Preconditioners for Variable Viscosity Stokes Flow

Dave May; dave.mayhem23@gmail.com
Louis Moresi; louis.moresi@gmail.com

MONASH University

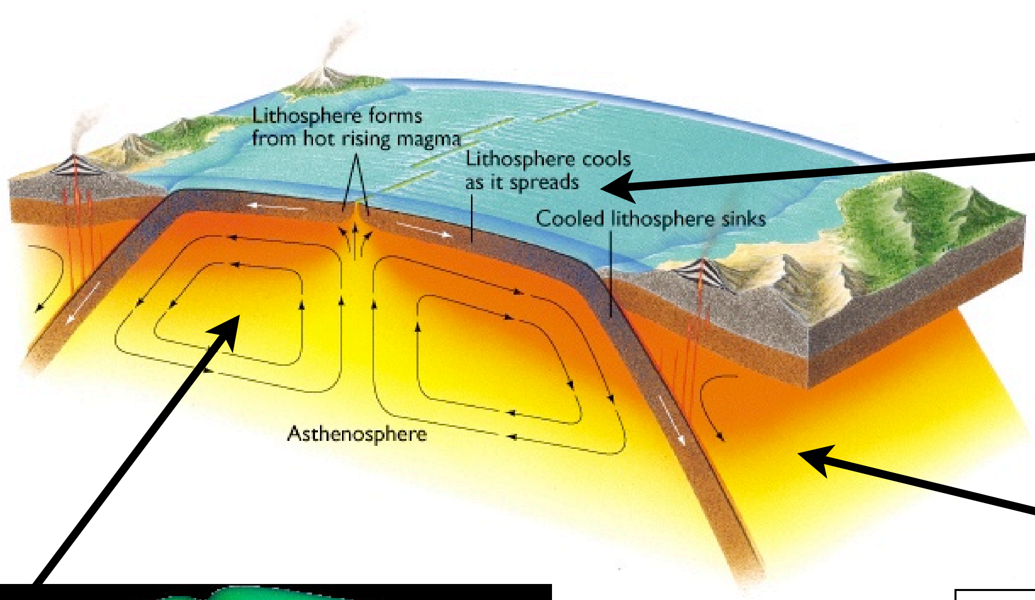
School of Mathematical Sciences
Monash Cluster Computing

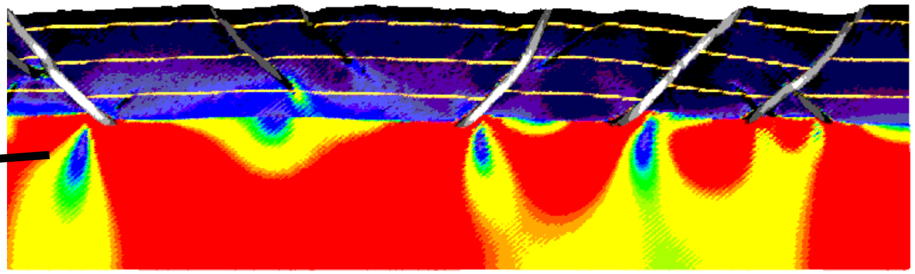
CIG - GeoMath
Santa Fe, NM
September 15, 2008

Outline

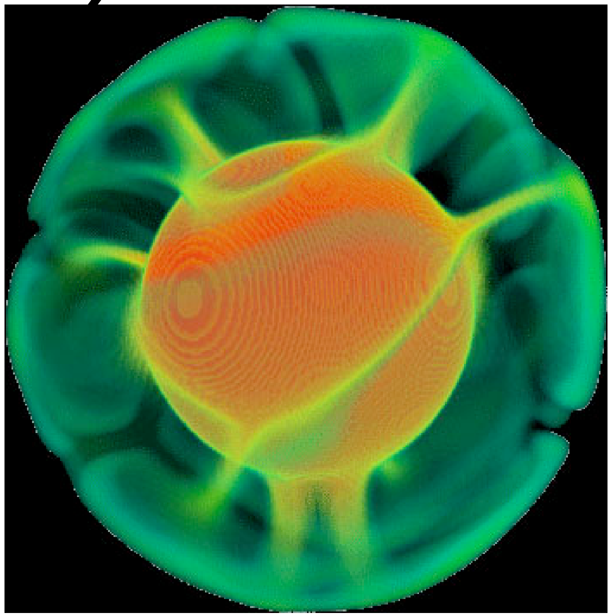
- Background
- Two solution strategies
- Preconditioners
- Some results

Motivating Examples

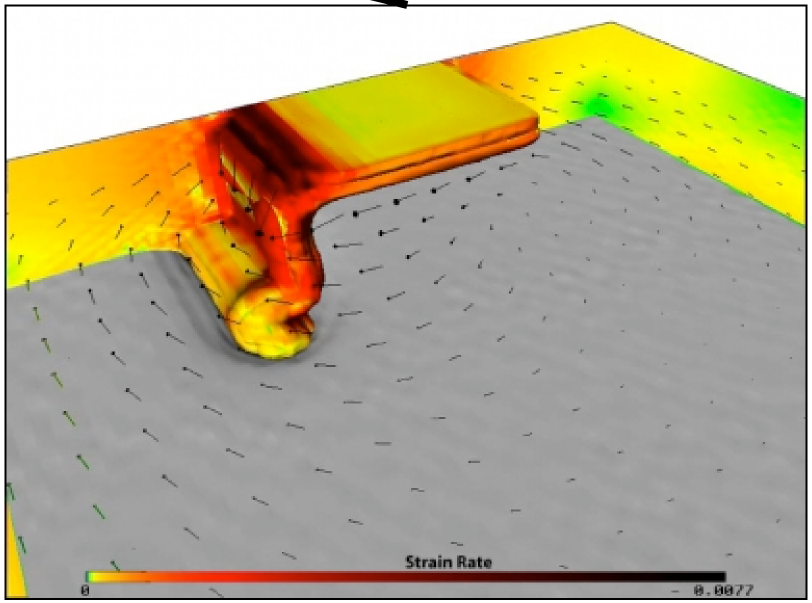




Lithospheric deformation



Convection in the mantle



Subduction

Continuum Description

- Incompressible Stokes flow with general constitutive tensor.

