

# Uncertainty Quantification for Large-Scale Ice Sheet Modeling and Simulation

Omar Ghattas<sup>\*†‡</sup>

Representing the rest of the UT-Austin/LANL team:  
Don Blankenship<sup>†</sup> Tan Bui-Thanh<sup>\*</sup> Carsten Burstedde<sup>\*</sup>  
David Higdon<sup>¶</sup> Tobin Isaac<sup>\*</sup> Charles Jackson<sup>†</sup>  
Stephen Price<sup>§</sup> Georg Stadler<sup>\*</sup> Lucas Wilcox<sup>\*</sup>  
(Bill Lipscomb)

<sup>\*</sup>Institute for Computational Engineering and Sciences, UT-Austin

<sup>†</sup>Jackson School of Geosciences, UT-Austin

<sup>‡</sup>Dept. of Mechanical Engineering, UT-Austin

<sup>§</sup>COSIM/Fluid Dynamics Group, Los Alamos National Laboratory

<sup>¶</sup>Statistical Sciences, Los Alamos National Laboratory

## Needs and challenges

- ▶ Scalable linear solvers for variable-coefficient Stokes
- ▶ Scalable nonlinear solvers for nonlinear rheology
- ▶ Adaptivity to resolve flow transitions and high velocity gradients
- ▶ Inverse methods to identify unknown constitutive parameters, basal boundary conditions, and geothermal heat flux
- ▶ Statistical inference methods to estimate uncertainty in unknowns
- ▶ Assimilation of observational data into models to quantify uncertainties in ice sheet dynamics predictions

# Outline

Review of current capability for modeling creeping non-Newtonian Stokesian flow and transport (illustrated with mantle convection)

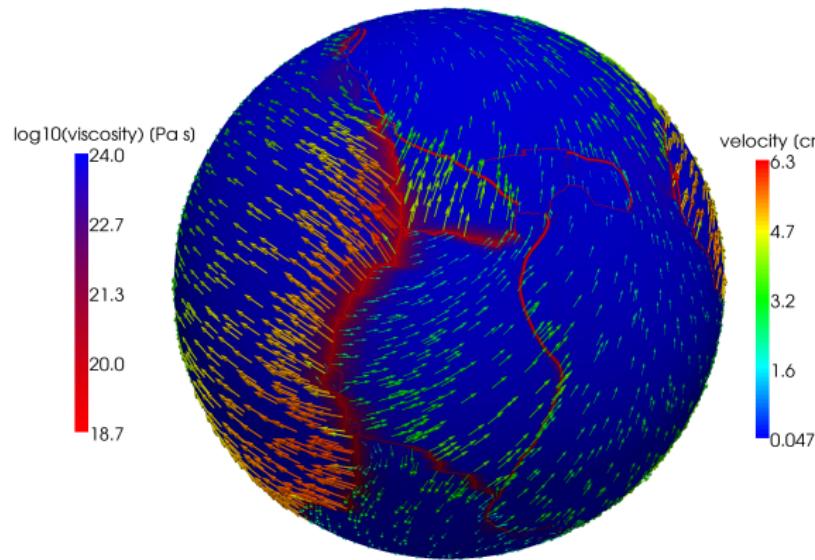
Enhancements of current implementation

New capabilities

Summary

# High resolution mantle convection simulation

- ▶ **Goal:** first global mantle convection simulations that resolve faulted plate boundaries under realistic conditions
- ▶ **Multiscale problem:** Spatial scales range from 1 km to  $10^4$  km; time scales from  $10^4$  years to  $10^9$  years  $\Rightarrow 10^{17}$  DOFs with regular grid



**Figure:** Plate motion with nonlinear rheology, showing velocity (vectors) and viscosity (colors)

# High resolution mantle convection simulation

- ▶ **Goal:** first global mantle convection simulations that resolve faulted plate boundaries under realistic conditions
  - ▶ **Multiscale problem:** Spatial scales range from 1 km to  $10^4$  km; time scales from  $10^4$  years to  $10^9$  years  $\Rightarrow 10^{17}$  DOFs with regular grid
  - ▶ **Adaptivity** is essential to render problem tractable
  - ▶ **Challenge:** parallel dynamic AMR and solvers for PDE systems on petascale systems
- 
- ▶ maximize accuracy per wall clock time

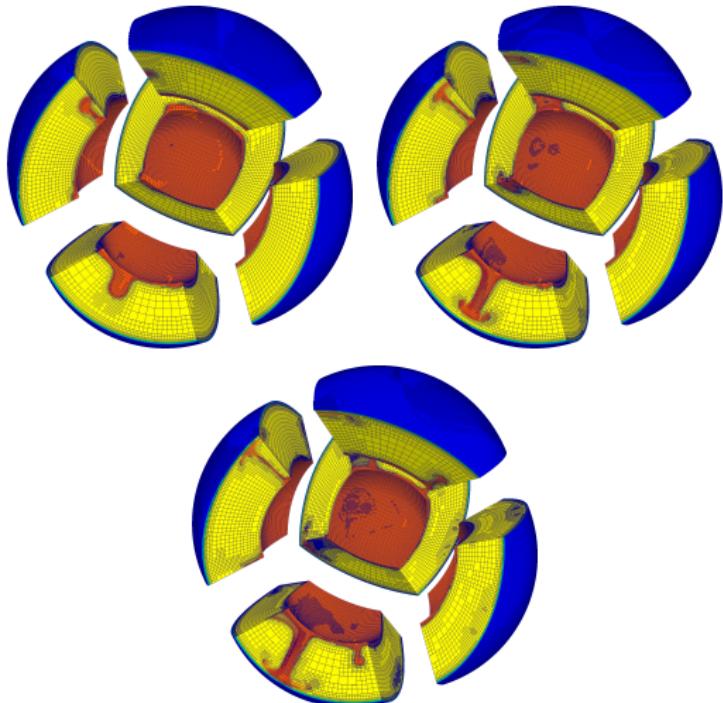
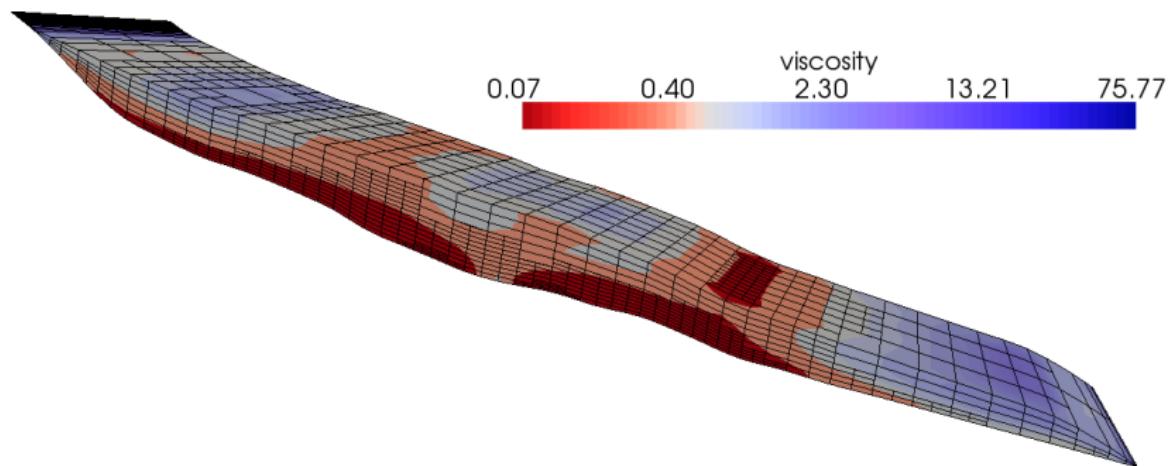


