SAND2013-6611C

Amortizing AMG Components Across Problem Sequences

Ray Tuminaro, Sandia National Labs

Collaborators:

Hu, Prokopenko, Siefert, Tsuji (Sandia)

Boman, Cyr, Lin, Shadid (Sandia), Gee, Wiesner (TU Munich), Olson (U. Illinois), Schroder (LLNL), Keyes (KAUST) , Gaidamour (CNRS), Maclachlan, Benson (Tufts), Shank (Temple), Waisman (Columbia U), Hiriyur(Weidlinger), Gerstenberg (Rolls Royce)

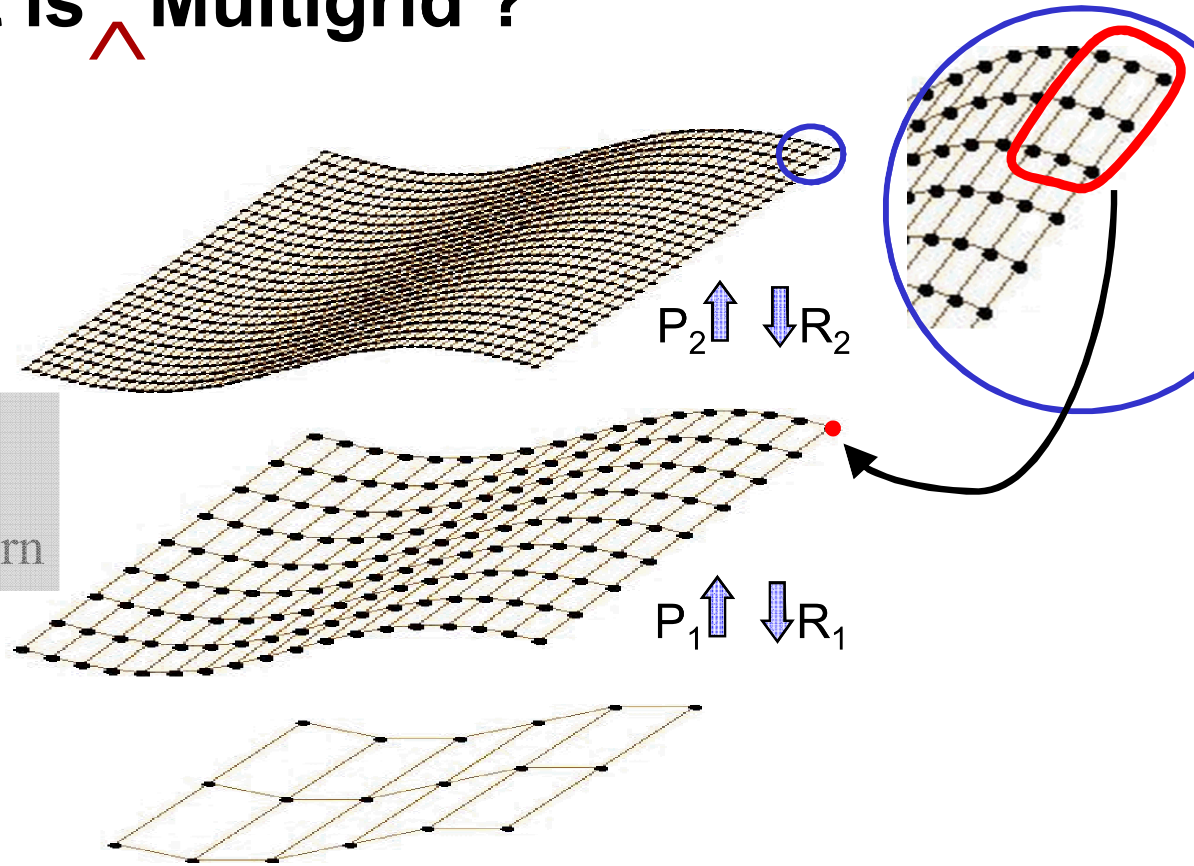
Sandia National Laboratories is a multi-program laboratory operated by Sandia Corp, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Outline

- AMG background
  - popular approaches: inexpensive, often effective, but rigid
- Energy minimization AMG
  - arbitrary coarsening, flexible coarse basis function support
- Energy minimization costs & amortization
  - Reuse & initial guess choice
- Leveraging flexibility
  - structured AMG for exascale architectures
  - mixed finite elements,
  - extended finite elements & anisotropic PDEs

What is Algebraic Multigrid ?