Momentum $\tau_{ij,j} - p_{,i} + f_i = 0,$

Mass $-u_{i,i} = 0,$

$$\left.\begin{array}{l}\tau_{ij,j} - p_{,i} + f_i = 0, \\ -u_{i,i} = 0,\end{array}\right\} \text{ in } \Omega$$

subject to

i) the boundary conditions

$$\begin{aligned} u_i &= g_i, && \text{on } \Gamma_{g_i} \\ \sigma_{ij} n_j &= h_i, && \text{on } \Gamma_{h_i}, \end{aligned}$$

ii) the pressure constraint

$$\int_\Omega p\, dV = p_s,$$ for some constant p_s

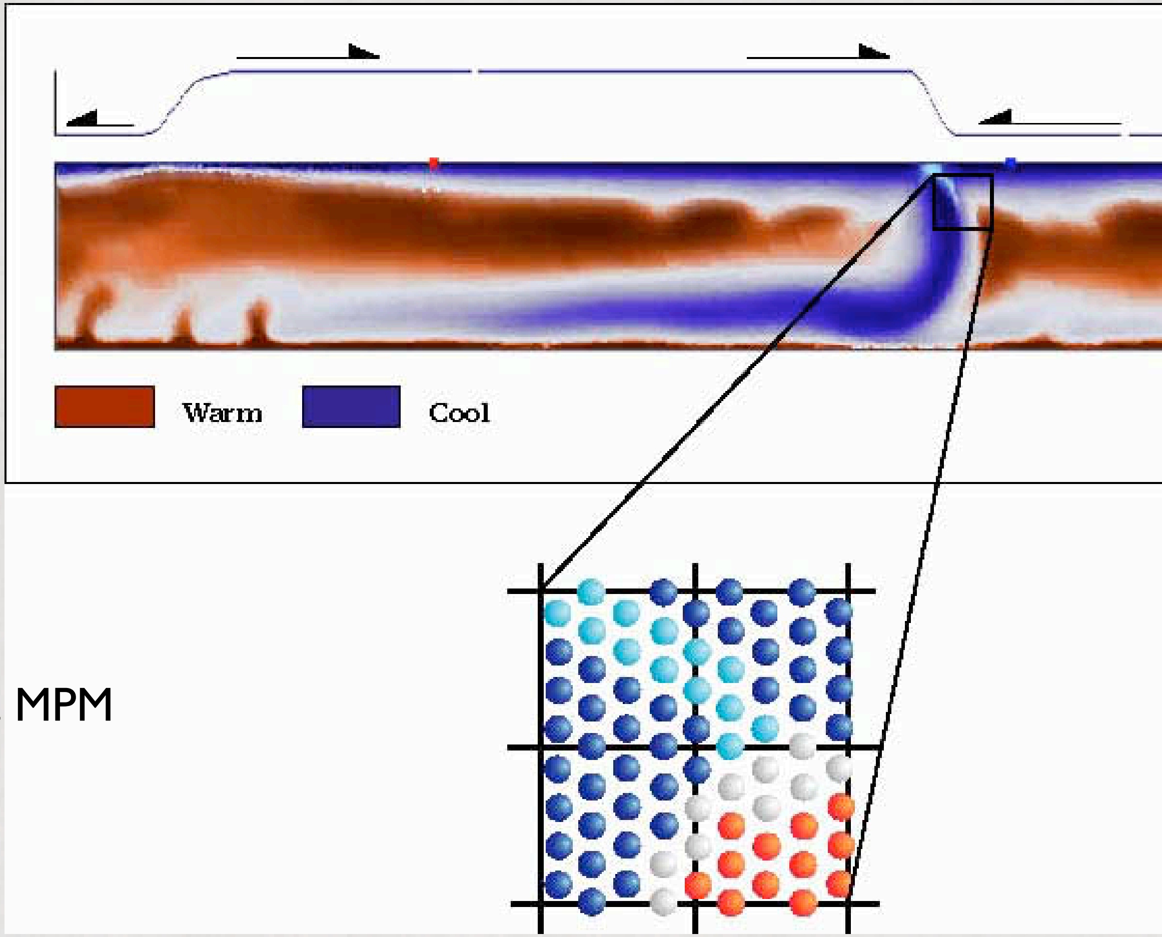
Constitutive $\tau_{ij} = \Lambda_{ijkl} \dot{\epsilon}_{kl},$

Strain rate $\dot{\epsilon}_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}),$

We formulate the problem entirely in terms of velocity u, and pressure p.

Discrete Stokes Flow Formulation

In the style of the Material Point Method...
"a fixed mesh with moving particles"



- Mixed Finite Element discretisation (Q1-P0)
  - Inherits robustness, versatility of FEM
  - Admits general constitutive relations
- Material properties tracked via particles - aka MPM
  - Compositional tracking
  - *Stress- history tensor*
  - *Plastic strain history (scalar / tensorial)*
  - *Material orientation (anisotropy)*

This is the approach adopted in Underworld
www.mcc.monash.edu.au

Discrete Stokes Flow Formulation

The connection between the FE formulation and the Lagrangian points is via evaluation of the weak form.

$$\mathbf{K}^E = \int_{\Omega_E} \mathbf{B}^T(\mathbf{x})\mathbf{C}(\mathbf{x})\mathbf{B}(\mathbf{x})d\Omega$$

$$\mathbf{K}^E = \sum_p w_p \mathbf{B}_p^T(\mathbf{x}_p)\mathbf{C}_p(\mathbf{x}_p)\mathbf{B}_p(\mathbf{x}_p)$$

Lagrangian points coincide with the quadrature points used to evaluate the weak form
- quadrature weights are defined locally over each element.
- weights are given by an approximate voronoi diagram.

The constitutive behaviour is associated with each particle "p" and is thus naturally incorporated into the quadrature.

The discrete form of the momentum and continuity generated by the FEM may be expressed as

$$\begin{pmatrix} K & G \\ G^T & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ h \end{pmatrix} \qquad K \in \mathbb{R}^{m \times m}, \quad G \in \mathbb{R}^{m \times n}, \quad m > n$$

The MPM formulation provided the means to discretize the Stokes flow.
But the issue of obtaining the flow field (u,p) is now a linear algebra problem...

Iterative methods for Discrete Stokes Flow

The ideal approach should be *optimal* in the sense that the convergence rate of method will be bounded independently of

- the discretisation parameters *(Example; grid resolution)*
- the constitutive parameters *(Example; smoothly varying vs. discontinuous viscosity*
- the constitutive behaviour *(Example; isotropic vs. anisotropic)*
- the solution is obtained in *O(n)* time... ie. multigrid

These are a challenging set of requirements !

