Parallel, Robust Geometric Multigrid for Adaptive High-Order Meshes and Highly Heterogeneous, Nonlinear Stokes Flow of Earth's Mantle

Johann Rudi

Institute for Computational Engineering and Sciences (ICES),
The University of Texas at Austin, USA

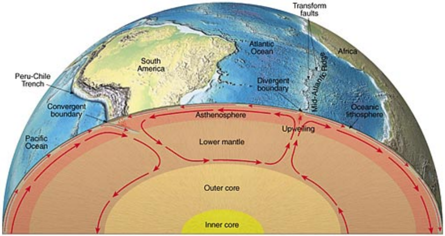
Advisor: Omar Ghattas

Main collaborators: Georg Stadler (NYU), Mike Gurnis (Caltech), Tobin Isaac (UT)

INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES

Introduction

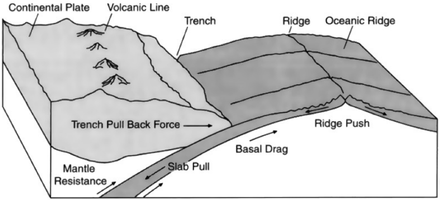
Introduction to mantle convection & plate tectonics



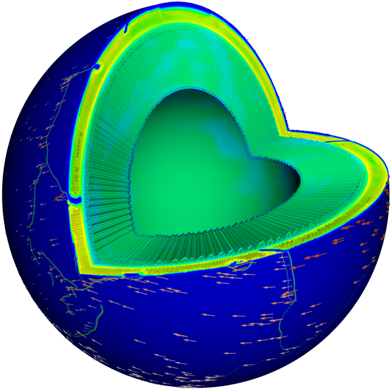
Central open questions:

- Energy dissipation in hinge zones
- Main drivers of plate motion: negative buoyancy forces or convective shear traction
- Role of slab geometries
- Accuracy of rheology extrapolations from experiments

- Mantle convection is the thermal convection in the Earth's upper $\sim$3000 km
- It controls the thermal and geological evolution of the Earth
- Solid rock in the mantle moves like viscous incompressible fluid on time scales of millions of years



Computational challenges of global-scale mantle flow



- Severe nonlinearity, heterogeneity, and anisotropy of the Earth's rheology with a wide range of spatial scales
- Highly localized features with respect to Earth's radius ($\sim$6371 km), like plate thickness $\sim$50 km and shearing zones at plate boundaries $\sim$5 km
- 6 orders of magnitude viscosity contrast within $\sim$5 km thin plate boundaries
- Resolution down to $\sim$1 km at plate boundaries (uniform mesh of Earth's mantle would result in computationally prohibitive $O(10^{12})$ degrees of freedom). Enabled by: adaptive mesh refinement
- Velocity approximation with high accuracy and local mass conservation. Enabled by: high-order discretizations

Mantle convection modeled as nonlinear Stokes flow

Rock in the mantle moves like a viscous, incompressible fluid (over millions of years) and can be modeled as a nonlinear Stokes system:

$$-\nabla \cdot \left[\mu(T, \boldsymbol{u}) \left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\top\right)\right] + \nabla p = \boldsymbol{f}(T)$$

$$\nabla \cdot \boldsymbol{u} = 0$$

The viscosity μ depends exponentially on the temperature, on a power of the second invariant of the strain rate tensor, incorporates plastic yielding and lower/upper bounds:

$$\mu(T, \boldsymbol{u}) = \max\left(\mu_{\min}, \min\left(\frac{\tau_{\text{yield}}}{2\dot{\varepsilon}(\boldsymbol{u})}, w \min\left(\mu_{\max}, a(T)\,\dot{\varepsilon}(\boldsymbol{u})^{\frac{1-n}{n}}\right)\right)\right)$$

Mantle convection modeled as nonlinear Stokes flow

Rock in the mantle moves like a viscous, incompressible fluid (over millions of years) and can be modeled as a nonlinear Stokes system:

$$-\nabla \cdot \left[\mu(T, \boldsymbol{u}) \left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\top\right)\right] + \nabla p = \boldsymbol{f}(T)$$
$$\nabla \cdot \boldsymbol{u} = 0$$

