Parallel High-Order Geometric Multigrid Methods on Adaptive Meshes for Highly Heterogeneous Nonlinear Stokes Flow Simulations of Earth's Mantle

Johann Rudi*, Hari Sundar†, Tobin Isaac*, Georg Stadler*, Michael Gurnis‡, and Omar Ghattas*§¶
*Institute for Computational Engineering and Sciences, The University of Texas at Austin, USA
†School of Computing, The University of Utah, USA
‡Seismological Laboratory, California Institute of Technology, USA
§Department of Mechanical Engineering, The University of Texas at Austin, USA
¶Jackson School of Geosciences, The University of Texas at Austin, USA

Poster Summary

We address the problem of modeling global-scale, convection-driven flow in Earth's mantle with associated plate motions. Modeling global mantle flow is a challenging problem due to several inherent difficulties: (1) the severe nonlinearity of the rheology of the mantle, (2) the orders-of-magnitude viscosity contrast and sharp viscosity gradients, and (3) the widely-varying spatial scales. We present scalable numerical methods and their efficient parallel implementation for solving global mantle flow problems that represent a significant improvement over our previous work (e.g., see [1]) in key areas as the discretization as well as the linear and nonlinear solver. In particular, we develop parallel scalable geometric multigrid methods for preconditioning the linearized Stokes systems that arise upon discretization with high-order finite elements on adaptive meshes. We provide numerical results based on real Earth data that are likely the most realistic global scale mantle flow simulations conducted to date, and we demonstrate scalability results on up to 16,384 cores.

Mantle flow is governed by nonlinear incompressible Stokes equations with shear-thinning rheology,

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

with temperature T, velocity $\boldsymbol{u}$, pressure p, and effective viscosity μ. The viscosity depends exponentially on the temperature (via an Arrhenius relationship), on a power of the second invariant of the strain rate tensor, and incorporates plastic yielding. The right-hand side forcing, $\boldsymbol{f}$, is derived from the Boussinesq approximation and depends on the temperature. We employ an inexact Newton-Krylov method to solve the discretized nonlinear equations, with the advantages of superlinear (or even quadratic) convergence, mesh size independence, and avoidance of oversolving in early iterations [2]. Most of the solver time is spent in the inner Krylov loop and therefore it is crucial to find effective linear preconditioners. Our discretization employs high-order finite elements on highly adaptively refined hexahedral meshes, which allow us to resolve plate boundaries down to a few hundred meters while still capturing global-scale behavior. We obtain the discrete Stokes system,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}.$$

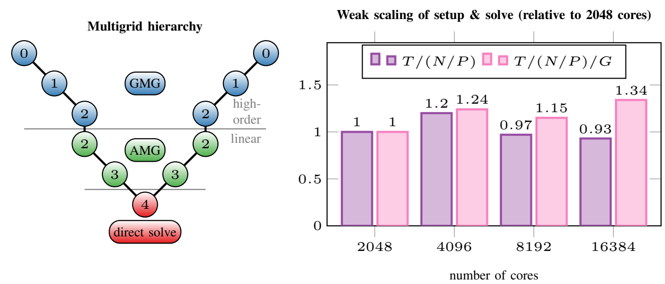

Figure 1. Left: Diagram of geometric-algebraic multigrid hierarchy. Initially the adaptive mesh is coarsened geometrically (GMG) and repartitioned uniformly among all cores at each level. High-order operators are used for smoothing and residual computation. When the problem becomes small enough, algebraic multigrid (AMG) is used on a small subset of cores and for a linear finite element discretization. Right: Weak scaling of geometric-algebraic multigrid (solution of $\mathbf{Au} = \mathbf{f}$). Normalized scaling of algorithm and implementation is $T/(N/P)$ and implementation only is $T/(N/P)/G$, where T is runtime, N is degrees of freedom, P is number of cores, and G is number of iterations. The largest problem has 5.4B degrees of freedom.