Figure: Parallel dynamic AMR

## Example: Adaptivity for a benchmark ice flow problem



Haut Glacier D'Arolla benchmark problem from ISMIP-HOM

# Mantle convection equations

$$\begin{aligned}\nabla \cdot \mathbf{u} &= 0 \\ -\nabla \cdot [\eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^\top) - p \mathbf{I}] &= \rho_0 [1 - \alpha (T - T_0)] \mathbf{g} \\ \rho_0 c \left( \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right) - \nabla \cdot (k \nabla T) &= H(\mathbf{u}) \\ \mathbf{n} \times \mathbf{n} \times [\eta (\nabla \mathbf{u} + \nabla \mathbf{u}^\top) - p \mathbf{I}] \mathbf{n} &= \mathbf{0} \text{ on } \partial\Omega \\ \mathbf{u} \cdot \mathbf{n} = 0 \text{ and } T = T_{\text{bc}} \text{ on } \partial\Omega \text{ and } T = T_{\text{inc}} \text{ at } t = 0\end{aligned}$$

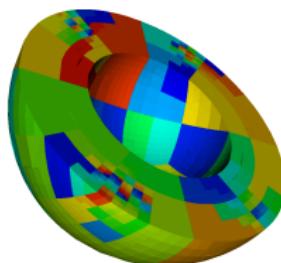
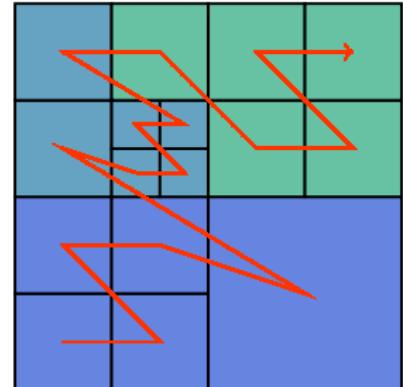
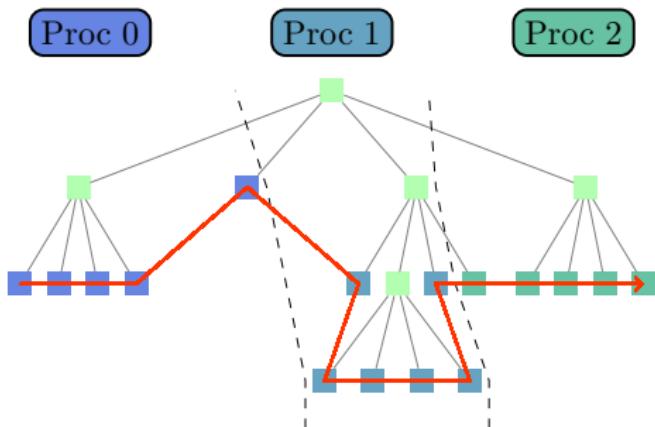
Variables:

- $T(\mathbf{x}, t)$ ,  $\mathbf{u}(\mathbf{x}, t)$ ,  $p(\mathbf{x}, t)$  — temperature, velocity, pressure

Parameters:

- Ra — Rayleigh number  $\sim 10^6$ – $10^9$
- $H(\mathbf{u})$ ,  $\eta(T, \mathbf{u})$  — heat production rate, viscosity
- $\rho_0$ ,  $T_0$  — reference density and temperature
- $k$ ,  $c$  — thermal conductivity, specific heat
- $\mathbf{g}$ ,  $\alpha$  — gravitational acceleration, coefficient of thermal expansion

# Adaptive FEM based on parallel octrees



# Strong scalability of AMR for advection–diffusion

Speedup vs. number of cores demonstrates excellent fixed-size scalability

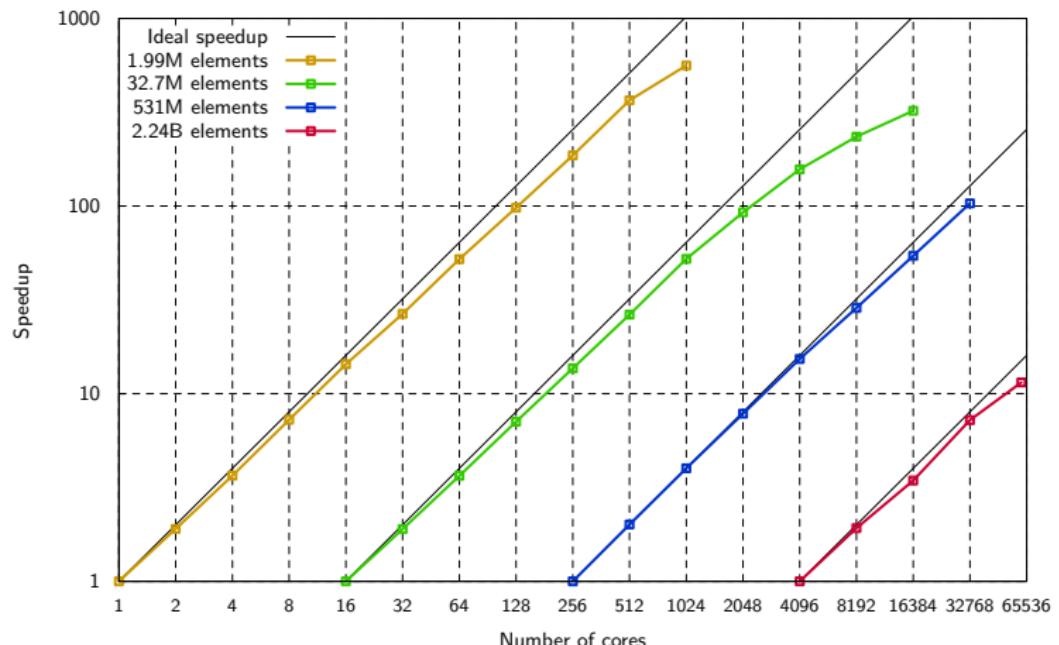


Figure: mesh is adapted every 32 time steps; explicit time stepping; trilinear elements

# Weak scalability of AMR for advection–diffusion

Excellent overall efficiency for weak scaling from 1 to 62K cores

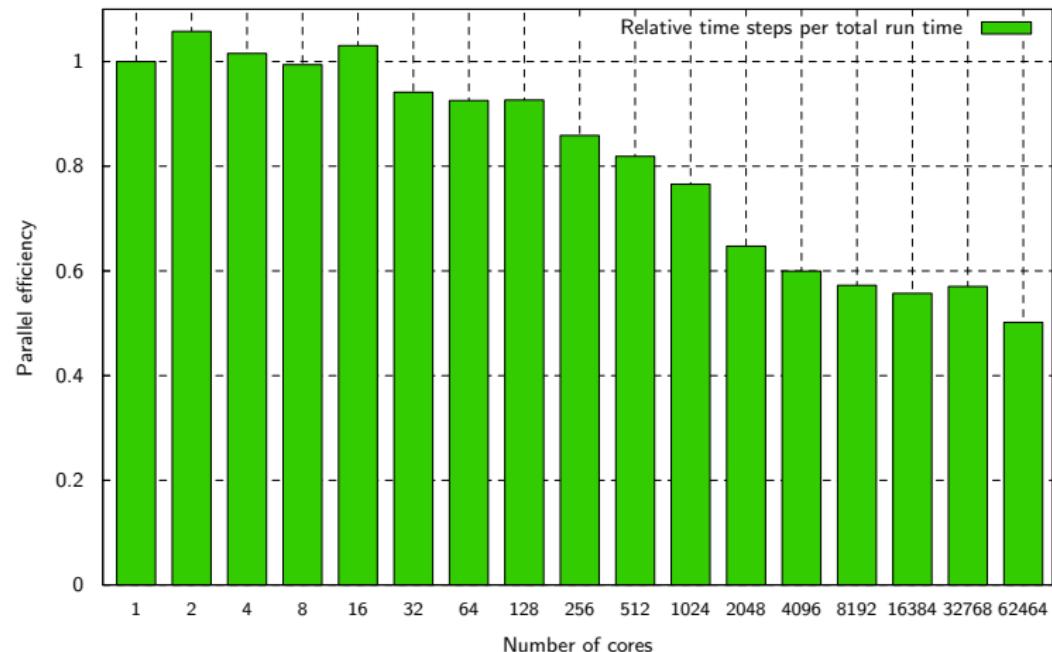


Figure: mesh is adapted every 32 time steps; explicit time stepping; trilinear elements; 131K elements/core (7.9B elements on 62K cores)

