Global Convection in Earth's Mantle: Advanced Numerical Methods for Forward and Inverse Modeling

Johann Rudi[1] Georg Stadler[2] Jiashun Hu[3] Michael Gurnis[3] Omar Ghattas[4]

[1]Mathematics and Computer Science Division, Argonne National Laboratory

[2]Courant Institute of Mathematical Sciences, New York University

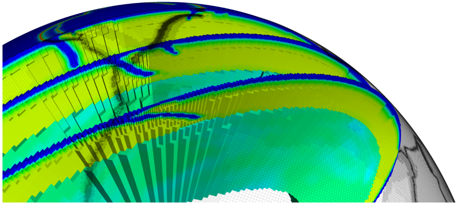
[3]Seismological Laboratory, California Institute of Technology

[4]Oden Institute, Jackson School of Geosciences, and Dept. of Mechanical Engineering, The University of Texas at Austin

Motivation: Understanding of Earth's fundamental properties

Simulating realistic high-resolution models:

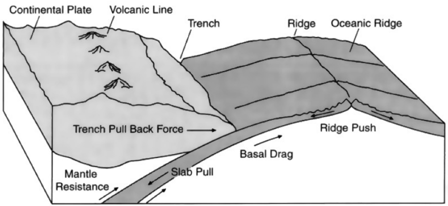
- Global mantle convection & associated plate tectonics
- Resolutions to faulted plate boundaries
- Nonlinear viscous fluid flow



Computational mesh that resolves viscosity (colors).

Learning properties of Earth's rheology:

- What we know: Observational data of plate motions at the surface, topography, surface strain rates, estimates of average viscosities, etc.
- What we want: Parameters in constitutive laws and plate coupling strength that are consistent with data and model; Uncertainties w.r.t. available data

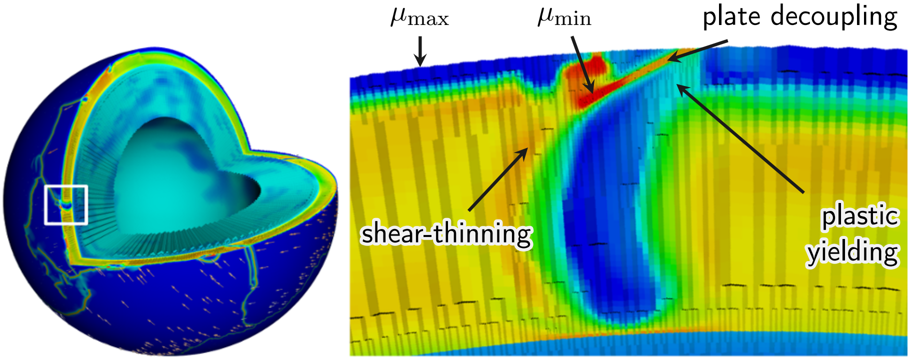


Underlying forces (Credit: Schubert, Turcotte, Olsen).

Simulation

Approach for simulation

Model PDE: present day instantaneous flow modeled by incompressible, nonlinear Stokes
Rheology / effective viscosity: shear-thinning with plastic yielding, and upper/lower bounds



Colors represent viscosity ($\mu_{\mathrm{max}}/\mu_{\mathrm{min}} = 10^6$).

Approach for simulation

Model PDE: present day instantaneous flow modeled by incompressible, nonlinear Stokes

$$-\nabla \cdot \left[\mu(\boldsymbol{x}, \dot{\varepsilon}_{\mathrm{II}}(\boldsymbol{u}))\left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^{\mathsf{T}}\right)\right] + \nabla p = \boldsymbol{f} \qquad \text{viscosity } \mu,\ \text{RHS forcing } \boldsymbol{f}$$

$$-\nabla \cdot \boldsymbol{u} = 0 \qquad \text{velocity } \boldsymbol{u},\ \text{pressure } p$$

- Free-slip boundary conditions (tangential velocities at the surface are output of the model)
- Temperature structure from seismic models and convergence rate of slabs
- Fault weak zones based on *Slab1.0* compilation of models for subduction zones

Rheology / effective viscosity: shear-thinning with plastic yielding, and upper/lower bounds

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

- Plates are modeled by high-viscosity: $a(T(\boldsymbol{x}))$ depends exponentially on T
- Plate boundaries are narrow zones of weak viscosity: $10^{-6} \leq w(\boldsymbol{x}) \leq 1$