Krylov methods are advantageous because they permit matrix-free Jacobian applications. Within the matrix-free apply routines, we exploit the tensor product structure of the velocity finite element shape functions to increase the floating point to memory operations ratio. We use a GMRES linear solver with right preconditioning based on the upper triangular block matrix [5],

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

Here, approximations of the inverse of the viscous block, $\tilde{\mathbf{A}}^{-1} \approx \mathbf{A}^{-1}$, and the inverse of the Schur complement, $\tilde{\mathbf{S}}^{-1} \approx \mathbf{S}^{-1} := (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top)^{-1}$, are required.

As mentioned previously, adaptive mesh refinement (AMR) is important because the locally highly-varying viscosity needs to be resolved by the mesh (down to ∼0.5 km resolution). We use the `p4est` parallel AMR library, which provides hexahedral meshes with a forest-of-octrees structure and space filling curves for fast, parallel neighbor search, partitioning, and 2:1 balancing [3], [4]. The `p4est` library has been demonstrated to scale well to $O(500\text{K})$ processor cores.

The inverse of the viscous block, $\tilde{\mathbf{A}}^{-1}$, is approximated by an efficient, collective communication-free, parallel implementation of geometric multigrid for high-order finite elements on adaptive meshes. This is a generalization of the geometric multigrid method in [6] for the Poisson operator using linear elements. We use a hybrid geometric-algebraic multigrid approach where the mesh is coarsened and repartitioned geometrically first, and then once the problem size is small enough, we revert to algebraic multigrid (PETSc's GAMG) using only a small subset of cores (figure 1, left). On the geometric multigrid levels, matrix-free apply functions for the viscous block have the advantage of being faster and more memory efficient than using assembled matrices. By limiting algebraic multigrid to small problems on small core counts we can avoid its high setup cost. Moreover, the assembled matrix for algebraic multigrid is obtained from a linear finite element discretization on the high-order nodes, which is beneficial for the sparsity of the matrix as well as multigrid convergence.

The weak scaling of this high-order geometric multigrid preconditioner when solving the system $\mathbf{A}\mathbf{u} = \mathbf{f}$ is shown in the chart on the right of figure 1. We consider the solution to be converged after reducing the residual by six orders of magnitude. The numerical experiments were carried out on TACC's Stampede supercomputer on up to 16,384 cores with problem sizes up to 5.4B degrees of freedom. The normalized scalability of the numerical algorithms and the implementation is better than optimal (i.e., <1.00 on 16,384 cores) since the number of iterations for convergence is decreasing as we refine the mesh and resolve the variations in the viscosity better. The scalability of the implementation only, which is normalized with respect to the number of iterations, is also shown in figure 1.

For solving the full Stokes system, it remains to find an effective preconditioner for the inverse of the Schur complement $\tilde{\mathbf{S}}^{-1}$. We use a method that is robust with respect to extreme viscosity variations. It is an improved version of BFBT / Least Squares Commutator methods (based on [7], [8]) and is defined as

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

where $\mathbf{D} := \mathrm{diag}(\mathbf{A})$. Using geometric multigrid for the pressure Laplacian, $(\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^{\top})^{-1}$, is problematic due to the discontinuous modal discretization of the pressure function space. We take a novel approach by re-discretizing the the Laplacian operator with continuous nodal basis functions. This continuous nodal Laplacian operator is then approximately inverted with geometric multigrid. As a result, similar scalability results as demonstrated in figure 1 hold.