The viscosity μ depends exponentially on the temperature, on a power of the second invariant of the strain rate tensor, incorporates plastic yielding and lower/upper bounds:

$$\mu(T, \boldsymbol{u}) = \max\left(\mu_{\min}, \min\left(\frac{\tau_{\text{yield}}}{2\dot{\varepsilon}(\boldsymbol{u})}, w \min\left(\mu_{\max}, a(T)\, \dot{\varepsilon}(\boldsymbol{u})^{\frac{1-n}{n}}\right)\right)\right)$$

The Newton update $(\tilde{\boldsymbol{u}}, \tilde{p})$ is computed as the inexact solution of:

$$-\nabla \cdot \left[\left(\mu\, \mathbf{I} + \dot{\varepsilon} \frac{\partial \mu}{\partial \dot{\varepsilon}} \frac{(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\top) \otimes (\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\top)}{\|(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\top)\|_F^2}\right) (\nabla \tilde{\boldsymbol{u}} + \nabla \tilde{\boldsymbol{u}}^\top)\right] + \nabla \tilde{p} = -\boldsymbol{r}_{\text{mom}}$$
$$\nabla \cdot \tilde{\boldsymbol{u}} = -r_{\text{mass}}$$

Methods & Algorithms

Solving the discretized Stokes system

Finite element discretization:

- High-order, inf-sup stable velocity-pressure pairings: $\mathbb{Q}_k \times \mathbb{P}_{k-1}^{\mathrm{disc}}$
- Local mass conservation at the element level, discont. pressure

Solving the discretized Stokes system

Finite element discretization:

- High-order, inf-sup stable velocity-pressure pairings: $\mathbb{Q}_k \times \mathbb{P}_{k-1}^{\mathrm{disc}}$
- Local mass conservation at the element level, discont. pressure

Coupled iterative solver with upper triangular block preconditioning:

$$\underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix}}_{\text{Stokes operator}} \underbrace{\begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1}}_{\text{preconditioner}} \begin{bmatrix} \mathbf{u}' \\ \mathbf{p}' \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}$$

Requires: (i) approx. inverse of the viscous stress block, $\tilde{\mathbf{A}}^{-1} \approx \mathbf{A}^{-1}$
(ii) approx. inverse of the Schur complement, $\tilde{\mathbf{S}}^{-1} \approx (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top)^{-1}$

Solving the discretized Stokes system

Finite element discretization:

- High-order, inf-sup stable velocity-pressure pairings: $\mathbb{Q}_k \times \mathbb{P}_{k-1}^{\text{disc}}$
- Local mass conservation at the element level, discont. pressure

Coupled iterative solver with upper triangular block preconditioning:

$$\underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix}}_{\text{Stokes operator}} \underbrace{\begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1}}_{\text{preconditioner}} \begin{bmatrix} \mathbf{u}' \\ \mathbf{p}' \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}$$

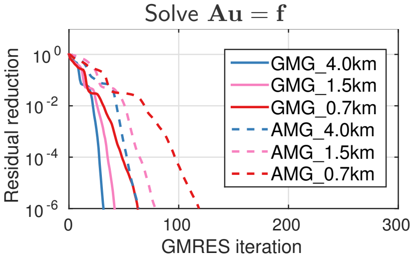
Requires: (i) approx. inverse of the viscous stress block, $\tilde{\mathbf{A}}^{-1} \approx \mathbf{A}^{-1}$
(ii) approx. inverse of the Schur complement, $\tilde{\mathbf{S}}^{-1} \approx (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top)^{-1}$

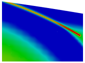
BFBT / Least Squares Commutator (LSC) method:

$$\tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^\top)^{-1}$$

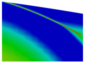
with diagonal scaling, $\mathbf{D}^{-1} := \operatorname{diag}(\mathbf{A})^{-1}$.

Comparison to state of the art for unstructured meshes

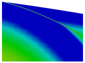




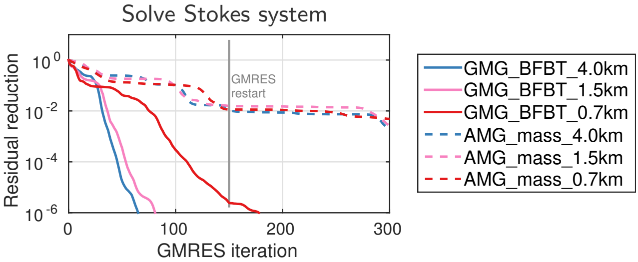
4.0 km



1.5 km



0.7 km



Review: BFBT/LSC methods for Schur complement $\tilde{\mathbf{S}}^{-1}$

BFBT method [Elman, 1999]: pseudoinverse

$$\tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top)^+ = (\mathbf{B}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{A}\mathbf{B}^\top)(\mathbf{B}\mathbf{B}^\top)^{-1}$$