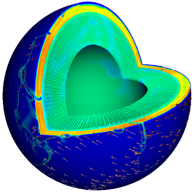
Severe challenges for parallel scalable implicit solvers

... arising from global models of Earth's mantle convection:

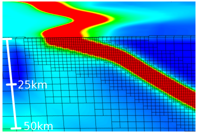
- Severe nonlinearity, heterogeneity, and anisotropy (upon linearization)
- Sharp viscosity gradients in narrow regions (change by $\mathcal{O}(10^6)$ in $\sim$5 km)
- Wide range of spatial scales and highly localized features, e.g., plate boundaries of size $\mathcal{O}(1\,\text{km})$ influence plate motion at continental scales of $\mathcal{O}(1000\,\text{km})$
- Adaptive refinement of computational mesh is essential
- High-order finite elements $\mathbb{Q}_k \times \mathbb{P}_{k-1}^{\text{disc}}$, order $k \geq 2$, with local mass conservation; yields a difficult to deal with discontinuous, modal pressure approximation

---

A resolved model has about 1 billion DOFs (with aggressive adaptivity), because smallest mesh elements $\sim$1 km; Runs on 10K–100K computing cores.



Colors represent viscosity, arrows show plate velocities.



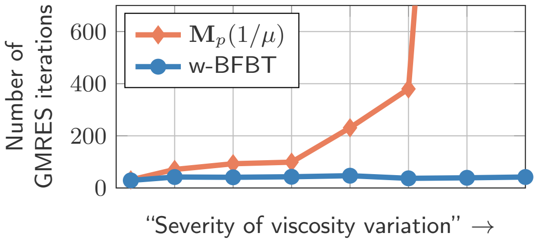
Fault zone (red) between stationary and subducting plate.

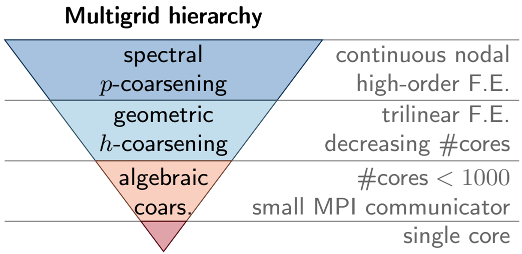
Solver highlights: Viscosity variation-robust preconditioning and scalable parallel multigrid for adaptive high-order F.E.

Stokes preconditioning: $\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\hat{\mathbf{u}}} \\ \tilde{\hat{\mathbf{p}}} \end{bmatrix} = \begin{bmatrix} -\mathbf{r}_1 \\ -\mathbf{r}_2 \end{bmatrix}$ $\quad \tilde{\mathbf{A}}^{-1} \approx \mathbf{A}^{-1} \rightarrow$ MG V-cycle, $\quad \tilde{\mathbf{S}}^{-1} \approx (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top)^{-1}$

Schur complement preconditioning: $\tilde{\mathbf{S}}^{-1} := \underbrace{(\mathbf{B}\mathbf{C}_w^{-1}\mathbf{B}^\top)^{-1}}_{\rightarrow \text{ MG V-cycle}} (\mathbf{B}\mathbf{C}_w^{-1}\mathbf{A}\mathbf{D}_w^{-1}\mathbf{B}^\top) \underbrace{(\mathbf{B}\mathbf{D}_w^{-1}\mathbf{B}^\top)^{-1}}_{\rightarrow \text{ MG V-cycle}}$

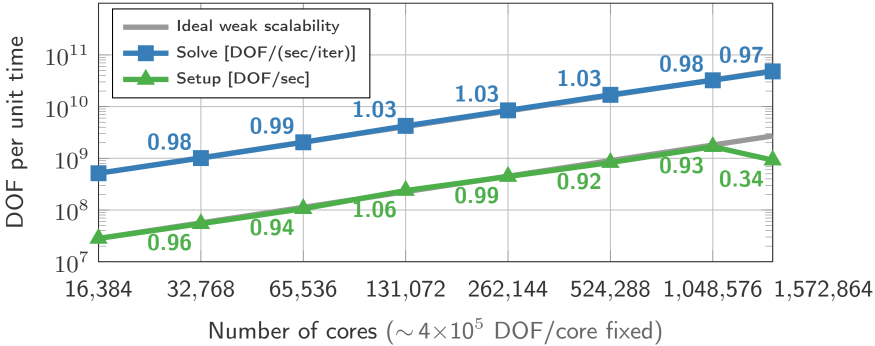
Choice of diagonal weighting matrices $\mathbf{C}_w = \mathbf{D}_w$ is critical for efficacy & robustness.