Finally, we present results on solver convergence for the nonlinear global mantle flow problem using an inexact Newton-Krylov method on 4096 cores. The graph in figure 2 shows the reduction of the nonlinear (Newton) residual in red and the convergence of the linear (Krylov) residual in blue. The nonlinear solver was terminated after reducing the nonlinear residual by eight orders of magnitude. The nonlinear dependence of the viscosity on the strain rate creates sharper viscosity gradients. Therefore, a grid continuation method is utilized, where the mesh is refined adaptively after the first four Newton steps (indicated by black vertical lines). This prevents stalling of the linear convergence. The problem has 642M velocity and pressure degrees of freedom at the converged solution. The surface velocities of the Earth at the solution are shown on the right of figure 2. Due to the robustness of the preconditioners, we can observe steady linear convergence, which is crucial to obtain sufficiently accurate Newton steps, especially at the last couple of Newton steps when the linear systems need to be solved with higher precision.

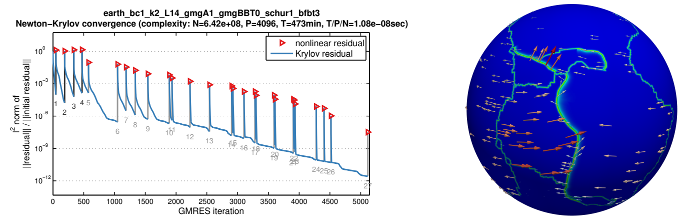

Figure 2. Left: Convergence of inexact Newton-Krylov method for nonlinear global mantle flow problem on 4096 cores. Red triangles indicate the nonlinear (Newton) residual ℓ^2-norm and blue lines represent the linear (Krylov) convergence in ℓ^2-norm. Each Newton step is separated by a gray vertical line. The mesh is adaptively refined after the first four Newton steps (black vertical lines) to resolve the variations in the viscosity due to the nonlinearity. Right: Surface velocities of the Earth (around South America) at solution of nonlinear global mantle flow problem.

In conclusion, we have shown that we are able to overcome the challenges of global mantle flow problems by designing solvers that combine two essential components: powerful numerical methods and their efficient scalable parallel implementation. Inexact Newton-Krylov methods are employed to handle the severe nonlinear rheology. AMR is used to resolve the locally highly-varying viscosity, as well as between Newton steps as a grid continuation technique, resulting in effective capture of global scale behavior. Central for reducing the time to solution is the effectiveness and robustness of the preconditioners for $\tilde{\mathbf{A}}^{-1}$ and $\tilde{\mathbf{S}}^{-1}$, which are built on high-order multigrid methods. This combination of inexact Newton-Krylov, high-order multigrid, and AMR permits us to carry out the most realistic global mantle flow simulations to date.

References

[1] C. Burstedde, O. Ghattas, M. Gurnis, E. Tan, T. Tu, G. Stadler, L. C. Wilcox, and S. Zhong, *Scalable adaptive mantle convection simulation on petascale supercomputers*, SC08: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2008.

[2] S. C. Eisenstat and H. F. Walker, *Globally convergent inexact Newton methods*, SIAM Journal on Optimization, 4 (1994), pp. 393–422.

[3] C. Burstedde, O. Ghattas, M. Gurnis, T. Isaac, G. Stadler, T. Warburton, and L. C. Wilcox, *Extreme-scale AMR*, SC10: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2010.

[4] C. Burstedde, L. C. Wilcox, and O. Ghattas, *`p4est`: Scalable algorithms for parallel adaptive mesh refinement on forests of octrees*, SIAM Journal on Scientific Computing, 33 (2011), pp. 1103-1133.

[5] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1–137.

[6] H. Sundar, G. Biros, C. Burstedde, J. Rudi, O. Ghattas, and G. Stadler, *Parallel geometric-algebraic multigrid on unstructured forests of octrees*, SC12: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2012.

[7] H. Elman, V. Howle, J. Shadid, R. Shuttleworth, and R. Tuminaro, *Block preconditioners based on approximate commutators*, SIAM Journal on Scientific Computing, 27 (2006), pp. 1651–1668.

[8] D. A. May and L. Moresi, *Preconditioned iterative methods for Stokes flow problems arising in computational geodynamics*, Physics of the Earth and Planetary Interiors, 171 (2008), pp. 33–47.