# Weak scalability of AMR for advection–diffusion

Breakdown of components: AMR consumes < 11% of overall time

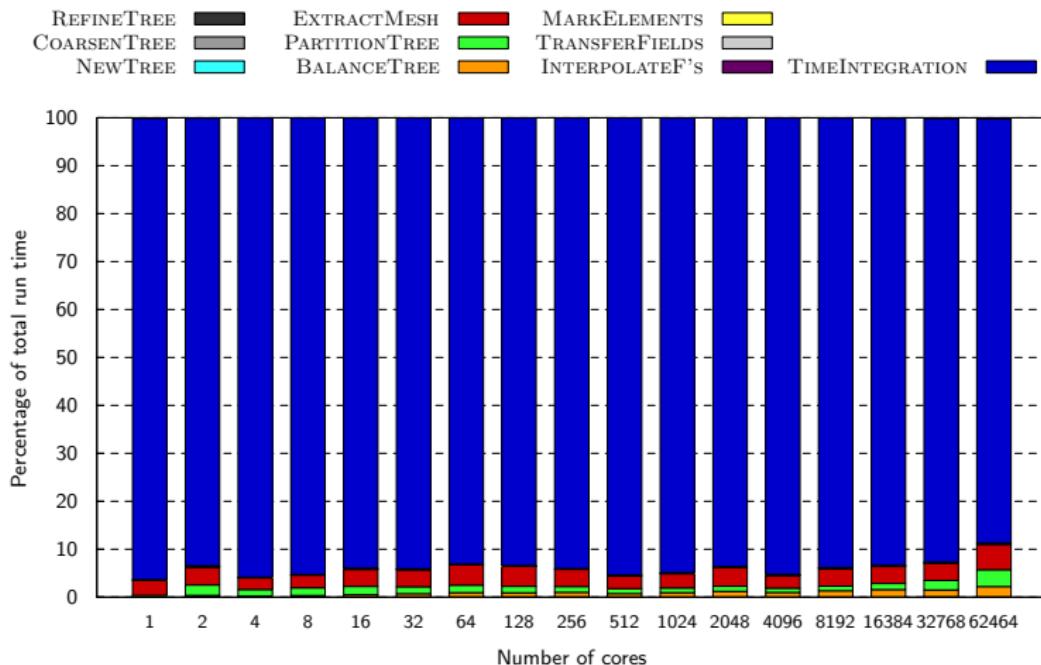


Figure: mesh is adapted every 32 time steps; explicit time stepping; trilinear elements; 131K elements/core (7.9B elements on 62K cores)

# Rhea: adaptive mantle convection code

## Summary of discretization and solution

- ▶ Trilinear FEM for temperature, velocity, and pressure
  - ▶ Conforming approximation by algebraic elimination of “hanging nodes”
- ▶ FEM stabilization:
  - ▶ Streamline Upwind/Petrov Galerkin (SUPG) for advection-diffusion system
  - ▶ Polynomial pressure projection for stabilization of Stokes equation (Dohrmann and Bochev)
- ▶ Explicit integration of energy equation decouples temperature update from nonlinear Stokes solve
- ▶ Nonlinear Stokes solver: lagged-viscosity (Picard) iteration
- ▶ Linear Stokes solver: MINRES iteration with block preconditioner based on AMG V-cycle and inverse viscosity mass matrix approximation of Schur complement

## MINRES preconditioner for linear Stokes solve

- ▶ Block factorization of Stokes system:

$$\begin{pmatrix} A & B^\top \\ B & -C \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -(BA^{-1}B^\top + C) \end{pmatrix} \begin{pmatrix} I & A^{-1}B^\top \\ 0 & I \end{pmatrix}$$

where  $A$  is discrete viscous operator,  $B$  is discrete divergence, and  $C$  is pressure stabilization

- ▶ Suggests preconditioner of form:

$$P = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \quad \text{with} \quad S = BA^{-1}B^\top + C$$

- ▶ Approximate inverse of  $A$  with one V-cycle of Algebraic Multigrid (e.g. *ML* or *BoomerAMG*)
- ▶ Approximate  $S$  with diagonal inverse-viscosity lumped-mass matrix  $\tilde{M}$  (spectrally equivalent)

# Independence of solver w.r.t. viscosity variation

Using BoomerAMG on Cartesian geometry

$\mu_{\min}$	$\mu_{\max}$	# MINRES iterations	AMG setup time (s)	solve time per iteration (s)
1.00e-0	1.00	86	25.29	5.82
4.98e-2	1.00	80	28.02	5.80
5.53e-4	1.00	75	25.26	5.62
5.53e-4	1.00	90	28.44	5.75
5.53e-4	1.00	91	26.97	5.35
6.14e-6	1.00	95	28.42	5.70
3.06e-7	1.00	93	31.35	6.46

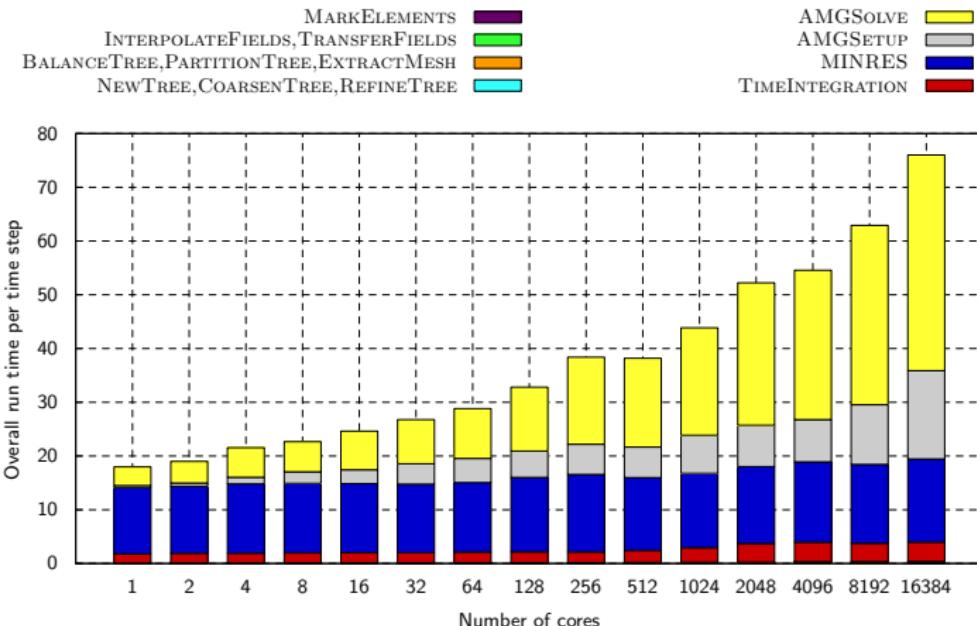
Figure: 216M unknowns, 512 cores

## Weak scalability of Stokes solver

#cores	#dofs	MinRes #iter	AMG setup	AMG V-cycle
1	170K	66	1.45	18.06
8	1.1M	76	1.60	22.91
32	4.6M	88	2.22	33.20
128	17.9M	81	3.41	30.22
2048	294M	63	15.12	70.53
16384	2.35B	71	26.91	84.96

Figure: Weak scaling,  $5 \times 10^3$  viscosity variation, 2 times adaptively refined mesh, Spherical geometry, full viscous block preconditioner, *ML* from *Trilinos* with *ZOLTAN* repartitioning used for V-cycle.

# Weak scaling of AMR mantle convection



**Figure:** Weak scaling, 50K elements/core, three levels of refinement, residual-based global error indicator, adaptation every 16 times steps, Cartesian geometry and *BoomerAMG* from *hypre* as preconditioner.

