first analysis

$$\mathcal{A} = \left[\begin{array}{ccc|c} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{12}^T & A_{22} & A_{23} & A_{24} \\ A_{13}^T & A_{23}^T & A_{33} & A_{34} \\ \hline A_{14}^T & A_{24}^T & A_{34}^T & A_{44} \end{array}\right] \left.\vphantom{\begin{array}{c} A \\ A \\ A \end{array}}\right\}\text{rank } r \qquad \mathcal{P}_1 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & A_{23} & \\ A_{13}^T & A_{23}^T & A_{33} & \\ & & & A_{44} \end{bmatrix}$$

Theorem (PMHTH 16)

$\mathcal{P}_1^{-1}\mathcal{A}$ has $4n - 2r$ eigenvalues equal to 1 and any remaining eigenvalues λ satisfy

$$0 < 1 - \sqrt{\mu_{max}} \leq \lambda \leq 1 + \sqrt{\mu_{max}} < 2.$$

A first analysis

$$\mathcal{A} = \left[\begin{array}{ccc|c} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{12}^T & A_{22} & A_{23} & A_{24} \\ A_{13}^T & A_{23}^T & A_{33} & A_{34} \\ \hline A_{14}^T & A_{24}^T & A_{34}^T & A_{44} \end{array}\right] \qquad \mathcal{P}_1 = \left[\begin{array}{cccc} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & A_{23} & \\ A_{13}^T & A_{23}^T & A_{33} & \\ & & & A_{44} \end{array}\right]$$

Theorem (PMHTH 16)

$\mathcal{P}_1^{-1}\mathcal{A}$ has $4n - 2r$ eigenvalues equal to 1 and any remaining eigenvalues λ satisfy

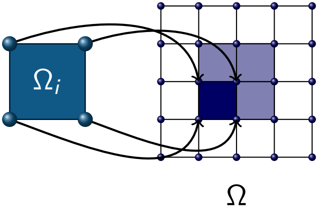
$$0 < 1 - \sqrt{\mu_{max}} \leq \lambda \leq 1 + \sqrt{\mu_{max}} < 2.$$

$\lambda_{max}(\mathcal{P}_1^{-1}\mathcal{A}) < 2$ **but the smallest eigenvalue could be arbitrarily small**

A closer look

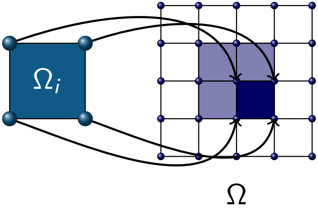
The assembly process

$$\mathcal{A}_{ij} = \int_{\Omega} \nabla^2 \phi_i \, \nabla^2 \phi_j = \sum_{i=1}^{M} \int_{\Omega_i} \nabla^2 \phi_i \, \nabla^2 \phi_j$$



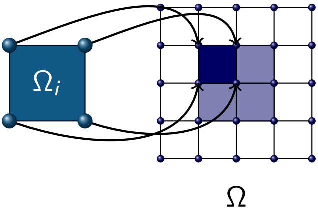
The assembly process

$$\mathcal{A}_{ij} = \int_{\Omega} \nabla^2 \phi_i \, \nabla^2 \phi_j = \sum_{i=1}^{M} \int_{\Omega_i} \nabla^2 \phi_i \, \nabla^2 \phi_j$$



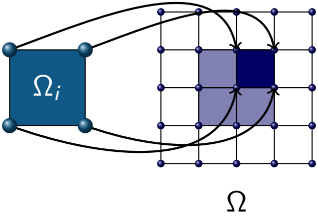
The assembly process

$$\mathcal{A}_{ij} = \int_{\Omega} \nabla^2 \phi_i \, \nabla^2 \phi_j = \sum_{i=1}^{M} \int_{\Omega_i} \nabla^2 \phi_i \, \nabla^2 \phi_j$$

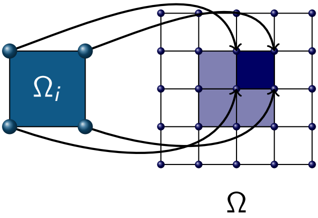


The assembly process

$$\mathcal{A}_{ij} = \int_{\Omega} \nabla^2 \phi_i \, \nabla^2 \phi_j = \sum_{i=1}^{M} \int_{\Omega_i} \nabla^2 \phi_i \, \nabla^2 \phi_j$$



The assembly process



To assemble the matrix $\mathcal{A}$, we form

$$\mathcal{A} = L^T \begin{bmatrix} A_{e_1} & & \\ & A_{e_2} & \\ & & \ddots \\ & & & A_{e_M} \end{bmatrix} L$$

Idea behind the analysis

$$\mathcal{A} = L^T \operatorname{diag}(A_e) L, \qquad \mathcal{P}_1 = L^T \operatorname{diag}(P_e) L$$