Solve $A_3 u_3 = f_3$

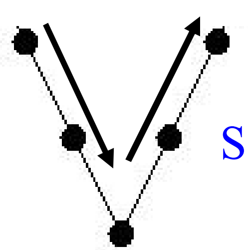


- Construct Graph & Coarsen
- Determine P_i & R_i sparsity pattern
- Determine P_i & R_i's coefs
- Project: $A_i = R_i A_{i+1} P_i$

Smooth $A_3 u_3 = f_3$. Set $f_2 = R_2 r_3$.

Smooth $A_2 u_2 = f_2$. Set $f_1 = R_1 r_2$.

Solve $A_1 u_1 = f_1$ directly.



Set $u_2 = u_2 + P_1 u_1$. Smooth $A_2 u_2 = f_2$.

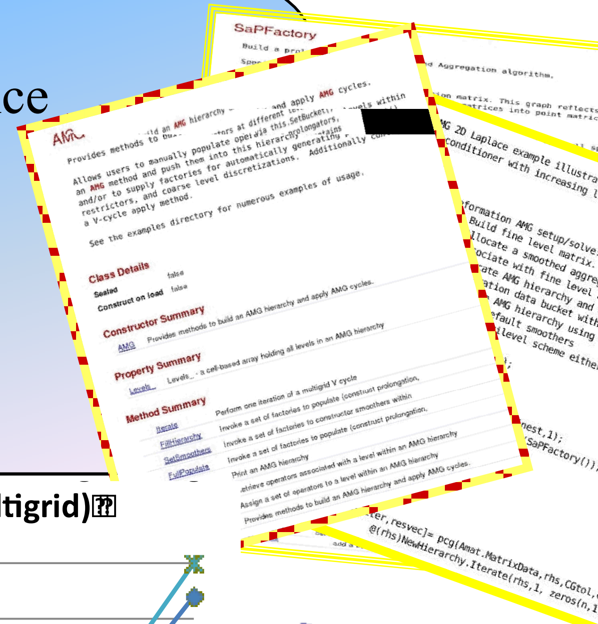
Set $u_3 = u_3 + P_2 u_2$. Smooth $A_3 u_3 = f_3$.