Least Squares Commutators (LSC) [Elman, et al., 2006]:
Find commutator matrix $\mathbf{X}$ s.t. $(\mathbf{A}\mathbf{B}^\top - \mathbf{B}^\top\mathbf{X}) \approx \mathbf{0}$, by solving the least squares problem:

$$\text{Find columns } \mathbf{x}_j \text{ of } \mathbf{X} \text{ s.t. } \quad \min_{\mathbf{x}_j} \left\| [\mathbf{A}\mathbf{B}^\top]_j - \mathbf{B}^\top\mathbf{x}_j \right\|_2^2$$

$$\Rightarrow \ \mathbf{X} = (\mathbf{B}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{A}\mathbf{B}^\top)$$

$$(\mathbf{A}\mathbf{B}^\top - \mathbf{B}^\top\mathbf{X}) \approx \mathbf{0} \ \Rightarrow \ (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top)^{-1} \approx (\mathbf{B}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{A}\mathbf{B}^\top)(\mathbf{B}\mathbf{B}^\top)^{-1}$$

LSC gives same result for $\tilde{\mathbf{S}}^{-1}$ as pseudoinverse.

Q: Does this work for FE discretizations?...

Review: BFBT/LSC methods for Schur complement $\tilde{\mathbf{S}}^{-1}$

BFBT method [Elman, 1999]: pseudoinverse

$$\tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\top})^{+} = (\mathbf{B}\mathbf{B}^{\top})^{-1}(\mathbf{B}\mathbf{A}\mathbf{B}^{\top})(\mathbf{B}\mathbf{B}^{\top})^{-1}$$

Least Squares Commutators (LSC) [Elman, et al., 2006]:
Find commutator matrix $\mathbf{X}$ s.t. $(\mathbf{A}\mathbf{B}^{\top} - \mathbf{B}^{\top}\mathbf{X}) \approx \mathbf{0}$, by solving the least squares problem:

$$\text{Find columns } \mathbf{x}_j \text{ of } \mathbf{X} \text{ s.t. } \quad \min_{\mathbf{x}_j} \left\| [\mathbf{A}\mathbf{B}^{\top}]_j - \mathbf{B}^{\top}\mathbf{x}_j \right\|_2^2$$

$$\Rightarrow\ \mathbf{X} = (\mathbf{B}\mathbf{B}^{\top})^{-1}(\mathbf{B}\mathbf{A}\mathbf{B}^{\top})$$

$$(\mathbf{A}\mathbf{B}^{\top} - \mathbf{B}^{\top}\mathbf{X}) \approx \mathbf{0}\ \Rightarrow\ (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\top})^{-1} \approx (\mathbf{B}\mathbf{B}^{\top})^{-1}(\mathbf{B}\mathbf{A}\mathbf{B}^{\top})(\mathbf{B}\mathbf{B}^{\top})^{-1}$$

LSC gives same result for $\tilde{\mathbf{S}}^{-1}$ as pseudoinverse.

Q: Does this work for FE discretizations?... no

Review: BFBT/LSC methods for Schur complement $\tilde{\mathbf{S}}^{-1}$

Diagonally scaled BFBT method [Elman, et al., 2006]:

Find columns $\mathbf{x}_j$ of $\mathbf{X}$ s.t. $\min_{\mathbf{x}_j} \left\| \mathbf{M}_1^{-1/2} [\mathbf{A}\mathbf{M}_2^{-1}\mathbf{B}^\top]_j - \mathbf{M}_1^{-1/2}\mathbf{B}^\top \mathbf{x}_j \right\|_2^2$

$$\Rightarrow\ \mathbf{X} = (\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{M}_1^{-1}\mathbf{A}\mathbf{M}_2^{-1}\mathbf{B}^\top)$$

$$\Rightarrow\ \tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{M}_1^{-1}\mathbf{A}\mathbf{M}_2^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{M}_2^{-1}\mathbf{B}^\top)^{-1}$$

Proposed scaling: For FE, use "diagonalized" velocity mass matrix,

diagonal: $\mathbf{M}_1 = \mathbf{M}_2 = \mathrm{diag}(\mathbf{M}_{\boldsymbol{u}})$ or lumped: $\mathbf{M}_1 = \mathbf{M}_2 = \tilde{\mathbf{M}}_{\boldsymbol{u}}$

Since $\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top$ can be understood as a Laplace operator for the pressure, approximate $(\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top)^{-1}$ by a multigrid V-cycle.