Solver highlights: Parallel weak scalability on Sequoia (BlueGene/Q)

Parallel scalability of linear iterative solvers to 1.6M cores (collaboration with IBM Research, Zurich):



- Excellent parallel scalability while maintaining optimal algorithmic performance w.r.t. mesh refinement; nearly optimal w.r.t. higher discretization order
- Full nonlinear Stokes solver: Inexact Newton–Krylov with nonlinear preconditioning and grid continuation for highly nonlinear models of Earth's mantle

Inference

Approach for inference: From data to parameters consistent with model

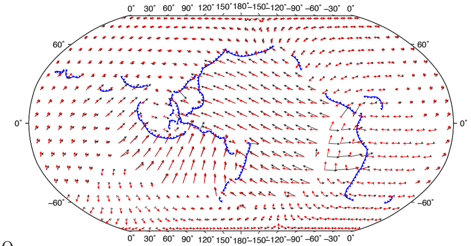
General procedure for inference from data consistent with a model (physical, statistical, etc.):

1. **Question**: Start with a scientific question, e.g., quantify forces in deep Earth.
2. **Data collection**: Gather data relevant to the question, e.g., plate motions at the surface.
3. **Model selection**: Construct a model that is *appropriate* to the problem and data, moreover, select parameters for inference, e.g., global mantle convection model with plate coupling strength as parameters.
4. **Parameter estimation**: Find the parameters in the model, which often requires minimizing the difference between model output and data (in some suitable metric).
5. **Validation**: Inspect if the model and parameters are consistent with knowledge or data that was excluded from estimation, e.g., average viscosity or topography.
6. **Prediction**: Estimate the quantities of interest and their uncertainties that are associated with the initial scientific question.

Approach for inference: Global mantle convection & plate tectonics

Data: Challenging because of limited amount

- Current plate motion of rigid plates from GPS and magnetic anomalies recorded in bands perpendicular to seafloor spreading



Data (black) vs. model (red) (Credit: J. Hu).

Model: Challenging because of computational complexity

- Incompressible, nonlinear Stokes PDE:

$$-\nabla\cdot\left[\mu(\boldsymbol{x},\dot{\varepsilon}_{\mathrm{II}}(\boldsymbol{u}))\left(\nabla\boldsymbol{u}+\nabla\boldsymbol{u}^{\mathsf{T}}\right)\right]+\nabla p=\boldsymbol{f},\quad -\nabla\cdot\boldsymbol{u}=0$$

- Effective viscosity: $\mu(\boldsymbol{x},\dot{\varepsilon}_{\mathrm{II}}(\boldsymbol{u})) := \mu_{\min}+\min\left(\frac{\tau_{\mathrm{yield}}}{2\dot{\varepsilon}_{\mathrm{II}}(\boldsymbol{u})}\,,\, w(\boldsymbol{x})\min\left(\mu_{\max}\,,\, a(T(\boldsymbol{x}))^{\frac{1}{n}}\,\dot{\varepsilon}_{\mathrm{II}}(\boldsymbol{u})^{\frac{1}{n}-1}\right)\right)$

Parameters: Challenging because of vastly different scales of sensitivity

- Global parameters: scaling factors, activation energy in Arrhenius law $a(T(\boldsymbol{x}))$, stress exponent n, yield strength τ_{yield}
- Local coupling strength $w(\boldsymbol{x})$, i.e., energy dissipation between adjacent plates

Systematic inversion of uncertain parameters consistent with data & model

Given:

- Model PDE (forward problem), here incompressible, nonlinear Stokes: $\mathcal{A}(m,u) = f$
- Map of model output (dependent on parameters m) to observations: $\mathcal{F}(u(m))$
- Assume data d contains normally distributed additive noise, $\mathcal{N}(0, \mathscr{C}_{\text{noise}})$
- Assume prior of the parameters m is normally distributed, $\mathcal{N}(m_{\text{pr}}, \mathscr{C}_{\text{pr}})$

