Algorithmic reuse for non-smooth problems in heterogeneous media

**Jed Brown** jedbrown@mcs.anl.gov (ANL and CU Boulder)
Mark Adams (LBL), Matt Knepley (UChicago)

PMAA, Lugano, 2014-07-02

This talk: [http://59A2.org/files/20140702-PMAA.pdf](http://59A2.org/files/20140702-PMAA.pdf)
Plan: ruthlessly eliminate communication

Why?
- Local recovery despite global coupling
- Tolerance for high-frequency load imbalance
  - From irregular computation or hardware error correction
- More scope for dynamic load balance

Requirements
- Must retain optimal convergence with good constants
- Flexible, robust, and debuggable
Multigrid Preliminaries

Multigrid is an $O(n)$ method for solving algebraic problems by defining a hierarchy of scale. A multigrid method is constructed from:

1. a series of discretizations
   - coarser approximations of the original problem
   - constructed algebraically or geometrically

2. intergrid transfer operators
   - residual restriction $I_h^H$ (fine to coarse)
   - state restriction $\hat{I}_h^H$ (fine to coarse)
   - partial state interpolation $I_h^H$ (coarse to fine, ‘prolongation’)
   - state reconstruction $\mathbb{I}_h^H$ (coarse to fine)

3. Smoothers ($S$)
   - correct the high frequency error components
   - Richardson, Jacobi, Gauss-Seidel, etc.
   - Gauss-Seidel-Newton or optimization methods
τ formulation of Full Approximation Scheme (FAS)

- classical formulation: “coarse grid accelerates fine grid”
- τ formulation: “fine grid feeds back into coarse grid”
- To solve \( Nu = f \), recursively apply

\[
\text{pre-smooth: } \tilde{u}^h \leftarrow S^h_{\text{pre}}(u^h_0, f^h)
\]

\[
\text{solve coarse problem for } u^H \quad N^H u^H = I^H_h f^h + N^H \hat{I}^H_h \tilde{u}^h - I^H_h N^h \tilde{u}^h
\]

\[
\text{correction and post-smooth: } u^h \leftarrow S^h_{\text{post}} \left( \tilde{u}^h + I^H_h (u^H - \hat{I}^H_h \tilde{u}^h), f^h \right)
\]

- At convergence, \( u^{H*}_h = \hat{I}^H_h u^{h*}_h \) solves the τ-corrected coarse grid equation \( N^H u^H = f^H + \tau^H_h \), thus \( \tau^H_h \) is the “fine grid feedback” that makes the coarse grid equation accurate.
- \( \tau^H_h \) is local and need only be recomputed where it becomes stale.
\( \tau \) corrections

- Plane strain elasticity, \( E = 1000, \nu = 0.4 \) inclusions in \( E = 1, \nu = 0.2 \) material, coarsen by \( 3^2 \).

- Solve initial problem everywhere and compute
  \[ \tau_h^H = A^H \hat{I}_h^H u^h - I_h^H A^h u^h \]

- Change boundary conditions and solve FAS coarse problem
  \[ N^H \dot{u}^H = I_h^H \dot{f}^h + N^H \hat{I}_h^H \ddot{u}^h - I_h^H N^h \ddot{u}^h \]

- Prolong, post-smooth, compute error \( e^h = \dot{u}^h - (N^h)^{-1} \dot{f}^h \)

- Coarse grid \textit{with} \( \tau \) is nearly \( 10 \times \) better accuracy
\[ \begin{align*}
\tau \text{ corrections} \\
\text{ ■ Plane strain elasticity, } E = 1000, \nu = 0.4 \text{ inclusions in } E = 1, \nu = 0.2 \text{ material, coarsen by } 3^2. \\
\text{ ■ Solve initial problem everywhere and compute } \\
\tau^H_h = A^H \hat{I}_h^H u^h - I^H_h A^h u^h \\
\text{ ■ Change boundary conditions and solve FAS coarse problem } \\
N^H H \dot{u}^H = \underbrace{I^H_h \dot{f}^h}_f^H + \underbrace{N^H H \hat{I}_h^H \hat{u}^h - I^H_h N^h \hat{u}^h}_\tau^H_h \\
\text{ ■ Prolong, post-smooth, compute error } e^h = \dot{u}^h - (N^h)^{-1} \dot{f}^h \\
\text{ ■ Coarse grid } \text{with } \tau \text{ is nearly } 10 \times \text{ better accuracy}
\end{align*} \]
τ adaptivity: an idea for heterogeneous media