Schur Complement Reduction

- Decouple u and p

- Solve the Schur complement system

$$\text{solve for } p: \quad (G^T K^{-1} G)p = G^T K^{-1} f - h,$$
$$\text{solve for } u: \quad Ku = f - Gp.$$

where $S = G^T K^{-1} G$ is the Schur complement.

- Represent S as a matrix-free object. To compute $y = Sx$ we

$$\text{compute:} \quad f^* = Gx,$$
$$\text{solve for } u^*: \quad Ku^* = f^*,$$
$$\text{compute:} \quad y = G^T u^*.$$

- Outer Krylov iterations performed on Sp=h', inner iterations performed on Ky=x.

- Need preconditioners for S and K.

Fully Coupled Approach

- Treat the Stokes problem as single coupled system

$$\begin{pmatrix} K & G \\ G^T & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ h \end{pmatrix} \quad \longrightarrow \quad \mathcal{A}x = b, \quad \mathcal{A} \in \mathbb{R}^{(m+n)\times(m+n)}$$

- Apply any suitable Krylov method to $\mathcal{A}x = b$

- We require preconditioner for $\mathcal{A}$.

- Block diagonal or block upper triangular

$$\hat{\mathcal{A}}_d = \begin{pmatrix} \hat{K} & 0 \\ 0 & -\hat{S} \end{pmatrix}, \qquad \hat{\mathcal{A}}_u = \begin{pmatrix} \hat{K} & G \\ 0 & -\hat{S} \end{pmatrix}.$$

Elman, Silvester (1994)
Rusten, Winther (1992)
Silvester, Wathan (1994)

Bramble, Pasciak (1988)
Murphy, Golub (2000)

- Both options require preconditioners for K and S.

A Block Preconditioner for K

- Billinear form of the deviatoric stress tensor gradient

$$a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega 2\eta \, \epsilon_{ij}(\boldsymbol{u})\epsilon_{ij}(\boldsymbol{v}) \, dV$$

- Discrete operator

$$\boldsymbol{K} = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}$$

- A spectrally equivalent billinear form is

$$\hat{a}(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \eta \; (\nabla u_k) \cdot (\nabla v_k) \; dV$$

[Axelsson, Padiy, "*On a robust and scalable linear elasticity solver based on a saddle point formulation*"]

with discrete operator given by

$$\hat{\boldsymbol{K}} = \begin{pmatrix} K_{11} & 0 \\ 0 & K_{22} \end{pmatrix}$$

A Block Preconditioner for K

- Block Gauss-Seidel

$$Ax = b$$

$$x^{k+1} = (D+L)^{-1}\left(b - Ux^k\right)$$

$$A = D + L + U$$

- Discrete counterpart of the stress gradient is given by

$$\begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f_x \\ f_y \end{pmatrix}$$

$$D = \mathrm{diag}\left[K_{11}, K_{22}\right]$$

- Treat each velocity component as scalar, variable coefficient diffusion problem.

- Each scalar problem permits effective multigrid preconditioning.

Each K_{ii}^{-1} given by CG, with $\frac{\|r_k\|}{\|r_0\|} < 10^{-2}$, preconditioned via *ML*.

A Block Preconditioner for K

- Two dimensional code using Q1-P0. Solve $Ku = f, \quad \Omega = [0,1] \times [0,1]$
- **Sensitivity of iterations to grid and viscosity contrast**

(i) $\eta = 10^6 \exp(\theta x)$

elements	θ $(\Delta\eta)$		
$M \times N$	$4.6\ (10^2)$	$13.8\ (10^6)$	$18.4\ (10^8)$
100^2	4	4	4
200^2	4	4	4
300^2	4	4	4

(ii) $\eta = \begin{cases} 1 & x < \frac{1}{2} \\ \Delta\eta & x \geq \frac{1}{2} \end{cases}$

elements	$\Delta\eta$		
$M \times N$	10^2	10^6	10^{10}
100^2	5	5	5
200^2	5	5	5
300^2	5	5	5

A Block Preconditioner for K

- **Timings**

$$\eta = 10^6 \exp(\theta x)$$

$$\theta = 13.8,\ \Delta\eta = 10^6$$

$M \times N$	unknowns	its.	CPU time (sec)
200^2	80,802	4	7.2
284^2	162,450	4	12.7
402^2	324,818	4	25.8

- In short, the block Gauss-Seidel preconditioner looks very promising.
- It is scalable, robust and exhibits ~ *O(n)* solution times when combined with *ML*.
- It's effectiveness for 3D problems *(u,v,w)* needs to be examined.
- Parallel efficiency needs to be explored. Others have demonstrated this with *ML*.

Preconditioner for S

- Consider using commutators to construct an approximation for
[Kay, et. al, *SIAM J. Sci. Comput.*, **24** (2002)]

For isoviscous Stokes, the commutator relation used is;

$$Z_c = \left(\eta\nabla^2\right)\nabla - \nabla\left(\eta\nabla^2\right) = 0$$

For variable viscosity the continuous operators will not commutator.
Instead we will utilize the the analgous discrete commutator of *Zc*

$$Z = KG - GK_p \qquad Z \in \mathbb{R}^{m \times n}$$

[Elman, et. al, *SIAM J. Sci. Comput.*, **27** (2006)]

Preconditioner for S

- To define the Schur complement, pre-multiply discrete commutator Z by $G^T K^{-1}$

$$G^T G - \left(G^T K^{-1} G\right) K_p = G^T K^{-1} Z$$

- A little manipulation yields

$$L_p - S K_p = G^T K^{-1} Z \qquad , L_p = G^T G$$

$$S = \left(L_p - G^T K^{-1} Z\right) K_p^{-1}$$

- And finally we have

$$S^{-1} = K_p \left(L_p - G^T K^{-1} Z\right)^{-1}$$

Preconditioner for S

- Define the operator *Kp* in a least squares manner assuming *Z ~ 0*.

$$GK_p = KG$$

$$G^T G K_p = G^T K G$$

$$K_p^* = \left(G^T G\right)^{-1} G^T K G$$

$$\begin{aligned} S^{-1} &\approx K_p^* L_p^{-1} \\ &= L_p^{-1} G^T K G L_p^{-1} \\ &= S_b^{-1} \end{aligned}$$