Idea: use the **element matrices** for eigenvalue analysis, since this is independent of the mesh size (Wathen 1987, 1992)

The smallest eigenvalue

$$\mathcal{A} = L^T \operatorname{diag}(A_e) L, \qquad \mathcal{P}_1 = L^T \operatorname{diag}(P_e) L$$

Variational characterisation of eigenvalues

$$\lambda_{\min}(\mathcal{P}_1^{-1}\mathcal{A}) = \min_{\boldsymbol{x} \neq 0} \frac{\boldsymbol{x}^T \mathcal{A} \boldsymbol{x}}{\boldsymbol{x}^T \mathcal{P}_1 \boldsymbol{x}}$$

The smallest eigenvalue

$$\mathcal{A} = L^T \operatorname{diag}(A_e) L, \qquad \mathcal{P}_1 = L^T \operatorname{diag}(P_e) L$$

Variational characterisation of eigenvalues

$$\lambda_{\min}(\mathcal{P}_1^{-1}\mathcal{A}) = \min_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}^T \mathcal{A} \boldsymbol{x}}{\boldsymbol{x}^T \mathcal{P}_1 \boldsymbol{x}} = \min_{\substack{\boldsymbol{y}=L\boldsymbol{x},\\ \boldsymbol{x}\neq 0}} \frac{\boldsymbol{y}^T \operatorname{diag}(A_e) \boldsymbol{y}}{\boldsymbol{y}^T \operatorname{diag}(P_e) \boldsymbol{y}}$$

Can just examine each 16×16 generalised eigenvalue problem

$$A_e \boldsymbol{z} = \mu P_e \boldsymbol{z}$$

A sticky situation

The smallest eigenvalue

$$\lambda_{\min}(\widehat{\mathcal{P}}_{BD}^{-1}\mathcal{A}) = \min_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}^T\mathcal{A}\boldsymbol{x}}{\boldsymbol{x}^T\mathcal{P}_1\boldsymbol{x}} = \min_{\substack{\boldsymbol{y}=L\boldsymbol{x},\\ \boldsymbol{x}\neq 0}} \frac{\boldsymbol{y}^T \operatorname{diag}(A_e)\boldsymbol{y}}{\boldsymbol{y}^T \operatorname{diag}(P_e)\boldsymbol{y}}$$

- Both A_e and P_e are singular
- $\operatorname{null}(\operatorname{diag}(P_e)) \subset \operatorname{null}(\operatorname{diag}(A_e))$

The smallest eigenvalue

$$\lambda_{\min}(\widehat{\mathcal{P}}_{BD}^{-1}\mathcal{A}) = \min_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}^T\mathcal{A}\boldsymbol{x}}{\boldsymbol{x}^T\mathcal{P}_1\boldsymbol{x}} = \min_{\substack{\boldsymbol{y}=L\boldsymbol{x},\\ \boldsymbol{x}\neq 0}} \frac{\boldsymbol{y}^T\operatorname{diag}(A_e)\boldsymbol{y}}{\boldsymbol{y}^T\operatorname{diag}(P_e)\boldsymbol{y}}$$

- Both A_e and P_e are singular
- $\operatorname{null}(\operatorname{diag}(P_e)) \subset \operatorname{null}(\operatorname{diag}(A_e))$

$$\frac{\boldsymbol{y}^T\operatorname{diag}(A_e)\boldsymbol{y}}{\boldsymbol{y}^T\operatorname{diag}(P_e)\boldsymbol{y}} = \frac{\boldsymbol{y}_R{}^T\operatorname{diag}(A_e)\boldsymbol{y}_R}{(\boldsymbol{y}_R+\boldsymbol{y}_M)^T\operatorname{diag}(P_e)(\boldsymbol{y}_R+\boldsymbol{y}_M)}$$

$\mathcal{R} = \operatorname{range}(\operatorname{diag}(A_e))$ $\qquad$ $\mathcal{M} = \operatorname{null}(\operatorname{diag}(A_e)) \cap (\operatorname{null}(\operatorname{diag}(P_e)))^\perp$

The smallest eigenvalue

$$\lambda_{\min}(\widehat{\mathcal{P}}_{BD}^{-1}\mathcal{A}) = \min_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}^T\mathcal{A}\boldsymbol{x}}{\boldsymbol{x}^T\mathcal{P}_1\boldsymbol{x}} = \min_{\substack{\boldsymbol{y}=L\boldsymbol{x},\\ \boldsymbol{x}\neq 0}} \frac{\boldsymbol{y}^T\,\mathrm{diag}(A_e)\boldsymbol{y}}{\boldsymbol{y}^T\,\mathrm{diag}(P_e)\boldsymbol{y}}$$

$$\frac{\boldsymbol{y}^T\,\mathrm{diag}(A_e)\boldsymbol{y}}{\boldsymbol{y}^T\,\mathrm{diag}(P_e)\boldsymbol{y}} = \frac{\boldsymbol{y}_R{}^T\,\mathrm{diag}(A_e)\boldsymbol{y}_R}{(\boldsymbol{y}_R+\boldsymbol{y}_M)^T\,\mathrm{diag}(P_e)(\boldsymbol{y}_R+\boldsymbol{y}_M)}$$