# Breakdown of components

Time per adaptation frame in seconds

#cores	Stokes solver	balance	partition	extract	interp/ transfer	mark/ adapt	AMR/solve ratio
1	269.00	0.03	0.00	0.48	0.05	0.04	0.23%
8	355.42	0.10	0.35	0.77	0.06	0.08	0.39%
64	434.75	0.20	0.46	0.97	0.07	0.14	0.43%
512	578.53	0.61	0.67	1.83	0.09	0.17	0.59%
4096	813.43	1.04	0.95	2.73	0.15	0.23	0.63%
16384	1134.30	1.23	1.22	2.85	0.20	0.32	0.51%

**Figure:** Isogranular scaling, 50K elements/core, three levels of refinement, residual-based global error indicator, adaptation every 16 times steps, Cartesian geometry and *BoomerAMG*

# Outline

Review of current capability for modeling creeping non-Newtonian Stokesian flow and transport (illustrated with mantle convection)

Enhancements of current implementation

New capabilities

Summary

# BoomerAMG preconditioner comparison

Variable viscosity adapted mesh Poisson vs. 7-point Laplacian

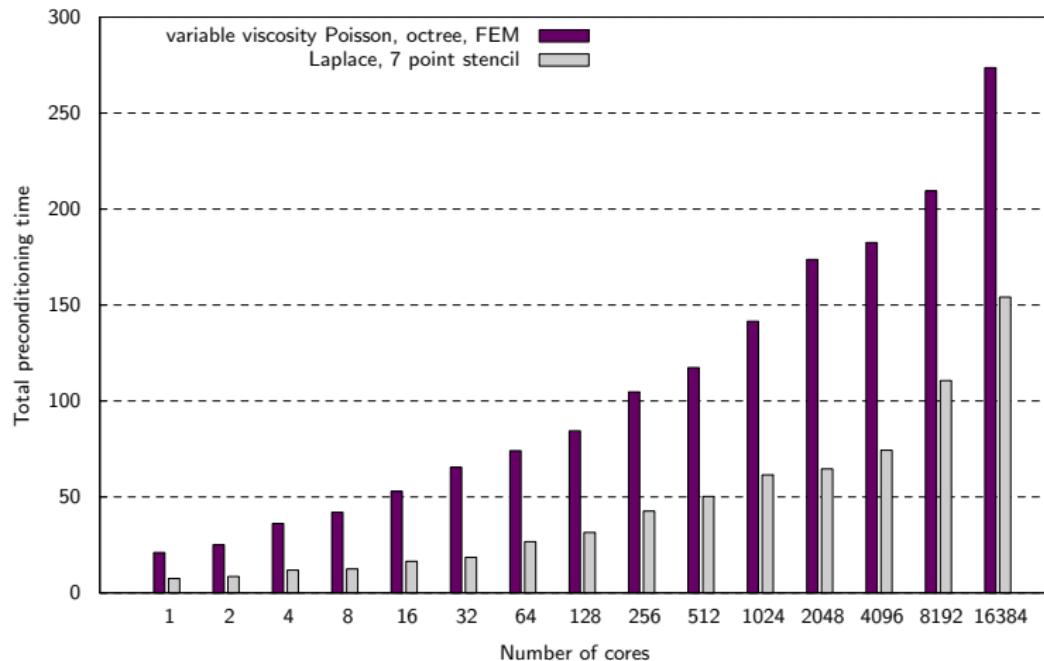
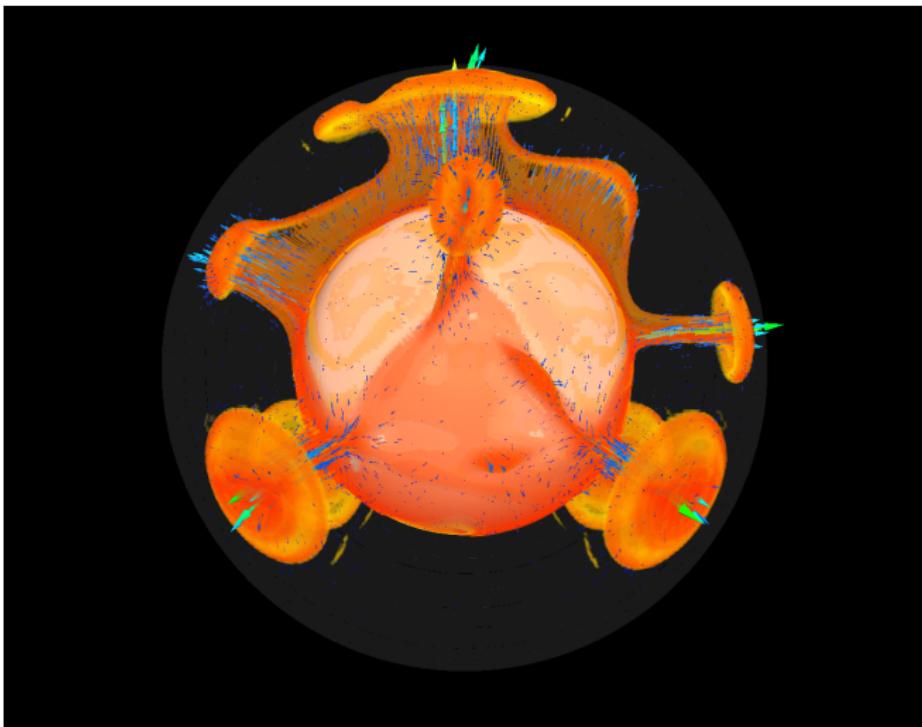
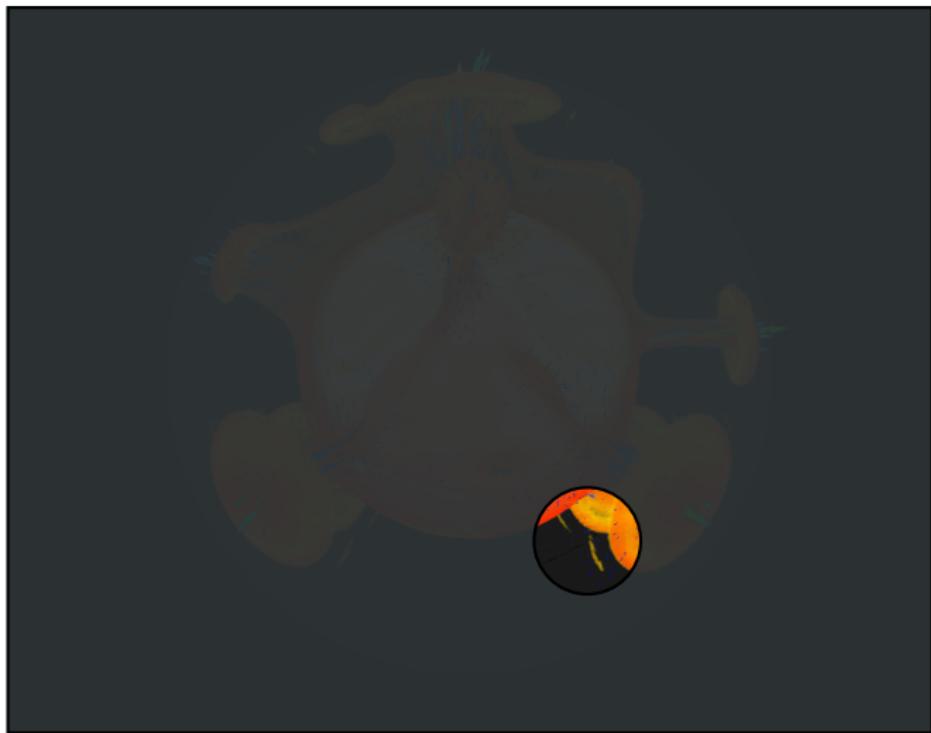


Figure: Timings for 1 AMG setup + 160 V-cycles; weak scaling with 50K elements/core, run on Ranger

# Dispersion and oscillations in temperature field



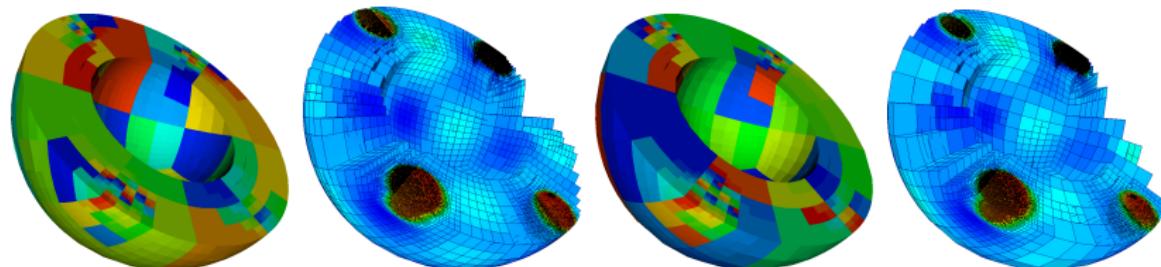
# Dispersion and oscillations in temperature field