Q: Is mass scaled BFBT effective for high viscosity variations?...

Review: BFBT/LSC methods for Schur complement $\tilde{\mathbf{S}}^{-1}$

Diagonally scaled BFBT method [Elman, et al., 2006]:

$$\text{Find columns } \mathbf{x}_j \text{ of } \mathbf{X} \text{ s.t. } \min_{\mathbf{x}_j} \left\| \mathbf{M}_1^{-1/2} [\mathbf{A}\mathbf{M}_2^{-1}\mathbf{B}^\top]_j - \mathbf{M}_1^{-1/2}\mathbf{B}^\top \mathbf{x}_j \right\|_2^2$$

$$\Rightarrow \; \mathbf{X} = (\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{M}_1^{-1}\mathbf{A}\mathbf{M}_2^{-1}\mathbf{B}^\top)$$

$$\Rightarrow \; \tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{M}_1^{-1}\mathbf{A}\mathbf{M}_2^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{M}_2^{-1}\mathbf{B}^\top)^{-1}$$

Proposed scaling: For FE, use "diagonalized" velocity mass matrix,

$$\text{diagonal: } \mathbf{M}_1 = \mathbf{M}_2 = \mathrm{diag}(\mathbf{M}_{\boldsymbol{u}}) \quad \text{or lumped: } \mathbf{M}_1 = \mathbf{M}_2 = \tilde{\mathbf{M}}_{\boldsymbol{u}}$$

Since $\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top$ can be understood as a Laplace operator for the pressure, approximate $(\mathbf{B}\mathbf{M}_1^{-1}\mathbf{B}^\top)^{-1}$ by a multigrid V-cycle.

Q: Is mass scaled BFBT effective for high viscosity variations?. . . no

Review: BFBT/LSC methods for Schur complement $\tilde{\mathbf{S}}^{-1}$

BFBT for scaled Stokes systems that arise in geodynamics [May, Moresi, 2008]:

$$\begin{bmatrix} \mathbf{D}_{\boldsymbol{u}}^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_p^{-1/2} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{\boldsymbol{u}}^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_p^{-1/2} \end{bmatrix}$$

Then the standard BFBT method yields its scaled version,

$$\Rightarrow \ \tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{A}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{B}^\top)^{-1}$$

Proposed scaling: heuristic, motivated by scaling of dimensional systems

$$[\mathbf{D}_{\boldsymbol{u}}]_{i,i} = \max_j |[\mathbf{A}]_{i,j}|$$

Q: Is BFBT with this scaling effective for high viscosity variations?...

Review: BFBT/LSC methods for Schur complement $\tilde{\mathbf{S}}^{-1}$

BFBT for scaled Stokes systems that arise in geodynamics [May, Moresi, 2008]:

$$\begin{bmatrix} \mathbf{D}_{\boldsymbol{u}}^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_p^{-1/2} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{\boldsymbol{u}}^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_p^{-1/2} \end{bmatrix}$$

Then the standard BFBT method yields its scaled version,

$$\Rightarrow \; \tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{A}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{D}_{\boldsymbol{u}}^{-1}\mathbf{B}^\top)^{-1}$$

Proposed scaling: heuristic, motivated by scaling of dimensional systems

$$[\mathbf{D}_{\boldsymbol{u}}]_{i,i} = \max_j |[\mathbf{A}]_{i,j}|$$

Q: Is BFBT with this scaling effective for high viscosity variations?... yes

New view on BFBT/LSC methods (1)

Let $\mathbf{C}$ be symm. pos. def. and let $\mathbf{D}$ be arbitrary,

$$\text{Find } \mathbf{X} \text{ s.t. } \quad \min_{\mathbf{X}} \left\| \mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top \mathbf{e}_j - \mathbf{B}^\top \mathbf{X} \mathbf{e}_j \right\|^2_{\mathbf{C}^{-1}} \text{ for all } j$$

$$\Rightarrow \; \mathbf{X} = (\mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{C}^{-1}\mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top)$$