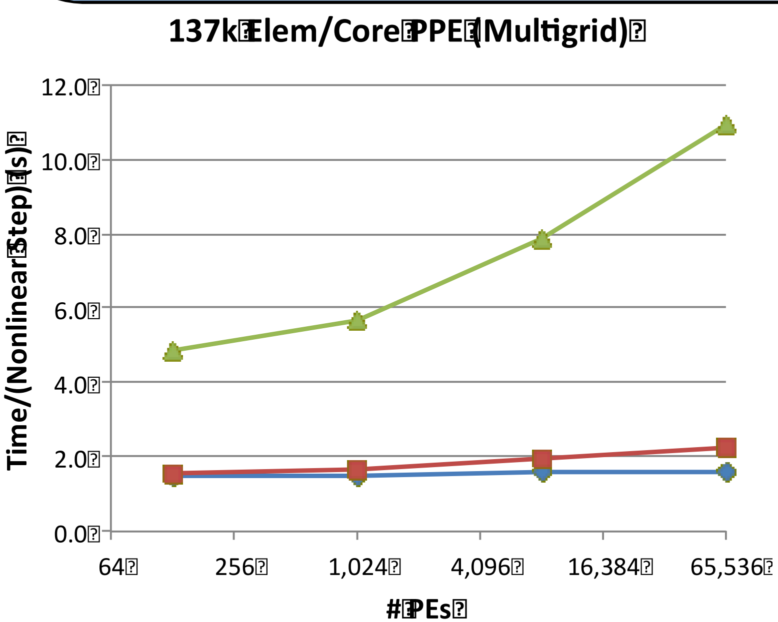
Sandia National Laboratories

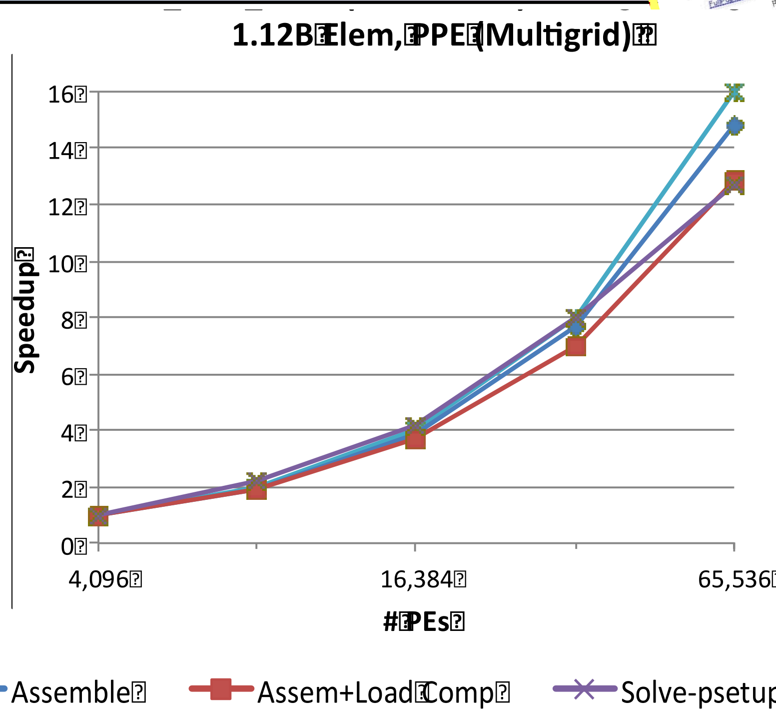
MueMat/ MueLu

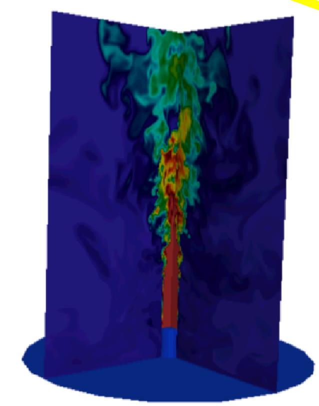
```
           (__)
           (oo)   "MueMat"
    /-------\/
   / |     ||
  *  ||----||
     ~~    ~~
```

- new AMG algorithms, e.g. energy minimization
- leverages Trilinos/Kokkos for multi-core/GPU performance
- templating for mixed precision
- designed for flexibility
  - *special* application circumstances
  - architecture considerations
  - reuse









9 billion dofs

AMG limitations …

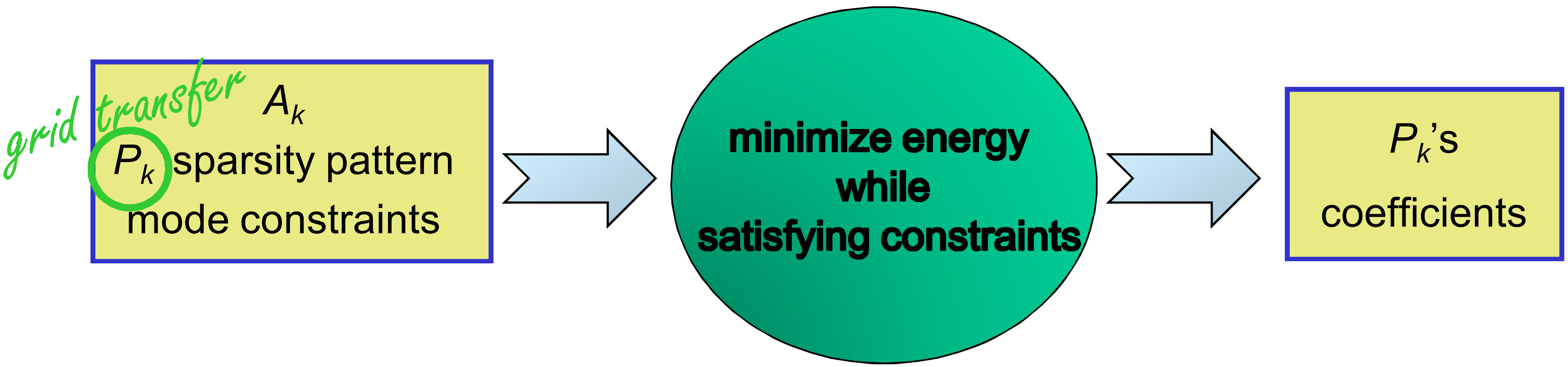
Best understood theoretically for
scalar elliptic PDEs with *standard* discretizations
(e.g. linear nodal FE & finite differences),
but often works on broader range of systems.