+ No inversion of K required

+ Coupling in K preserved

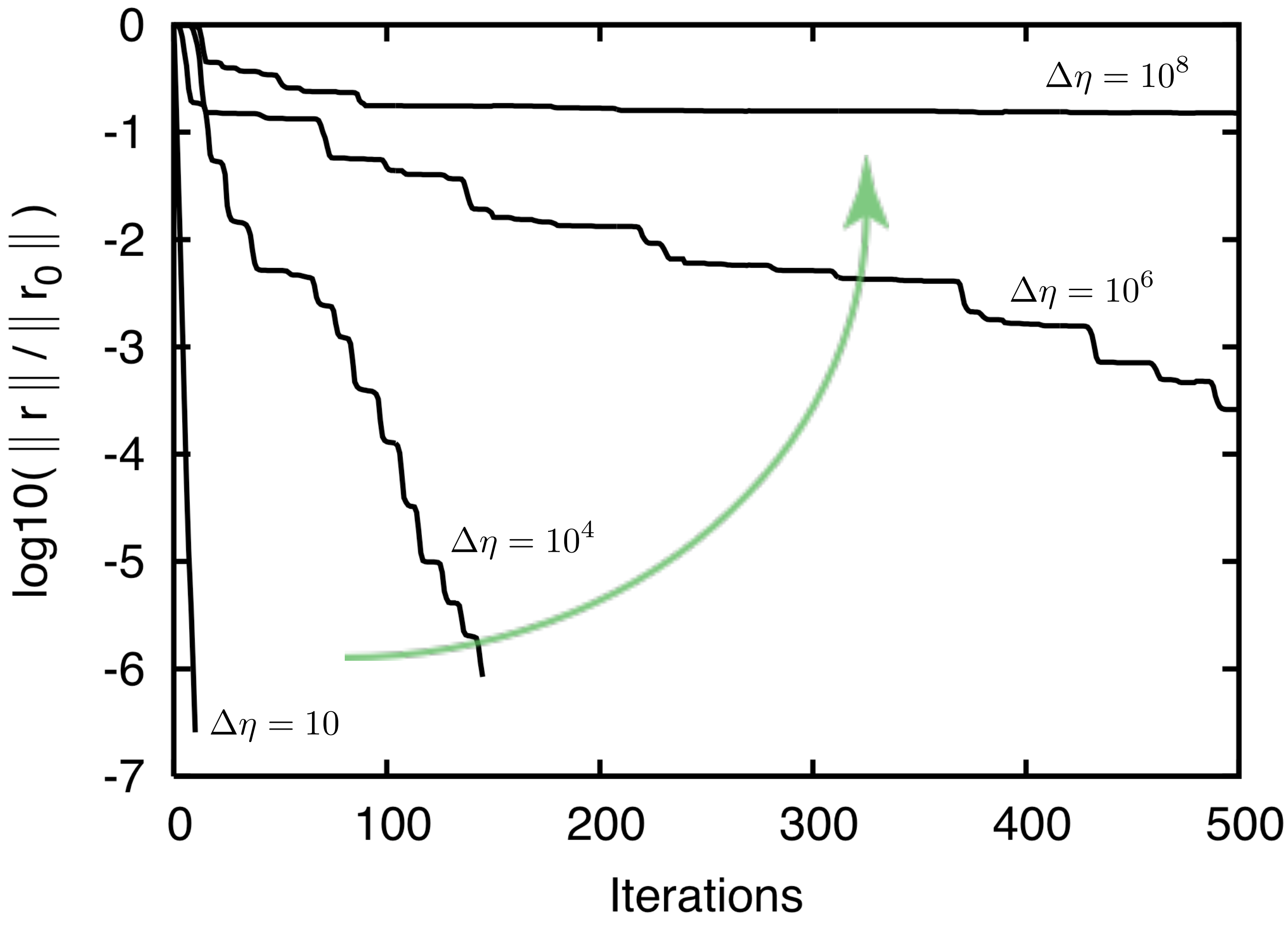
- Two Poisson solves required

+ Most problems require low precision Poisson solves

which is the BFBt preconditioner from; Elman, *SIAM J. Sci. Comput.*, **17** (1996)

In practice, the assumption *Z~0* is usually not true. As a result the approximate Schur complement inverse will **not** be close to the true Schur complement.

Break Down of the Commutator Assumption



Convergence deteriorates as the viscosity contrast increase

Scaled BFBt

- Some observations...
- Lets suppose

$$K = I_m$$

As $K \to I_m$, we have:

$$K_p^* \to I_n,$$
$$Z \to 0,$$
$$S_b^{-1} \to S^{-1}.$$

a) $$\begin{aligned} S &= G^T K^{-1} G \\ &= G^T G \end{aligned}$$

b) $$\begin{aligned} K_p^* &= L_p^{-1} G^T K G \\ &= L_p^{-1} G^T G \\ &= I_n \end{aligned}$$

c) $$\begin{aligned} Z &= KG - GK_p^* \\ &= I_m G - G I_n \\ &= 0 \end{aligned}$$

d) $$\begin{aligned} S_b^{-1} &= K_p^* L_p^{-1} \\ &= I_n L_p^{-1} \\ &= S^{-1} \end{aligned}$$

Scaled BFBt

- Previous observations motivate us to apply a scaling operation to (i) better condition the operator *A*, and (ii) to help drive the discrete commutator towards 0.

- We have chosen the following scaling to preserve symmetry:

$$\mathcal{X}\mathcal{A}\mathcal{X}^T = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} K & G \\ G^T & 0 \end{pmatrix} \begin{pmatrix} X_1^T & 0 \\ 0 & X_2^T \end{pmatrix}$$

$$= \begin{pmatrix} \bar{K} & \bar{G} \\ \bar{G}^T & 0 \end{pmatrix} = \bar{\mathcal{A}}$$

where

$$\bar{K} = X_1 K X_1^T,$$
$$\bar{G} = X_1 K X_2^T.$$

- We wish to build an X1 such that we have: $\bar{K} \approx I_m$

- For the sake of scaling rather than preconditioning we, prescribe *X* to be diagonal.