And we have a $\mathbf{C}^{-1}$-orthogonal projection, i.e., the residual satisfies

$$\left\langle \mathbf{B}^\top \mathbf{e}_i, (\mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top - \mathbf{B}^\top \mathbf{X})\mathbf{e}_j \right\rangle_{\mathbf{C}^{-1}} = \mathbf{0} \quad \text{for all } i, j,$$

therefore

$$\left(\mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top - \mathbf{B}^\top \mathbf{X} \right) \mathbf{e}_j \perp_{\mathbf{C}^{-1}} \operatorname{Ran}(\mathbf{B}^\top) \quad \text{for all } j$$

New view on BFBT/LSC methods (2)

Goal: Effective and robust preconditioning of the Schur complement in Stokes systems with high viscosity variations

Note: Condition for optimal preconditioning $(\mathbf{B}\tilde{\mathbf{A}}^{-1}\mathbf{B}^\top)\tilde{\mathbf{S}}^{-1} = \mathbf{I}$.
By choosing $\mathbf{C} = \tilde{\mathbf{A}}$, we obtain equivalence between orthogonality and the condition for optimal preconditioning:

$$\left\langle \mathbf{B}^\top \mathbf{e}_i, (\mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top - \mathbf{B}^\top \mathbf{X})\mathbf{e}_j \right\rangle_{\tilde{\mathbf{A}}^{-1}} = \mathbf{0} \quad \forall i, j \quad \Leftrightarrow \quad \tilde{\mathbf{S}} = \mathbf{B}\tilde{\mathbf{A}}^{-1}\mathbf{B}^\top$$

New view on BFBT/LSC methods (2)

Goal: Effective and robust preconditioning of the Schur complement in Stokes systems with high viscosity variations

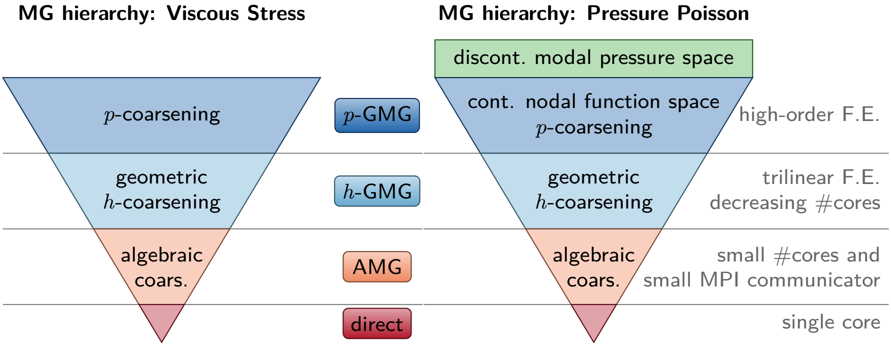
Note: Condition for optimal preconditioning $(\mathbf{B}\tilde{\mathbf{A}}^{-1}\mathbf{B}^\top)\tilde{\mathbf{S}}^{-1} = \mathbf{I}$.
By choosing $\mathbf{C} = \tilde{\mathbf{A}}$, we obtain equivalence between orthogonality and the condition for optimal preconditioning:

$$\left\langle \mathbf{B}^\top \mathbf{e}_i, (\mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top - \mathbf{B}^\top\mathbf{X})\mathbf{e}_j \right\rangle_{\tilde{\mathbf{A}}^{-1}} = \mathbf{0} \quad \forall i, j \quad \Leftrightarrow \quad \tilde{\mathbf{S}} = \mathbf{B}\tilde{\mathbf{A}}^{-1}\mathbf{B}^\top$$

Choices of $\mathbf{C}, \mathbf{D}$ that are computationally feasible are limited.
Our choice: $\mathbf{C} = \mathbf{D} := \mathrm{diag}(\mathbf{A})$, thus

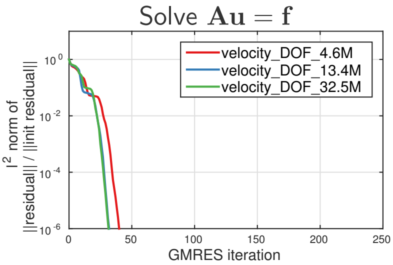
$$\tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^\top)^{-1}$$

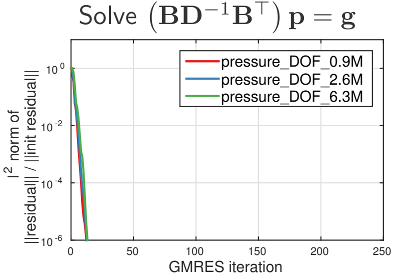
Parallel geometric multigrid (GMG)