Deficiency: **Classical AMG & smoothed aggregation are rigid & not easily adapted to *advanced* situations.**

- coarsening rules: diameter 3 aggregates or
  - (C1) for each $i \in F$, each point $j \in S_i$ should either be in C, or should be strongly connected to at least one point in $C \cap S_i$
  - (C2) C should be a maximal subset of all points with the property that no two C-points are strongly connected to each other
- coarsening, P sparsity pattern, p_{ij} choices are often tied together
- strong/weak decisions influence multiple phases of algorithm
- accurate interpolation of constants automatically addressed, but considering other important modes can be problematic

Sandia National Laboratories

AMG & Energy Minimization



Tradeoffs:

- + flexibility
  - any coarsening
  - any sparsity pattern
  - constraints
    - important modes requiring accurate interpolation
- + robustness

$$\min_P \; p_1^T A p_1 + p_2^T A p_2 + p_3^T A p_3 + \ldots$$

$$P = [p_1 \;\; p_2 \;\; p_3 \;\; \ldots]$$

Brandt, Brannick, Brezina, Chan, Gee, Kahl, Kolev, Livshits, Mandel, Olson, Schroder, Smith, T, Vanek, Vassilevski, Wagner, Wall, Wan, Wiesner, Xu, Zikatanov

Sandia National Laboratories

“Solve” AP = 0

1) with minimization algorithm

2) in space satisfying constraints

or equivalently

$$\text{“solve”}\ \hat{A}\,\mathbf{P} = 0 \quad \text{where} \quad \hat{A} = \begin{pmatrix} A & & & \\ & A & & \\ & & A & \\ & & & \ddots \end{pmatrix} \qquad \mathbf{P} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \end{pmatrix}$$

and $P = [\,p_1\ p_2 \cdots p_m\,]$

Minimization candidates include *CG & Chebyshev, GMRES & CGNR*

Constraints

– $P\ B_c = B \Rightarrow\ X\,\mathbf{P} = g$ *via* $Q = (\,I - X^T(X\,X^T)^{-1}\,X\,)$

Constraint Satisfying Space *might* be

$$Q\,\hat{A}\,\mathbf{P_0},\ (Q\hat{A})^2\mathbf{P_0},\ (Q\hat{A})^3\mathbf{P_0},\ (Q\hat{A})^4\mathbf{P_0},\ \ldots$$

Sandia National Laboratories

Prolongator construction cost

- Costs:

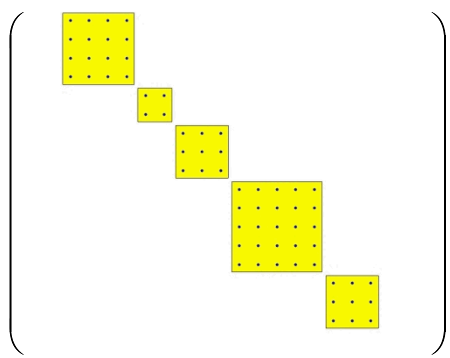
1 Krylov iteration $\begin{cases} \text{Matrix-matrix multiply: A P} \\ \text{Remove nnzs beyond desired sparsity pattern} \\ \text{Apply Q} = I - X^T(X X^T)^{-1}X \end{cases}$