- The scaling we use is given by:

$$X_1 = \Big[\mathrm{diag}(K)\Big]^{-1/2}$$

Scaled BFBt

- The scaled BFBt preconditioner is given by:

$$S_{sb}^{-1} = \bar{L}_p^{-1} \bar{G}^T \bar{K} \bar{G} \bar{L}_p^{-1}$$

- The scaled Lp operator is:

$$\begin{aligned} \bar{L}_p &= \bar{G}^T \bar{G} \\ &= X_2 \left[G^T X_1^T X_1 G \right] X_2^T \end{aligned}$$

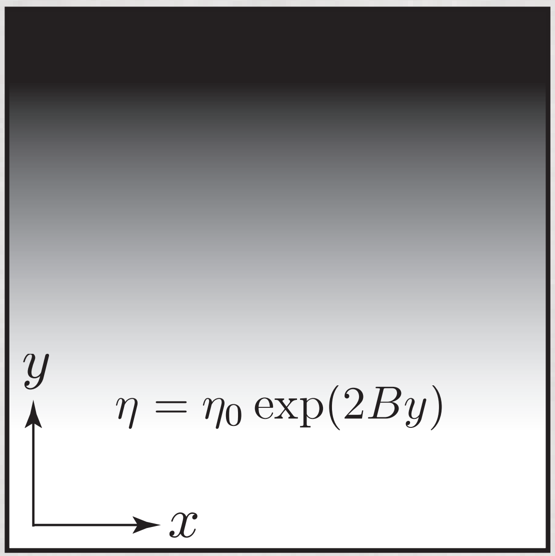
The scaling changes the discrete Poisson operator to a discrete variable-coefficient diffusion operator. The coefficient is related to the inverse of the viscosity.

The original and scaled Lp system can both be efficiently handled with multigrid. Either geometric MG or algebraic are suitable. We use *ML*.

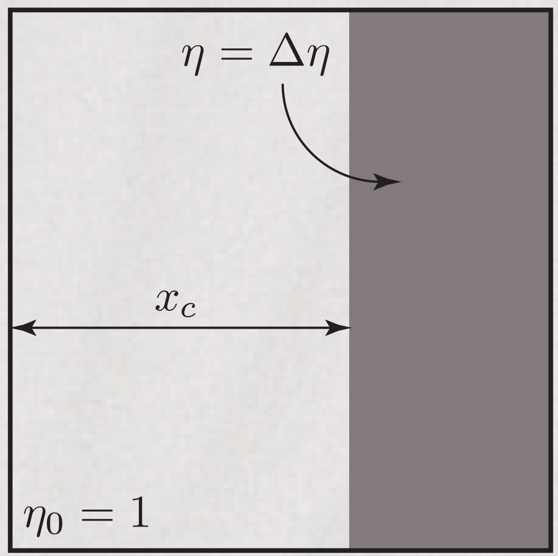
Prototype Problems

- We consider three isotropic models with different viscosity structures.
- Two dimensional code using Q1-P0 elements.

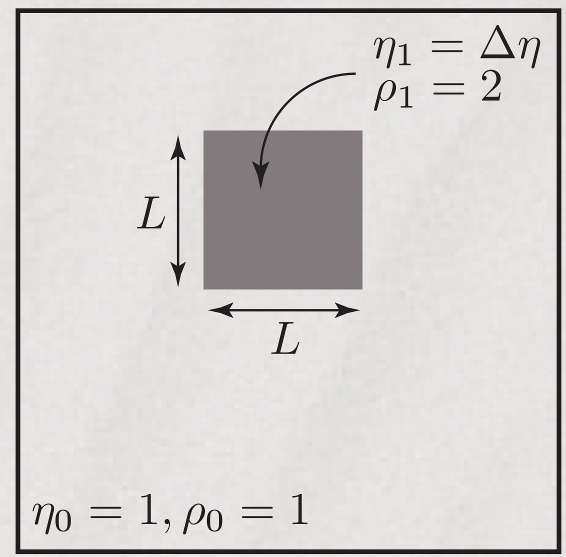
Exponentially varying with depth
"Exp(y)"



Step function in x
"Step(x)"



Step function in x & y
"Viscous sinker"



$$\text{Element resolutions} = \left\{\frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}\right\}$$

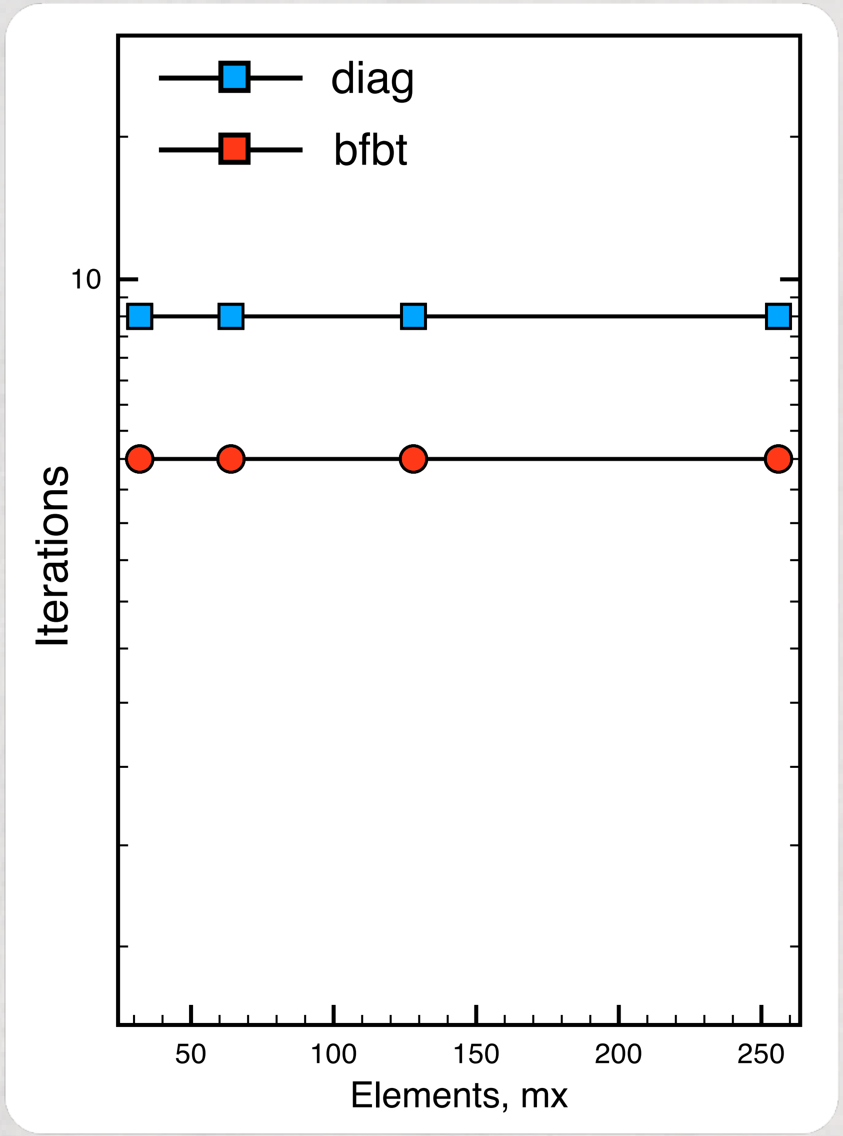
$$\text{Viscosity contrasts} = \left\{10, 10^3, 10^6, 10^8\right\}$$