$\mathcal{R} = \mathrm{range}(\mathrm{diag}(A_e))$ $\qquad$ $\mathcal{M} = \mathrm{null}(\mathrm{diag}(A_e)) \cap (\mathrm{null}(\mathrm{diag}(P_e)))^\perp$

For all grids tested $\boldsymbol{y}_R{}^T\,\mathrm{diag}(P_e)\boldsymbol{y}_R > \boldsymbol{y}_M{}^T\,\mathrm{diag}(P_e)\boldsymbol{y}_M$ for any $\boldsymbol{y} = L\boldsymbol{x}$

Pulling things together

Theorem (PMHTH 16)

Assuming that $\boldsymbol{y}_R{}^T \operatorname{diag}(P_e)\boldsymbol{y}_R > \boldsymbol{y}_M{}^T \operatorname{diag}(P_e)\boldsymbol{y}_M$, *the eigenvalues* λ *of* $\mathcal{P}_1^{-1}\mathcal{A}$ *satisfy*

$$0.0118 \leq \lambda < 2.$$

Similar results hold for stretched grids.

For comparison:

	4×4	8×8	16×16	32×32
$\lambda_{min}(\mathcal{P}_1^{-1}\mathcal{A})$	0.72	0.64	0.61	0.60
$\lambda_{max}(\mathcal{P}_1^{-1}\mathcal{A})$	1.28	1.36	1.39	1.40

A more practical preconditioner

Second preconditioner

$$\mathcal{P}_1 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & A_{23} & \\ A_{13}^T & A_{23}^T & A_{33} & \\ & & & A_{44} \end{bmatrix} \qquad \mathcal{P}_2 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & & \\ A_{13}^T & & A_{33} & \\ & & & A_{44} \end{bmatrix}$$

	4×4	8×8	16×16	32×32
$\lambda_{min}(\mathcal{P}_2^{-1}\mathcal{A})$	0.72	0.62	0.58	0.56
$\lambda_{max}(\mathcal{P}_2^{-1}\mathcal{A})$	1.28	1.38	1.40	1.41

Second preconditioner

$$\mathcal{P}_1 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & A_{23} & \\ A_{13}^T & A_{23}^T & A_{33} & \\ & & & A_{44} \end{bmatrix} \qquad \mathcal{P}_2 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \\ A_{12}^T & A_{22} & & \\ A_{13}^T & & A_{33} & \\ & & & A_{44} \end{bmatrix}$$

	4×4	8×8	16×16	32×32
$\lambda_{min}(\mathcal{P}_2^{-1}\mathcal{A})$	0.72	0.62	0.58	0.56
$\lambda_{max}(\mathcal{P}_2^{-1}\mathcal{A})$	1.28	1.38	1.40	1.41

- Can bound eigenvalues by combining bounds for $\mathcal{P}_2^{-1}\mathcal{P}_1$ and $\mathcal{P}_1^{-1}\mathcal{A}$
- These are complicated

Implementation

$$\mathcal{P}_2 = UL = \begin{bmatrix} I & A_{12}A_{22}^{-1} & A_{13}A_{33}^{-1} & \\ & I & & \\ & & I & \\ & & & I \end{bmatrix} \begin{bmatrix} S_{11} & & & \\ A_{12}^T & A_{22} & & \\ A_{13}^T & & A_{33} & \\ & & & A_{44} \end{bmatrix},$$

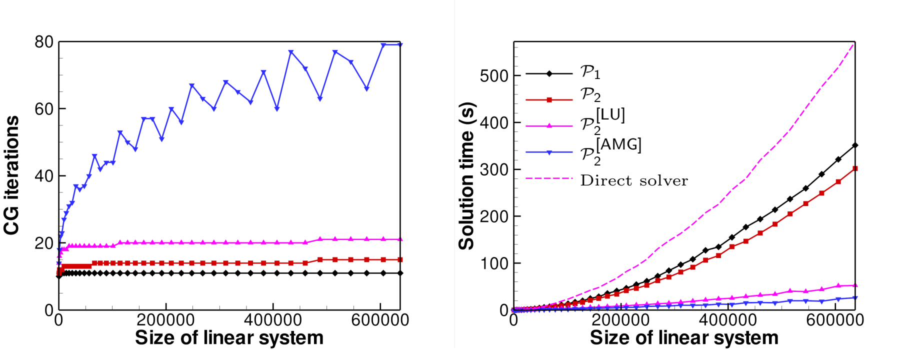
where

$$S_{11} = A_{11} - A_{12}A_{22}^{-1}A_{12}^T - A_{13}A_{33}^{-1}A_{13}^T$$

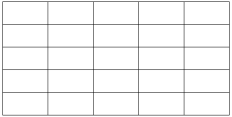
- Can lump A_{22}, A_{33}, A_{44}, or use diagonal, without greatly affecting eigenvalues/bounds
- Can approximate S_{11} using multigrid