- Practical algorithms follow …

  – Few iterations needed, e.g. Smoothed Aggregation

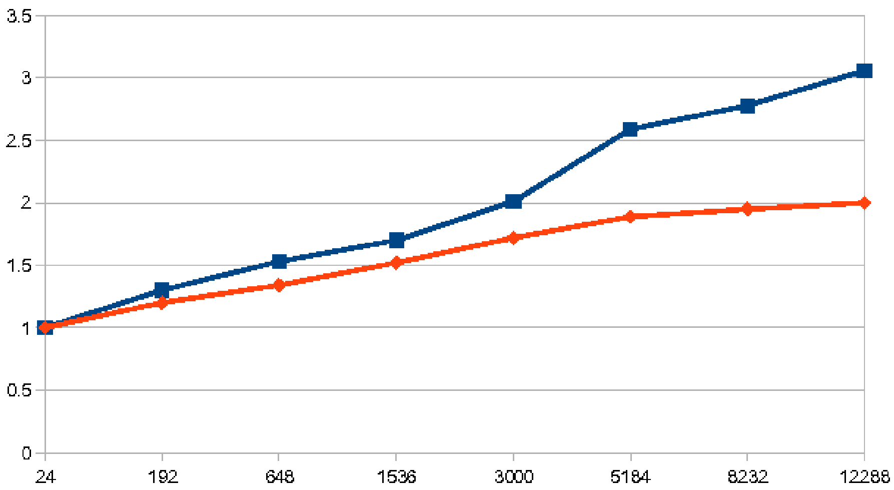
  – *Computation of* $(X X^T)^{-1}$ easy as

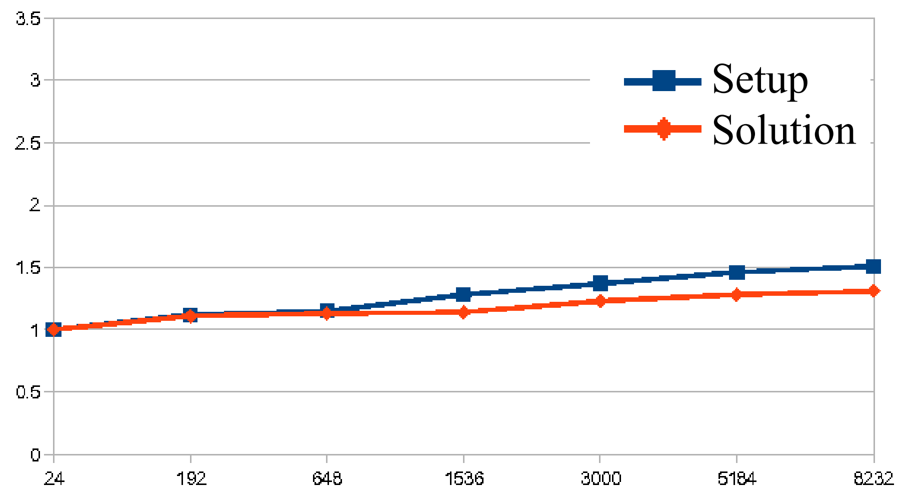
$X X^T =$



  – Amortization opportunities: init. guess, reuse of $X X^T$

MueLu *Emin* Setup Scaling





<u>3D Laplace via 7 pt stencil</u>

- 2 CG *Emin* iterations
- P_0 via tentative prolongator
- same sparsity pattern as SA

<u>3D elasticity with Poisson ratio .25</u>

- 2 CG *Emin* iterations
- P_0 via tentative prolongator
- same sparsity pattern as SA

2 *Emin* iterations typically fine for AMG convergence

SA sparsity pattern $\Rightarrow$ solver cost comparable with SA

1 *Emin* iter. $\Rightarrow$ Setup cost $\approx$1.5-2x > SA setup $\Rightarrow$ not too significant

Sandia National Laboratories

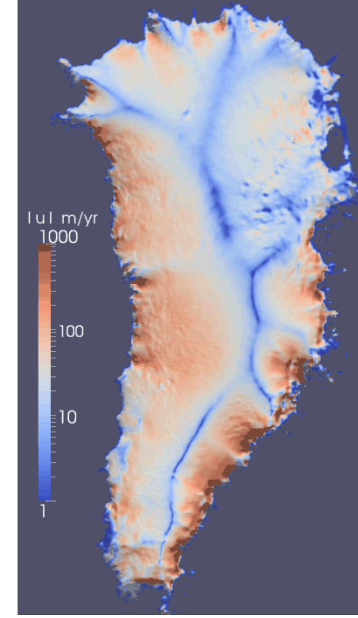
Setup Amortization

- E_{min} generally more expensive than SA
- Typically solve a sequence of related linear systems
  - time stepping, Newton, continuation, inverse problems, UQ, …
- Many setup components can often be reused
  - P_0, sparsity pattern, matrix graphs, constraint builder, even some substeps of matrix-matrix multiply
  - depends on changes to …

    mesh & matrix coefficients and strength-of-connection
- Previous P_{final} can be used for next $P_0 \Rightarrow$ fewer E_{min} iterations