Reference preconditioner for S

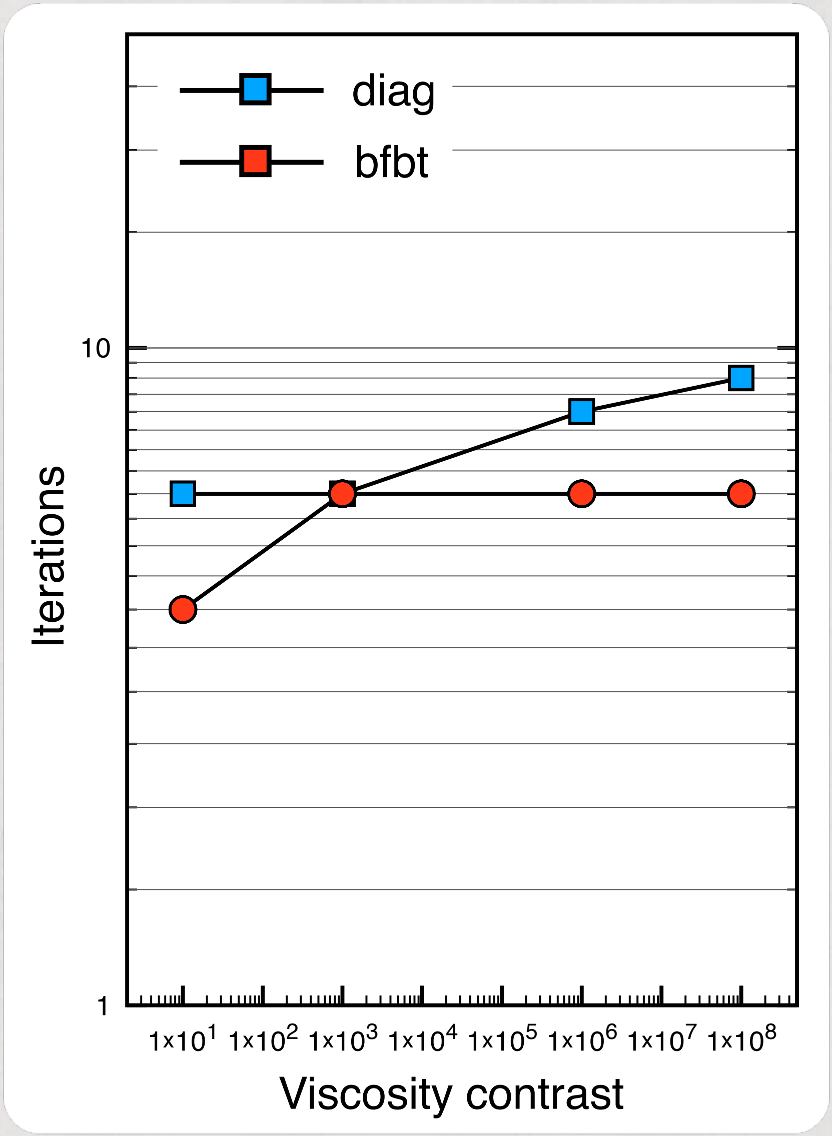
$$S_{\text{diag}}^{-1} = \text{diag}\big(G^T \text{diag}(K)^{-1} G\big)^{-1}$$

Results: "Exp(y)"