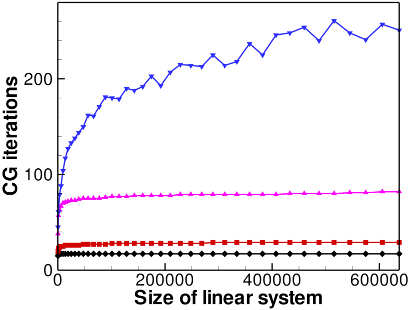
Results

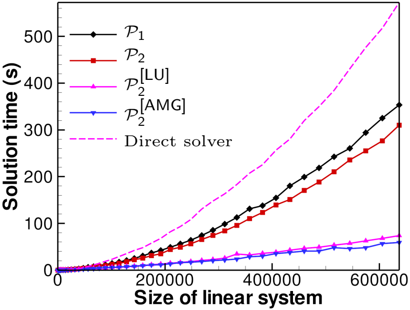
Square elements



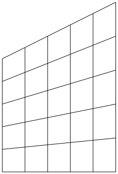
Stretched elements

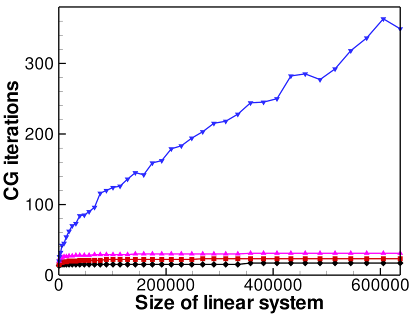


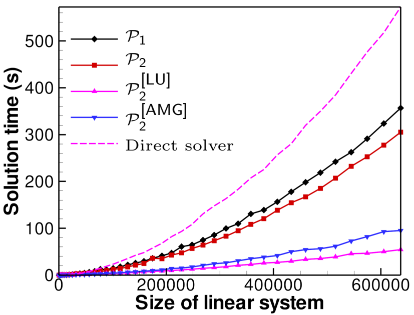




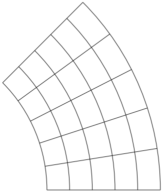
Distorted elements

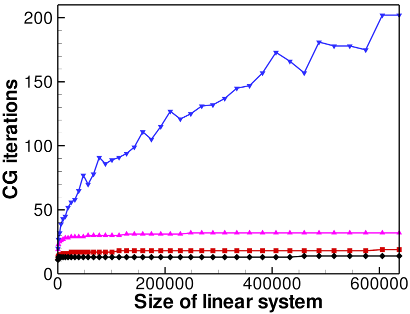


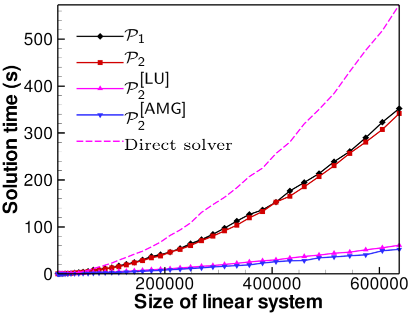




Non-convex domain







Conclusions

Conclusions

- Preconditioners based on blocking $\mathcal{A}$ according to basis function type can be very effective
- Both $\mathcal{P}_1$ and $\mathcal{P}_2$ seem to give mesh independent convergence; proved for $\mathcal{P}_1$
- Preconditioners robust to stretching and distortion of elements
- Some loss of robustness of AMG version but still the fastest!

Future work

- More complete analysis of $\mathcal{P}_1$ and $\mathcal{P}_2$
- Alternative block preconditioners
- More robust AMG

Thank you!

References

J. Pestana *et al.*, "Efficient block preconditioning for a C^1 finite element discretization of the Dirichlet biharmonic problem," *SIAM J. Sci. Comput.*, vol. 38, A325–A345, 2016.

A. J. Wathen, "Spectral bounds and preconditioning methods using element-by-element analysis for Galerkin finite element equations," in *The Mathematics of Finite Elements and Applications VI, MAFELAP 1987*, J. R. Whiteman, Ed., London, UK: Academic Press, 1988, pp. 157–168.

——, "Singular element preconditioning for the finite element method," in *Iterative Methods in Linear Algebra*, R. Beauwens and P. de Groen, Eds., Amsterdam: North Holland, 1992, pp. 531–540.