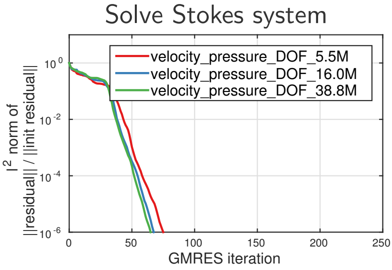


- Accurate high-order L^2-projection operators for restriction and interpolation during V-cycles, and for coarsening of the viscosity
- Coarsening of full fourth-order tensor coefficient of Jacobian
- Chebyshev accelerated point-Jacobi smoothers
- Velocity null spaces are projected out to establish stable convergence

h-dependence of GMG components & Stokes precond.

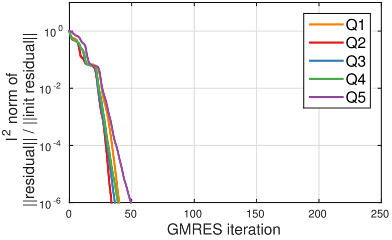




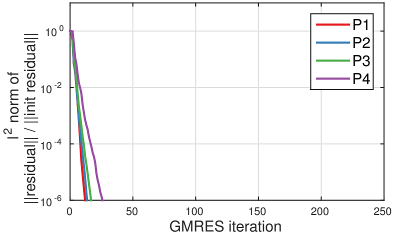


p-dependence of GMG components & Stokes precond.

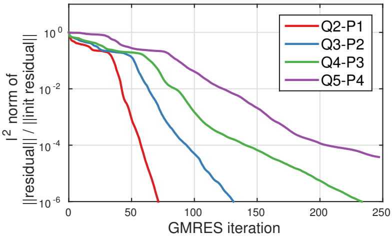
Solve $\mathbf{Au} = \mathbf{f}$



Solve $\left(\mathbf{BD}^{-1}\mathbf{B}^\top\right)\mathbf{p} = \mathbf{g}$



Solve Stokes system

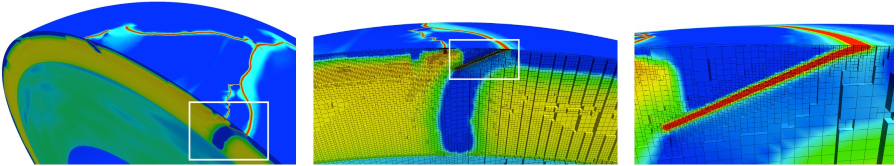


Computational results using real Earth data

Computational results: Solver Robustness

Robustness of linear Stokes solver w.r.t. plate boundary thickness

Plate boundary thickness (km)	**DOF**	**GMRES #iter** for solving $\mathbf{Au} = \mathbf{f}$	**GMRES #iter** for solving Stokes
15	1.16B	115	461
10	1.41B	129	488
5	3.01B	123	445

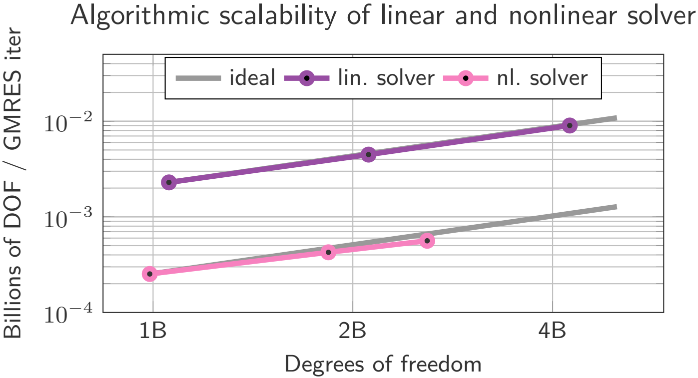


Robustness of inexact Newton-Krylov nonlinear solver w.r.t plate boundary thickness

Plate boundary thickness (km)	**DOF** (at nl. solution)	**Newton #steps**	**GMRES #iter** (for all Newton steps)
15	1.00B	25	3915
10	1.63B	27	4645
5	4.78B	29	6112