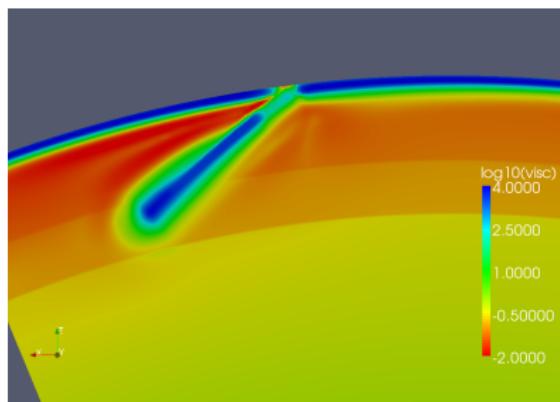
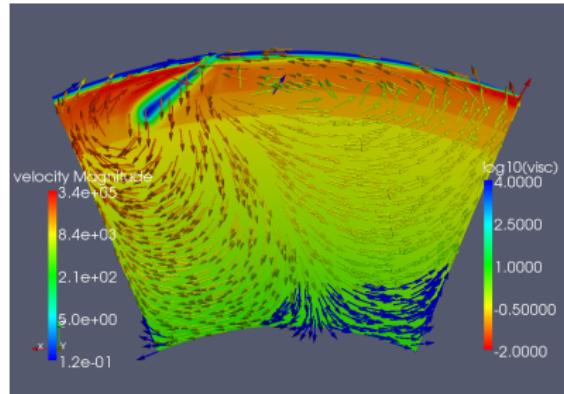
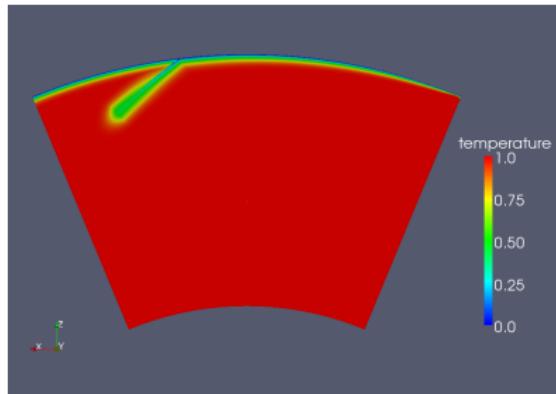
# Dispersion and oscillations in temperature field

Possible solutions:

- ▶ High-order discretization for dispersion
- ▶ Discontinuous Galerkin + flux limiter for energy equation
- ▶ Refine more near sharp fronts (will reduce time step)
- ▶ We currently developing a high-order discretization library (*mangll*) tailored to forest-of-octrees mesh (*p4est*)
  - ▶ Hexahedral elements with 2:1 adapted faces
  - ▶ Arbitrary order spectral element basis functions (nodal, GLL)
  - ▶ Continuous (in progress) and discontinuous Galerkin approximations
  - ▶ Isoparametric or piecewise diffeomorphic geometry mapping



# Limitations of explicit Picard for highly nonlinear rheology



Ex. of descending slab (L. Alisic)

- ▶ weak zone thickness and strength control yielding
- ▶ requires fine mesh, CFL becomes onerous
- ▶ Picard requires many iters
- ▶ Solution: Newton–Krylov + fully implicit

# Forward Euler, Newton step

$$\left[ \begin{array}{c|cc|c} 0 & & \nabla \cdot & 0 \\ \hline \nabla & -\nabla \cdot [\eta_m^{l+1} (\nabla + \nabla^\top)] & \rho_0 \alpha \mathbf{g} - \nabla \cdot [\nabla_T \eta_m^{l+1} (\nabla \mathbf{u}_m^{l+1} + [\nabla \mathbf{u}_m^{l+1}]^\top)] & \begin{bmatrix} \delta p_{m+1}^{l+1} \\ \delta \mathbf{u}_{m+1}^{l+1} \\ \delta T_{m+1}^{l+1} \end{bmatrix} \\ \hline 0 & 0 & \rho_0 c/h & \end{array} \right]$$

$$p_{m+1}^{l+1} = \delta p_{m+1}^{l+1} + p_m^{l+1}$$

$$\mathbf{u}_{m+1}^{l+1} = \delta \mathbf{u}_{m+1}^{l+1} + \mathbf{u}_m^{l+1}$$

$$T_{m+1}^{l+1} = \delta T_{m+1}^{l+1} + T_m^{l+1}$$

# Backward Euler, Newton step

$$\left[ \begin{array}{c|cc|c} 0 & & \nabla \cdot & 0 \\ \hline \nabla & -\nabla \cdot [\eta_m^{l+1} (\nabla + \nabla^\top)] & \rho_0 \alpha \mathbf{g} - \nabla \cdot [\nabla_T \eta_m^{l+1} (\nabla \mathbf{u}_m^{l+1} + [\nabla \mathbf{u}_m^{l+1}]^\top)] & \begin{bmatrix} \delta p_{m+1}^{l+1} \\ \delta \mathbf{u}_{m+1}^{l+1} \\ \delta T_{m+1}^{l+1} \end{bmatrix} \\ \hline 0 & \rho_0 c \nabla T_m^{l+1} - \nabla \cdot H_m^{l+1} & \rho_0 c (1/h + \mathbf{u}_m^{l+1} \cdot \nabla) - \nabla \cdot k \nabla & \end{array} \right]$$

$$p_{m+1}^{l+1} = \delta p_{m+1}^{l+1} + p_m^{l+1}$$

$$\mathbf{u}_{m+1}^{l+1} = \delta \mathbf{u}_{m+1}^{l+1} + \mathbf{u}_m^{l+1}$$

$$T_{m+1}^{l+1} = \delta T_{m+1}^{l+1} + T_m^{l+1}$$

## Summary: Possible enhancements of current capabilities

- ▶ Greater scalability in AMG subpreconditioner
  - ▶ Rely on improvements in *ML* and *BoomerAMG*, or
  - ▶ Implement geometric multigrid tailored to (forest of) octree hierarchy
- ▶ High order and continuous/discontinuous Galerkin discretization
- ▶ Inexact Newton-Krylov nonlinear solver
- ▶ Fully implicit time integration of the coupled mass–momentum–energy system

# Outline

Review of current capability for modeling creeping non-Newtonian Stokesian flow and transport (illustrated with mantle convection)