Want: Description of the posterior density of the parameters (using Bayes' theorem)

$$\pi_{\text{post}}(m) \propto \exp\left(-\frac{1}{2}\left\|d - \mathcal{F}(u(m))\right\|^2_{\mathscr{C}^{-1}_{\text{noise}}} - \frac{1}{2}\left\|m - m_{\text{pr}}\right\|^2_{\mathscr{C}^{-1}_{\text{pr}}}\right)$$

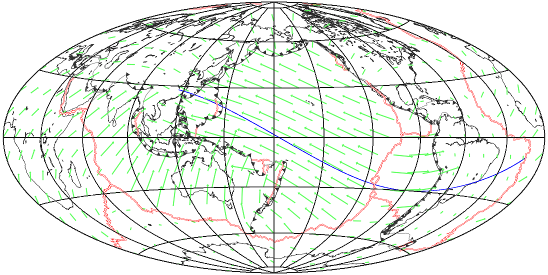
Computationally feasible: Using adjoints for gradient computation & BFGS approximation of Hessian

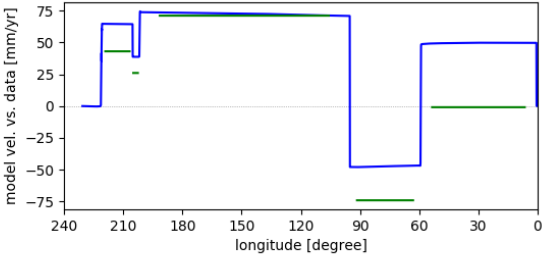
- Find the maximum of $\pi_{\text{post}}(m)$ by solving an optimization problem constraint by the model PDE:

$$\arg\min_m \mathcal{J}(m, u(m)) := \frac{1}{2}\left\|d - \mathcal{F}(u(m))\right\|^2_{\mathscr{C}^{-1}_{\text{noise}}} + \frac{1}{2}\left\|m - m_{\text{pr}}\right\|^2_{\mathscr{C}^{-1}_{\text{pr}}} \quad \text{subject to} \quad \mathcal{A}(m, u(m)) = f$$

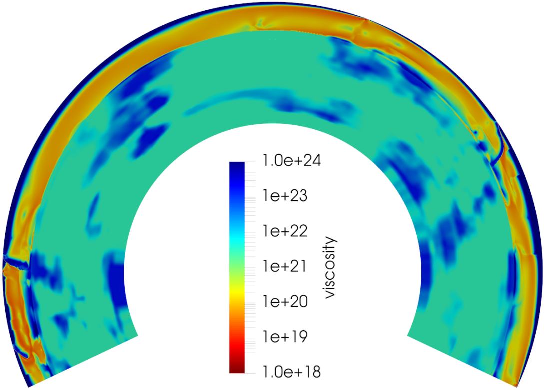
- Construct a Gaussian approximation of $\pi_{\text{post}}(m)$ around this maximum by approximating the Hessian of the optimization problem

Computational results: Inverse problem for a cross sectional model



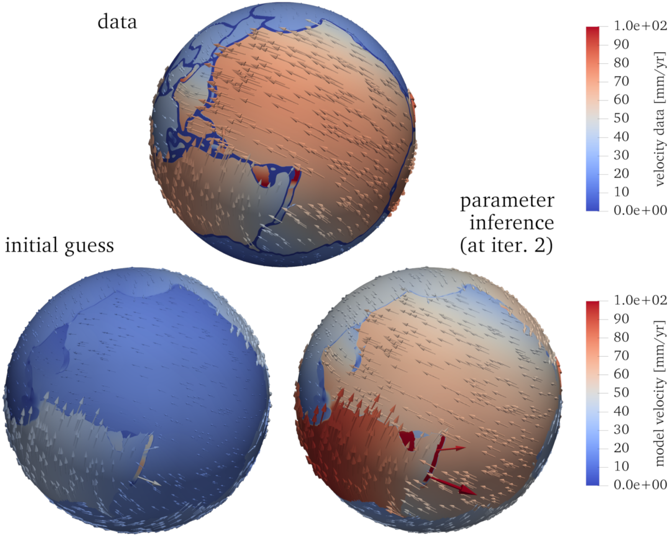


Effective viscosity after inference



- Cross sectional model is not able to adequately capture observed physics, i.e., fast motion of stationary South American plate

Computational results: Inverse problem for global Earth model



- Inversion of global scalars (scaling factors, activation energy, stress exponent, yield strength) successfully recovers main trends, i.e., velocities of large plates
- Local parameters (plate decoupling) remain challenging due to lower sensitivities

Thank you