- Applications
  - Geo: reservoir engineering, lithosphere dynamics (subduction, rupture/fault dynamics)
  - carbon fiber, biological tissues, fracture
  - Conventional adaptivity fails

- Traditional adaptive methods fail
  - Solutions are not “smooth”
  - Cannot build accurate coarse space without scale separation

- τ adaptivity
  - Fine-grid work needed everywhere at first
  - Then τ becomes accurate in nearly-linear regions
  - Only visit fine grids in “interesting” places: active nonlinearity, drastic change of solution
Comparison to nonlinear domain decomposition

- **ASPIN (Additive Schwarz preconditioned inexact Newton)**
  - Cai and Keyes (2003)
  - More local iterations in strongly nonlinear regions
  - Each nonlinear iteration only propagates information locally
  - Many real nonlinearities are activated by long-range forces
    - locking in granular media (gravel, granola)
    - binding in steel fittings, crack propagation
  - Two-stage algorithm has different load balancing
    - Nonlinear subdomain solves
    - Global linear solve

- τ adaptivity
  - Minimum effort to communicate long-range information
  - Nonlinearity sees effects as accurate as with global fine-grid feedback
  - Fine-grid work always proportional to “interesting” changes
Low communication MG

- **red arrows** can be removed by $\tau$-FAS with overlap
- **blue arrows** can also be removed, but then algebraic convergence stalls when discretization error is reached
- no simple way to check that discretization error is obtained
- if fine grid state is not stored, use compatible relaxation to complete prolongation $P$
Nonlinear and matrix-free smoothing

- matrix-based smoothers require global linearization
- nonlinearity often more efficiently resolved locally
- nonlinear additive or multiplicative Schwarz
- nonlinear/matrix-free is good if

\[ C = \frac{(\text{cost to evaluate residual at one point}) \cdot N}{(\text{cost of global residual})} \sim 1 \]

- finite difference: \( C < 2 \)
- finite volume: \( C \sim 2 \), depends on reconstruction
- finite element: \( C \sim \) number of vertices per cell

- larger block smoothers help reduce \( C \)
- additive correction like Jacobi reduces \( C \), but need to assemble corrector/scaling
Multiscale compression and recovery using $\tau$ form

-Checkpoint converged coarse state
-Recovers using FMG anchored at $\ell_{cp} + 1$
-Needs only $\ell_{cp}$ neighbor points
-$\tau$ correction is local

Solve $F(u^n; b^n) = 0$

Next solve

$\ell_{fine}$

$\tau$

$\ell_{cp} + 1$

$\tau$

$\ell_{cp}$

$\tau$

$\ell_{cp}$

Normal multigrid cycles visit all levels moving from $n \rightarrow n + 1$

FMG recovery only accesses levels finer than $\ell_{CP}$

Lightweight checkpointing for transient adjoint computation

Postprocessing applications, e.g., in-situ visualization at high temporal resolution in part of the domain
Outlook on $\tau$-FAS adaptivity and compression

- Benefits of AMR without fine-scale smoothness
- Coarse-centric restructuring is a major interface change
- Nonlinear smoothers (and discretizations)
  - Smooth in neighborhood of “interesting” fine-scale features
  - Which discretizations can provide efficient matrix-free smoothers?
- Dynamic load balancing
- Reliability of error estimates for refreshing $\tau$
  - We want a coarse indicator for whether $\tau$ needs to change
- Worthwhile for resilience and to better use hardware