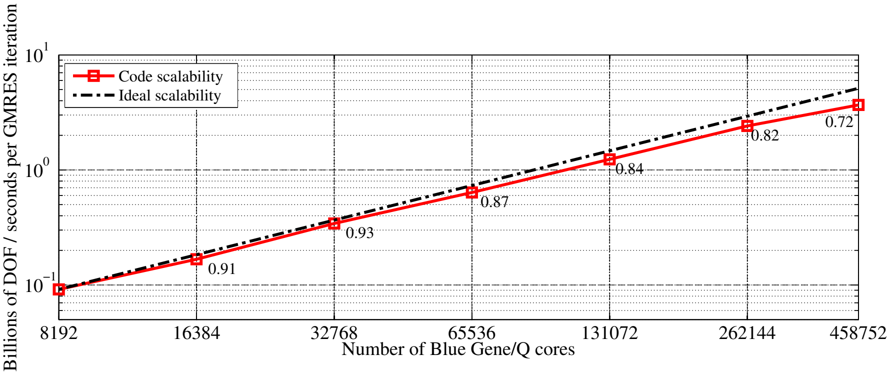
Computational results: Algorithmic scalability

(Fix problem parameters and refine the mesh)



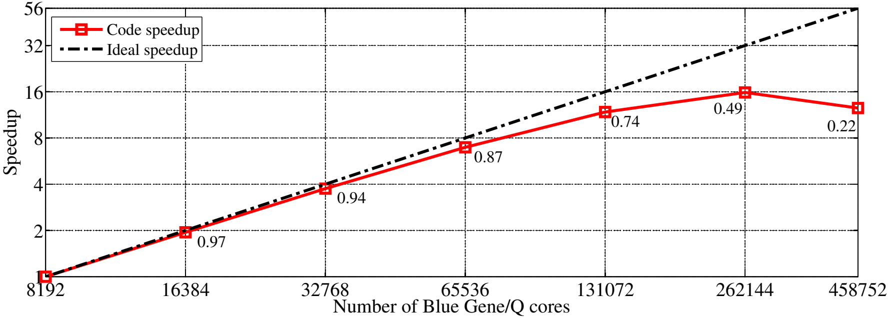
Computational results: Weak scalability

(Increase core count and problem size simultaneously)



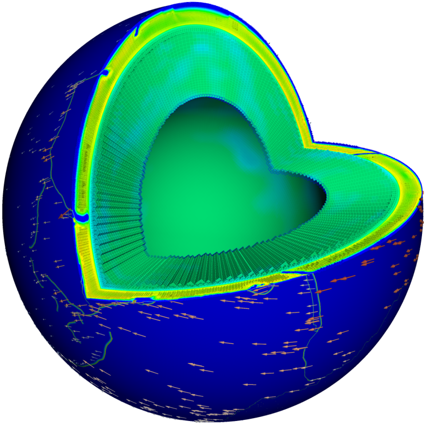
Computational results: Strong scalability

(Increase core count while problem size stays fixed)

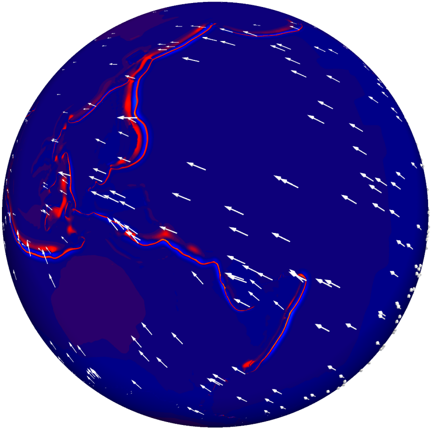


Computational results: Contribution to science

Effective viscosity at nonlinear solution and surface velocity



Normal stress at the surface and surface velocity



Summary of key contributions

- Parallel geometric multigrid for the viscous stress block on adaptive meshes (my impl.; based on AMG-only solver and parallel AMR library)
- p- and h-independent multigrid convergence through improvement of projection operators (my dev.)
- Stable convergence in presence of rotation null spaces (my dev.)
- Stable coarsening of anisotropic fourth-order tensor coefficient in Jacobian (my dev.)
- Geometric multigrid based BFBT; first matrix-free implementation of BFBT (my dev.)
- Inexact Newton-Krylov for complex Earth rheology with dynamic mesh refinement (my impl.; ideas in collaboration with G. Stadler)
- Parallel optimizations, e.g., MPI communicator reduction, OpenMP threading (my impl.)
- Parallel scalability on full system JUQUEEN supercomputer with 458,752 CPU cores (lead dev. in collaboration with IBM Research – Zürich)