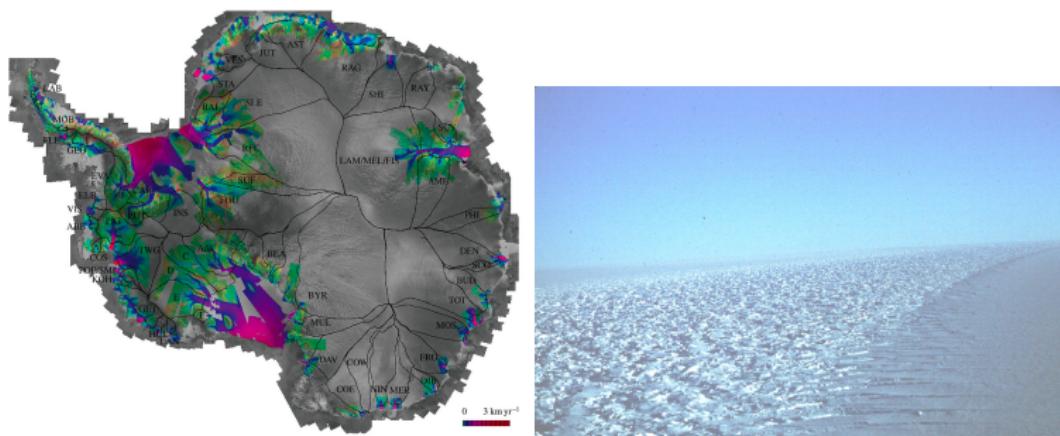
Enhancements of current implementation

New capabilities

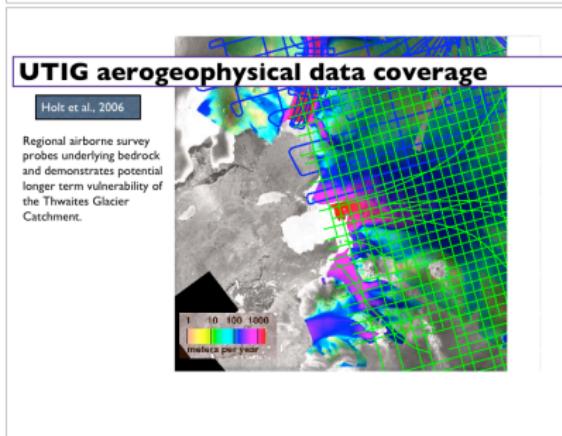
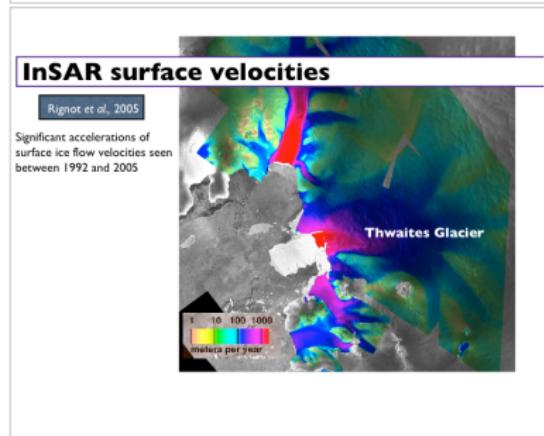
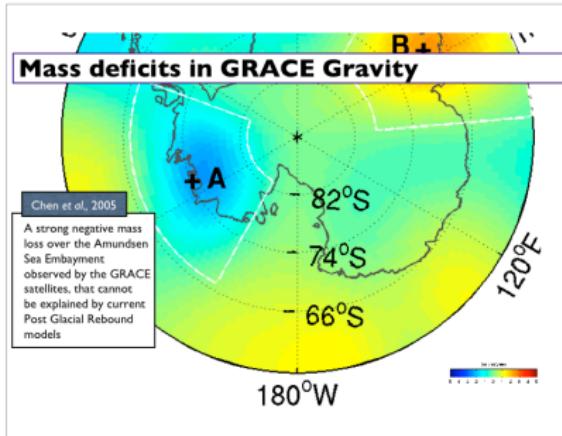
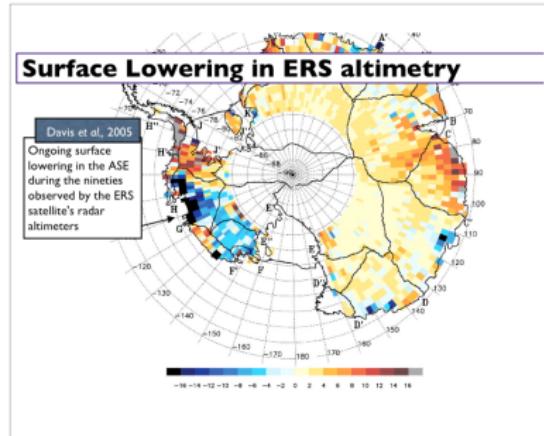
Summary

# Inverse ice sheet modeling

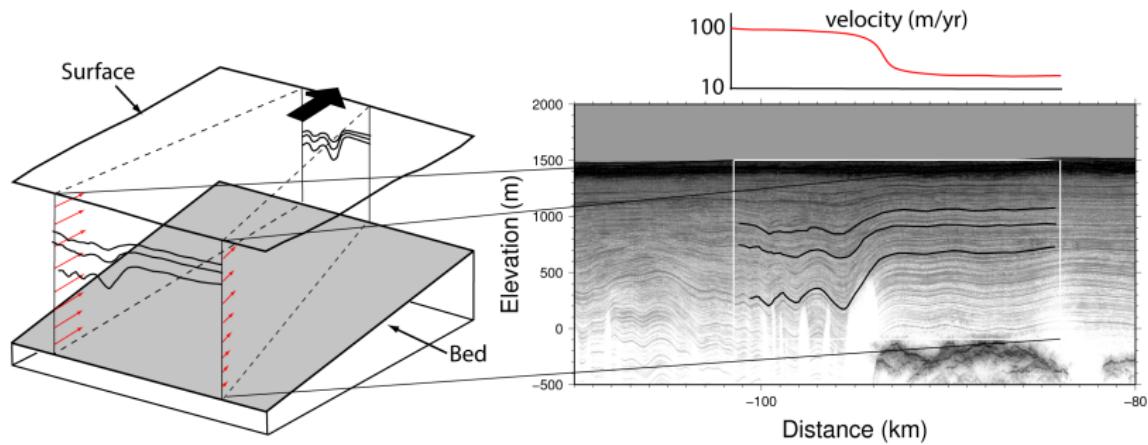
- ▶ Observations: internal layering structure (ice-penetrating radar), surface flow velocity (InSAR), age (from ice cores), surface elevation (altimetry)
- ▶ Inversion variables: constitutive parameters, basal boundary conditions, geothermal heat flux, initial temperature



# Inverse ice sheet modeling, cont.



## Inverse ice sheet modeling, cont.



UTIG ice penetrating radar

## A general ice sheet inverse problem

$$\begin{aligned} \text{minimize } \mathcal{F}(n, A_0, Q, q_B, \beta_B) := & \| \mathbf{u} - \mathbf{u}^{\text{obs}} \|_{\Gamma_{FS}} + \| z - z^{\text{obs}} \|_{\Gamma_{FS}} + \| \phi - \phi^{\text{obs}} \|_{\Omega} \\ & + \| n - n^{\text{pr}} \|_{\Omega} + \| A_0 - A_0^{\text{pr}} \|_{\Omega} + \| Q - Q^{\text{pr}} \|_{\Omega} + \| q - q^{\text{pr}} \|_{\Gamma_B} + \| \beta - \beta^{\text{pr}} \|_{\Gamma_B} \end{aligned}$$

# A general ice sheet inverse problem

minimize  $\mathcal{F}(n, A_0, Q, q_B, \beta_B) := \| \mathbf{u} - \mathbf{u}^{\text{obs}} \|_{\Gamma_{FS}} + \| z - z^{\text{obs}} \|_{\Gamma_{FS}} + \| \phi - \phi^{\text{obs}} \|_{\Omega}$   
+  $\| n - n^{\text{pr}} \|_{\Omega} + \| A_0 - A_0^{\text{pr}} \|_{\Omega} + \| Q - Q^{\text{pr}} \|_{\Omega} + \| q - q^{\text{pr}} \|_{\Gamma_B} + \| \beta - \beta^{\text{pr}} \|_{\Gamma_B}$   
subject to:

$$\nabla \cdot \mathbf{u} = 0, \quad -\nabla \cdot [\eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \mathbf{I} p] = \rho \mathbf{g}$$

# A general ice sheet inverse problem

minimize  $\mathcal{F}(n, A_0, Q, q_B, \beta_B) := \| \mathbf{u} - \mathbf{u}^{\text{obs}} \|_{\Gamma_{FS}} + \| z - z^{\text{obs}} \|_{\Gamma_{FS}} + \| \phi - \phi^{\text{obs}} \|_{\Omega}$   
+  $\| n - n^{\text{pr}} \|_{\Omega} + \| A_0 - A_0^{\text{pr}} \|_{\Omega} + \| Q - Q^{\text{pr}} \|_{\Omega} + \| q - q^{\text{pr}} \|_{\Gamma_B} + \| \beta - \beta^{\text{pr}} \|_{\Gamma_B}$   
subject to:

$$\nabla \cdot \mathbf{u} = 0, \quad -\nabla \cdot [\eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \mathbf{I} p] = \rho \mathbf{g}$$

$$\eta(T, \mathbf{u}) = \frac{1}{2} A^{-\frac{1}{n}} \dot{\varepsilon}_{\text{II}}^{\frac{1-n}{2n}}, \quad \dot{\varepsilon}_{\text{II}} = \frac{1}{2} \text{tr}(\dot{\varepsilon}^2), \quad \dot{\varepsilon} = \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T), \quad A = A_0 \exp\left(-\frac{Q}{RT}\right)$$