h - dependence

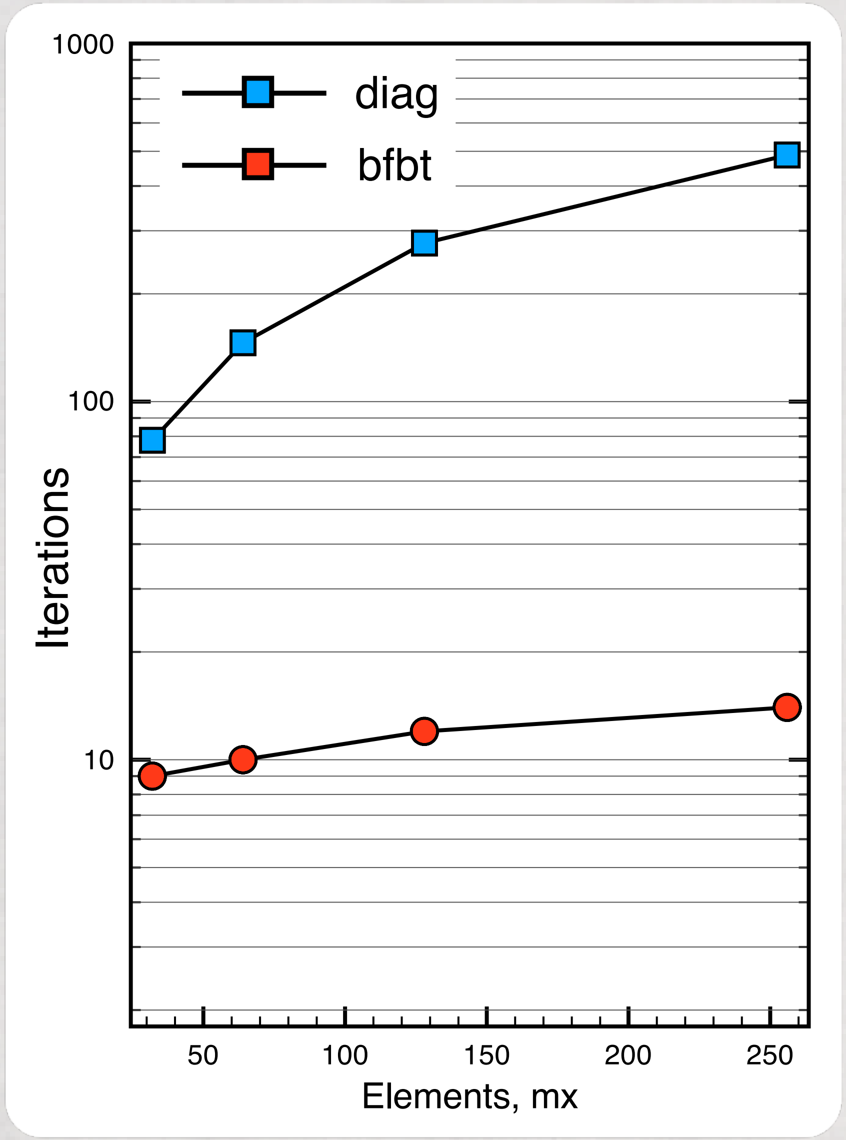


η - dependence

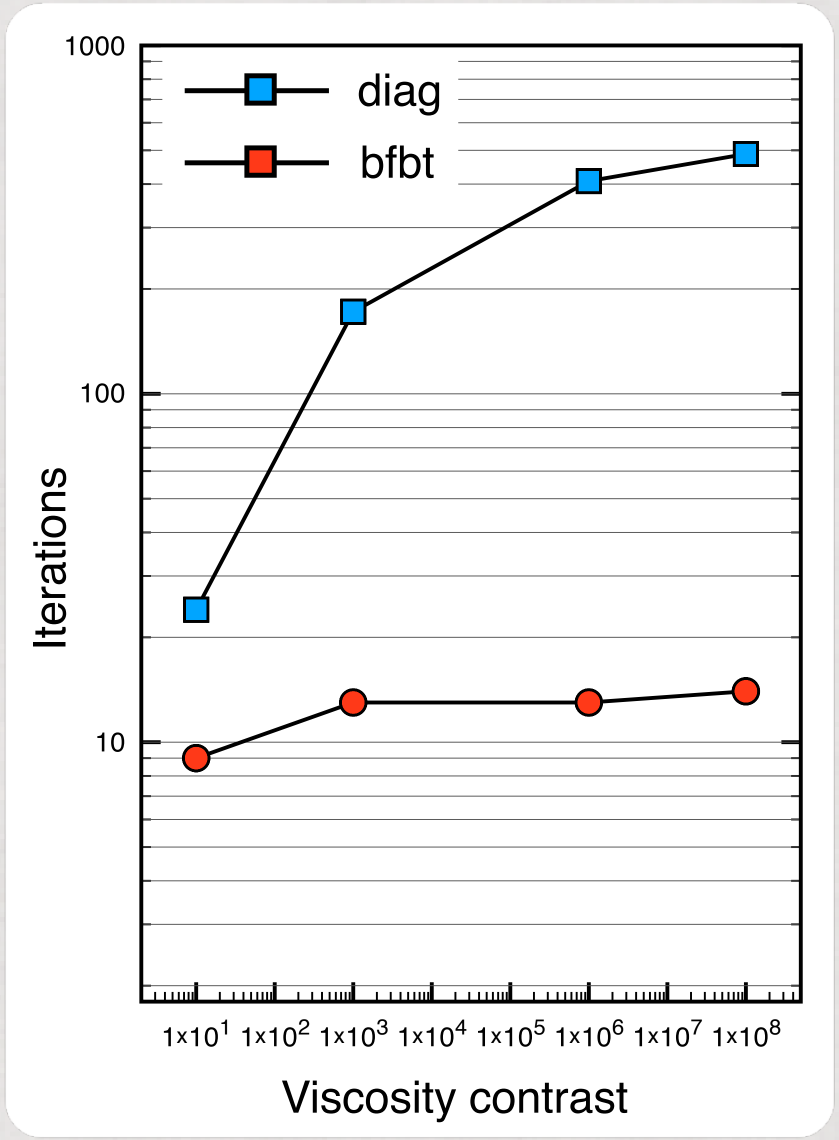


Results: "Step(x)"