# A general ice sheet inverse problem

minimize  $\mathcal{F}(n, A_0, Q, q_B, \beta_B) := \| \mathbf{u} - \mathbf{u}^{\text{obs}} \|_{\Gamma_{FS}} + \| z - z^{\text{obs}} \|_{\Gamma_{FS}} + \| \phi - \phi^{\text{obs}} \|_{\Omega}$   
+  $\| n - n^{\text{pr}} \|_{\Omega} + \| A_0 - A_0^{\text{pr}} \|_{\Omega} + \| Q - Q^{\text{pr}} \|_{\Omega} + \| q - q^{\text{pr}} \|_{\Gamma_B} + \| \beta - \beta^{\text{pr}} \|_{\Gamma_B}$   
subject to:

$$\nabla \cdot \mathbf{u} = 0, \quad -\nabla \cdot [\eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \mathbf{I} p] = \rho \mathbf{g}$$

$$\eta(T, \mathbf{u}) = \frac{1}{2} A^{-\frac{1}{n}} \dot{\varepsilon}_{\text{II}}^{\frac{1-n}{2n}}, \quad \dot{\varepsilon}_{\text{II}} = \frac{1}{2} \text{tr}(\dot{\varepsilon}^2), \quad \dot{\varepsilon} = \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T), \quad A = A_0 \exp\left(-\frac{Q}{RT}\right)$$

$$\frac{Dz}{Dt}|_{\Gamma_{FS}} = a, \quad \sigma \mathbf{n}|_{\Gamma_{FS}} = \mathbf{0}, \quad \mathbf{u} \cdot \mathbf{n}|_{\Gamma_B} = 0, \quad \beta_B \mathbf{n} \times \mathbf{n} \times \sigma \mathbf{n} + \mathbf{n} \times \mathbf{n} \times \mathbf{u}|_{\Gamma_B} = \mathbf{0}$$

# A general ice sheet inverse problem

minimize  $\mathcal{F}(n, A_0, Q, q_B, \beta_B) := \| \mathbf{u} - \mathbf{u}^{\text{obs}} \|_{\Gamma_{FS}} + \| z - z^{\text{obs}} \|_{\Gamma_{FS}} + \| \phi - \phi^{\text{obs}} \|_{\Omega}$   
+  $\| n - n^{\text{pr}} \|_{\Omega} + \| A_0 - A_0^{\text{pr}} \|_{\Omega} + \| Q - Q^{\text{pr}} \|_{\Omega} + \| q - q^{\text{pr}} \|_{\Gamma_B} + \| \beta - \beta^{\text{pr}} \|_{\Gamma_B}$   
subject to:

$$\nabla \cdot \mathbf{u} = 0, \quad -\nabla \cdot [\eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \mathbf{I} p] = \rho \mathbf{g}$$

$$\eta(T, \mathbf{u}) = \frac{1}{2} A^{-\frac{1}{n}} \dot{\varepsilon}_{\text{II}}^{\frac{1-n}{2n}}, \quad \dot{\varepsilon}_{\text{II}} = \frac{1}{2} \text{tr}(\dot{\varepsilon}^2), \quad \dot{\varepsilon} = \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T), \quad A = A_0 \exp\left(-\frac{Q}{RT}\right)$$

$$\frac{Dz}{Dt}|_{\Gamma_{FS}} = a, \quad \sigma \mathbf{n}|_{\Gamma_{FS}} = \mathbf{0}, \quad \mathbf{u} \cdot \mathbf{n}|_{\Gamma_B} = 0, \quad \beta_B \mathbf{n} \times \mathbf{n} \times \sigma \mathbf{n} + \mathbf{n} \times \mathbf{n} \times \mathbf{u}|_{\Gamma_B} = \mathbf{0}$$

$$\rho c \mathbf{u} \cdot \nabla T - \nabla \cdot (K \nabla T) = \eta \text{tr}(\dot{\varepsilon}^2), \quad T|_{\Gamma_{FS}} = T_{FS}, \quad K \nabla T \cdot \mathbf{n}|_{\Gamma_B} = q_B$$

# A general ice sheet inverse problem

minimize  $\mathcal{F}(n, A_0, Q, q_B, \beta_B) := \| \mathbf{u} - \mathbf{u}^{\text{obs}} \|_{\Gamma_{FS}} + \| z - z^{\text{obs}} \|_{\Gamma_{FS}} + \| \phi - \phi^{\text{obs}} \|_{\Omega}$   
+  $\| n - n^{\text{pr}} \|_{\Omega} + \| A_0 - A_0^{\text{pr}} \|_{\Omega} + \| Q - Q^{\text{pr}} \|_{\Omega} + \| q - q^{\text{pr}} \|_{\Gamma_B} + \| \beta - \beta^{\text{pr}} \|_{\Gamma_B}$   
subject to:

$$\nabla \cdot \mathbf{u} = 0, \quad -\nabla \cdot [\eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \mathbf{I} p] = \rho \mathbf{g}$$

$$\eta(T, \mathbf{u}) = \frac{1}{2} A^{-\frac{1}{n}} \dot{\varepsilon}_{\text{II}}^{\frac{1-n}{2n}}, \quad \dot{\varepsilon}_{\text{II}} = \frac{1}{2} \text{tr}(\dot{\varepsilon}^2), \quad \dot{\varepsilon} = \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T), \quad A = A_0 \exp\left(-\frac{Q}{RT}\right)$$

$$\frac{Dz}{Dt}|_{\Gamma_{FS}} = a, \quad \sigma \mathbf{n}|_{\Gamma_{FS}} = \mathbf{0}, \quad \mathbf{u} \cdot \mathbf{n}|_{\Gamma_B} = 0, \quad \beta_B \mathbf{n} \times \mathbf{n} \times \sigma \mathbf{n} + \mathbf{n} \times \mathbf{n} \times \mathbf{u}|_{\Gamma_B} = \mathbf{0}$$

$$\rho c \mathbf{u} \cdot \nabla T - \nabla \cdot (K \nabla T) = \eta \text{tr}(\dot{\varepsilon}^2), \quad T|_{\Gamma_{FS}} = T_{FS}, \quad K \nabla T \cdot \mathbf{n}|_{\Gamma_B} = q_B$$

$$(\nabla \phi \cdot \nabla \phi)^{\frac{1}{2}} = \frac{1}{|\mathbf{u}|}, \quad \phi|_{\Gamma_{FS}} = 0$$

## Example: Inverse problem for parameter $n$

$$\text{minimize } \mathcal{F} := \| \mathbf{u} - \mathbf{u}^{\text{obs}} \| + \| n \|$$

subject to:

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

$$-\nabla \cdot [\eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \mathbf{I} p] = \rho g$$

$$\rho c \left( \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right) - \nabla \cdot (K \nabla T) = \eta \text{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

$$\eta(T, \mathbf{u}) = \frac{1}{2} A^{-\frac{1}{n}} \dot{\boldsymbol{\varepsilon}}_{\text{II}}^{\frac{1-n}{2n}}$$

$$\dot{\boldsymbol{\varepsilon}}_{\text{II}} = \frac{1}{2} \text{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

$$\dot{\boldsymbol{\varepsilon}} = \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T)$$

$$A = A_0 \exp \left( -\frac{Q}{RT} \right)$$

# Optimality conditions

state equations:

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

$$-\nabla \cdot \left[ \eta(T, \mathbf{u}) (\nabla \mathbf{u} + \nabla \mathbf{u}^T) - \mathbf{I} p \right] = \rho \mathbf{g}$$

$$\rho c \left( \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right) - \nabla \cdot (K \nabla T) = \eta \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

adjoint equations:

$$\nabla \cdot \mathbf{v} = \mathcal{D}_p \mathcal{F}$$

$$\begin{aligned} -\nabla \cdot \left[ \eta (\nabla \mathbf{v} + \nabla \mathbf{v}^T) - \mathbf{I} q \right] + \frac{\mathcal{D}_{\mathbf{u}} \eta}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T) : (\nabla \mathbf{v} + \nabla \mathbf{v}^T) &= \\ \rho c S \nabla T + S \mathcal{D}_{\mathbf{u}} \eta \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}^2) - \nabla \cdot \left[ S \eta (\nabla \mathbf{u} + \nabla \mathbf{u}^T) \right] - \mathcal{D}_{\mathbf{u}} \mathcal{F} & \\ -\rho c \left( \frac{\partial S}{\partial t} + \mathbf{u} \cdot \nabla S \right) - \nabla \cdot (K \nabla S) - S \mathcal{D}_T \eta \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}^2) &= \\ -\frac{\mathcal{D}_T \eta}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T) : (\nabla \mathbf{v} + \nabla \mathbf{v}^T) - \mathcal{D}_T \mathcal{F} & \end{aligned}$$

control equation (e.g. for parameter  $n$ ):

$$\int_{t_0}^{t_1} \left[ \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T) : (\nabla \mathbf{v} + \nabla \mathbf{v}^T) - S \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}^2) \right] \mathcal{D}_n \eta - \mathcal{D}_n \mathcal{F} = 0$$

## Mathematical and computational issues: Deterministic inversion

Beyond generic inverse solver issues, the following must be addressed in the context of the inverse ice sheet problem:

- ▶ Gradient-consistent adjoint discretization schemes
- ▶ Inexact Newton-CG for the inverse problem
- ▶ Hessian preconditioning, multilevel methods
- ▶ Appropriate regularization

## Bayesian inference for inverse problem

- ▶ Bayesian framework for statistical inverse problem: when data and/or model have uncertainties, solution of inverse problem expressed as a posterior probability density function
- ▶ Central challenge: for inverse problems characterized by high-dimensional parameter spaces, method of choice is to sample the posterior density using Markov chain Monte Carlo (MCMC)
- ▶ For inverse problems characterized by expensive forward simulations, contemporary MCMC methods become prohibitive
- ▶ Intractability of MCMC methods for large-scale statistical inverse problems can be traced to their black-box treatment of the parameter-to-observable map (the forward code)
- ▶ Goal: develop methods that exploit the structure of the parameter-to-observation map (including its derivatives), as has been done successfully in deterministic PDE-constrained optimization
  - ▶ Hessian-informed Gaussian process response surface approximation
  - ▶ Hessian-preconditioned Langevin methods

# Bayesian formulation of statistical inverse problem

Given:

$\pi_{\text{pr}}(\mathbf{x}) :=$  prior p.d.f. of model parameters  $\mathbf{x}$

$\pi_{\text{obs}}(\mathbf{y}) :=$  prior p.d.f. of the observables  $\mathbf{y}$

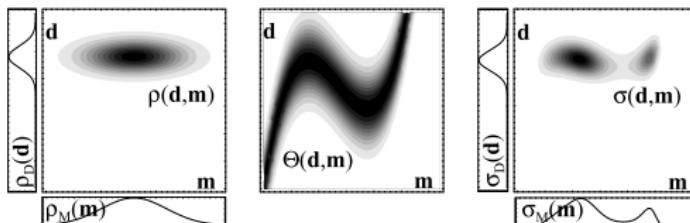
$\pi_{\text{model}}(\mathbf{y}|\mathbf{x}) :=$  conditional p.d.f. relating  $\mathbf{y}$  and  $\mathbf{x}$

Then *posterior p.d.f. of model parameters* is given by:

$$\pi_{\text{post}}(\mathbf{x}) \stackrel{\text{def}}{=} \pi_{\text{post}}(\mathbf{x}|\mathbf{y}_{\text{obs}})$$

$$\propto \pi_{\text{pr}}(\mathbf{x}) \int_{\mathcal{Y}} \frac{\pi_{\text{obs}}(\mathbf{y}) \pi_{\text{model}}(\mathbf{y}|\mathbf{x})}{\mu(\mathbf{y})} d\mathbf{y}$$

$$\propto \pi_{\text{pr}}(\mathbf{x}) \pi(\mathbf{y}_{\text{obs}}|\mathbf{x})$$



From A. Tarantola, *Inverse Problem Theory*, SIAM, 2005

## Gaussian additive noise

Given the parameter-to-observable map  $\mathbf{y} = \mathbf{f}(\mathbf{x})$ , a common noise model is Gaussian additive noise:

$$\mathbf{y}_{\text{obs}} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\text{noise}})$$

If the prior is taken as Gaussian with mean  $\mathbf{x}_{\text{pr}}$  and covariance  $\boldsymbol{\Gamma}_{\text{pr}}$ , then the posterior can be written

$$\pi_{\text{post}}(\mathbf{x}) \propto \exp \left( -\frac{1}{2} \| \mathbf{f}(\mathbf{x}) - \mathbf{y}_{\text{obs}} \|_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}}^2 - \frac{1}{2} \| \mathbf{x} - \mathbf{x}_{\text{pr}} \|_{\boldsymbol{\Gamma}_{\text{pr}}^{-1}}^2 \right)$$

Note that “most likely” point is given by

$$\begin{aligned} \mathbf{x}_{\text{MAP}} &\stackrel{\text{def}}{=} \arg \max_{\mathbf{x}} \pi_{\text{post}}(\mathbf{x}) \\ &= \arg \min_{\mathbf{x}} \frac{1}{2} \| \mathbf{f}(\mathbf{x}) - \mathbf{y}_{\text{obs}} \|_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}}^2 + \frac{1}{2} \| \mathbf{x} - \mathbf{x}_{\text{pr}} \|_{\boldsymbol{\Gamma}_{\text{pr}}^{-1}}^2 \end{aligned}$$

This is an (appropriately weighted) deterministic inverse problem!

## Gaussian additive noise, linear inverse problem

Suppose further the parameter-to-observable map is linear, i.e.

$$\mathbf{y} = \mathbf{F}\mathbf{x}$$

Then the posterior can be written

$$\pi_{\text{post}}(\mathbf{x}) \propto \exp \left( -\frac{1}{2} \| \mathbf{F}\mathbf{x} - \mathbf{y}_{\text{obs}} \|_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}}^2 - \frac{1}{2} \| \mathbf{x} - \mathbf{x}_{\text{pr}} \|_{\boldsymbol{\Gamma}_{\text{pr}}^{-1}}^2 \right)$$

The posterior is then Gaussian with

$$\mathbf{x} \sim \mathcal{N}(\mathbf{x}_{\text{MAP}}, \boldsymbol{\Gamma}_{\text{post}})$$

The covariance is the inverse Hessian of the negative log posterior:

$$\begin{aligned} \boldsymbol{\Gamma}_{\text{post}}^{-1} &= \mathbf{F}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{F} + \boldsymbol{\Gamma}_{\text{pr}}^{-1} \\ &= \nabla_{\mathbf{x}}^2 (-\log \pi_{\text{post}}) \end{aligned}$$

I.e., the covariance is given by the inverse Hessian of the regularized misfit function that is minimized by deterministic methods

# Outline

Review of current capability for modeling creeping non-Newtonian Stokesian flow and transport (illustrated with mantle convection)

Enhancements of current implementation

New capabilities

Summary

# Summary of mathematical/computational research agenda

- ▶ Overall goal: Scalable, parallel, adaptive full Stokes ice sheet simulator equipped with deterministic and statistical inversion capabilities
- ▶ Base level enhancements to existing nonlinear Stokes code
  - ▶ free surface
  - ▶ ice constitutive laws
- ▶ Further forward solver enhancements (as needed)
  - ▶ improved scalability of AMG preconditioner for  $> 10^4$  cores
  - ▶ high order spectral element DG/CG discretization
  - ▶ fully implicit time integration
  - ▶ Newton solver
- ▶ Deterministic inversion
  - ▶ gradient-consistent/adjoint-appropriate discretization
  - ▶ globalization of inexact Newton-CG solver
  - ▶ multilevel Hessian preconditioner
  - ▶ Regularization
- ▶ Statistical inversion
  - ▶ reduce-then-sample: Gaussian process response surface methods
  - ▶ sample-then-reduce: Langevin-based proposals for MCMC