h - dependence

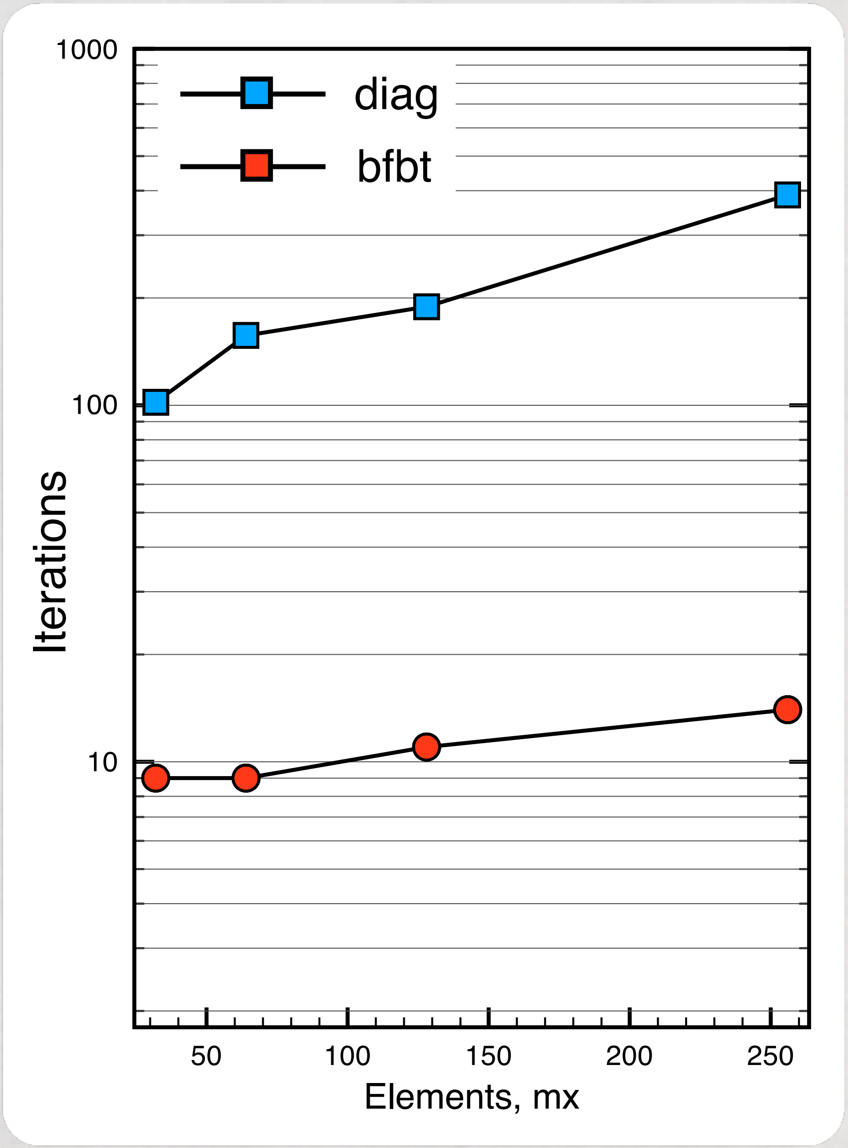


η - dependence

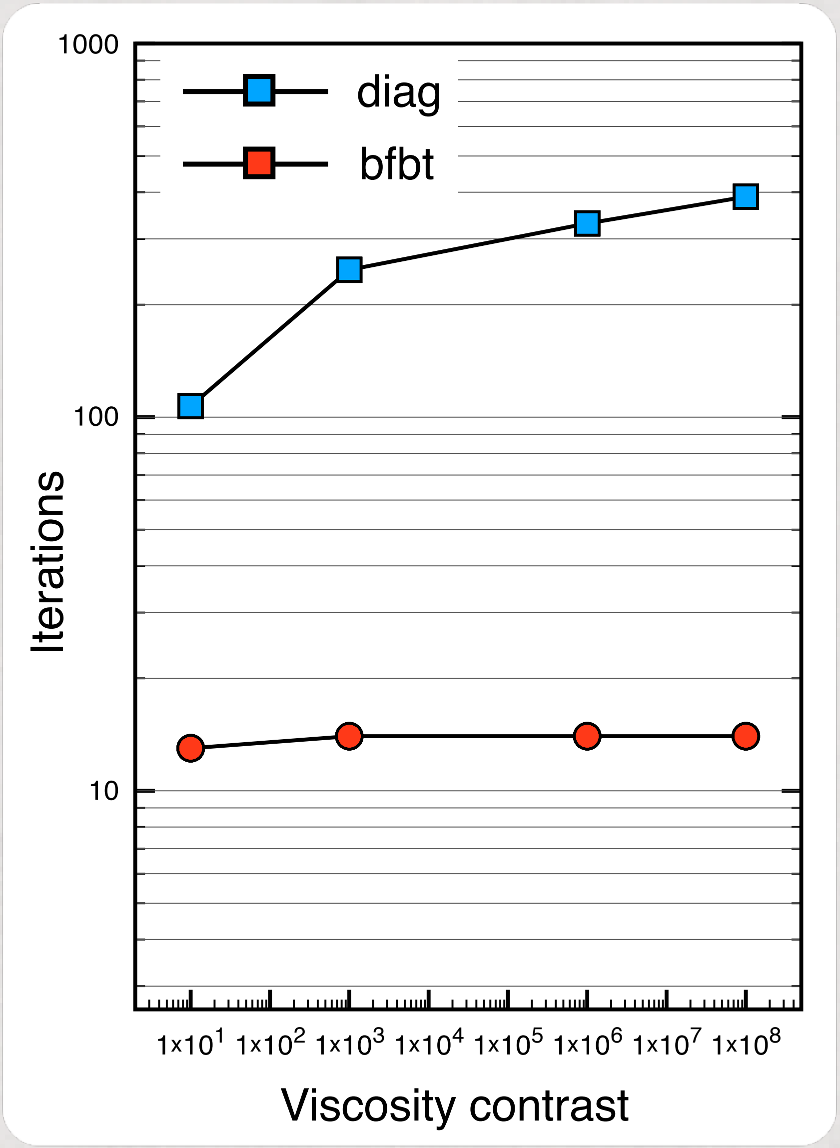


Results: "Sinker"

h - dependence



η - dependence



A Comparison

- Schur complement reduction versus fully coupled approach

"Sinker"

Iterations	SCR(FG,$\hat{S}_{sb}^{-1}$)					FC(FG,$\hat{S}_{sb}^{-1}$)				
$\Delta\eta$	10	10^3	10^6	10^8		10	10^3	10^6	10^8	
$h = 1/32$	8	9	9	9	+1	12	13	13	14	+2
$h = 1/64$	10	10	9	9	0	14	15	13	13	0
$h = 1/128$	11	11	11	11	0	16	17	16	16	0
$h = 1/256$	13	14	14	14	+1	19	20	20	21	+2
	+5	+5	+5	+5		+7	+7	+7	+7	

"Step(x)"

Iterations	SCR(FG,$\hat{S}_{sb}^{-1}$)					FC(FG,$\hat{S}_{sb}^{-1}$)				
$\Delta\eta$	10	10^3	10^6	10^8		10	10^3	10^6	10^8	
$h = 1/32$	6	8	9	9	+3	10	12	14	15	+5
$h = 1/64$	7	9	10	10	+3	11	14	14	17	+6
$h = 1/128$	8	10	12	12	+4	13	16	18	22	+9
$h = 1/256$	9	13	13	14	+5	16	19	22	27	+11
	+3	+5	+4	+5		+6	+7	+8	+13	

Summary

1) Stokes flow is an common component of numerous types of geodynamic models.

Viscosity structures used in geodynamic models can provide challenging problems for iterative methods.

2) A block Gauss-Seidel iteration combined with algebraic multigrid was shown to be robust and scalable preconditioner for the (1,1) block associated with the discretized gradient of the stress tensor.

3) The scaled BFBt preconditioner was shown to be fairly robust and scalable preconditioner for the Schur complement.

4) All preconditioners examined permit the use of multigrid.

5) Using a fully coupled (FC) approach to solve Stokes generally requires more iterations than the Schur complement reduction (SCR). The rate of the inreasing iterations required by FC w.r.t viscosity contrast and increased resolution grows faster than that observed with SCR. The extra iterations incurred by FC may be offset by using relaxed solves on the preconditioner for the (1,1) block.

We need to conduct more detailed studies to examine the effectiveness of these ideas in 3D and in a massively parallel environment. This work is presently underway within Underwold and CitcomS...

A note on software

- The following software packages were used to obtain the results presented...

PETSc: www.mcs.anl.gov/petsc

PetscExt: www.maths.monash.edu.au/~dmay/PetscExt

Underworld: www.mcc.monash.edu.au

ML (via PETSc): http://software.sandia.gov/Trilinos

Thanks for your attention