Sandia National Laboratories

Ice Sheet Amortization Test Case



Two coupled non-linear PDEs

$$\begin{cases} -\nabla\cdot(2\mu\dot{\boldsymbol{\epsilon}}_1) & = & -\rho g\,\frac{\partial s}{\partial x}, \\ -\nabla\cdot(2\mu\dot{\boldsymbol{\epsilon}}_2) & = & -\rho g\,\frac{\partial s}{\partial y} \end{cases}$$

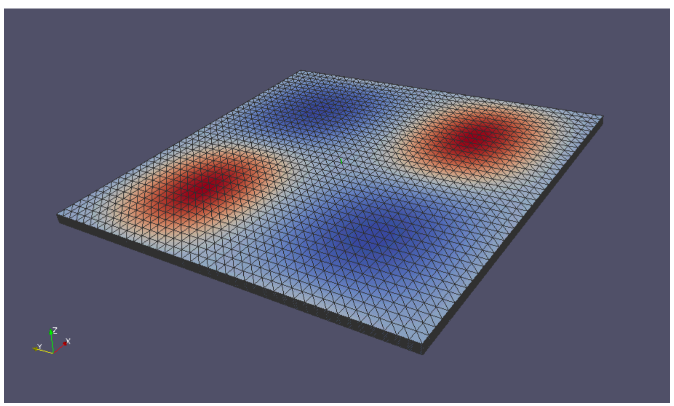
with Glen's law viscosity

$$\mu = \frac{1}{2}A^{-\frac{1}{n}}\left(\dot{\epsilon}_{xx}^2 + \dot{\epsilon}_{yy}^2 + \dot{\epsilon}_{xx}\dot{\epsilon}_{yy} + \dot{\epsilon}_{xy}^2 + \dot{\epsilon}_{xz}^2 + \dot{\epsilon}_{yz}^2 + \gamma\right)^{\left(\frac{1}{2n}-\frac{1}{2}\right)}$$

$$\dot{\boldsymbol{\epsilon}}_1^T = \left(\; 2\dot{\epsilon}_{xx} + \dot{\epsilon}_{yy}, \quad \dot{\epsilon}_{xy}, \quad \dot{\epsilon}_{xz} \;\right)$$

$$\dot{\boldsymbol{\epsilon}}_2^T = \left(\; \dot{\epsilon}_{xy}, \quad \dot{\epsilon}_{xx} + 2\dot{\epsilon}_{yy}, \quad \dot{\epsilon}_{yz} \;\right)$$

$$\dot{\epsilon}_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$$



$$\begin{aligned} A &= \text{flow rate factor} \\ n &= \text{Glen's law exponent} = 3 \\ \gamma &= \text{regularization parameter} \\ \beta &= \text{sliding coefficient} \geq 0 \end{aligned}$$

Newton's method & continuation used

Sandia National Laboratories

Ice Sheet Results

Step	Emin(6)	Emin(1)	Emin(6,1)
2	17	30	17
8	16	32	17
12	17	33	18
18	17	36	18
23	17	36	18
28	17	34	18

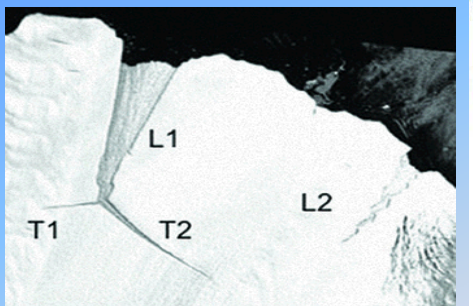
- Emin(k): k E_{min} iterations, no reuse
- Emin(k,1) k E_{min} iterations

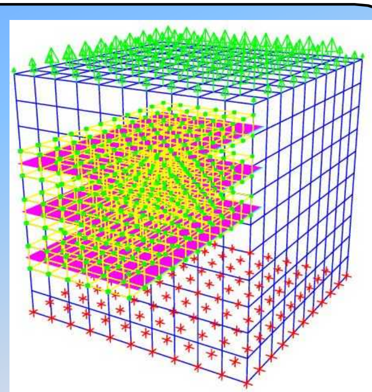
- Solve times comparable for Emin(6) and Emin(6,1)
- Emin(6,1) setup times *4x* better than Emin(6)

Sandia National Laboratories

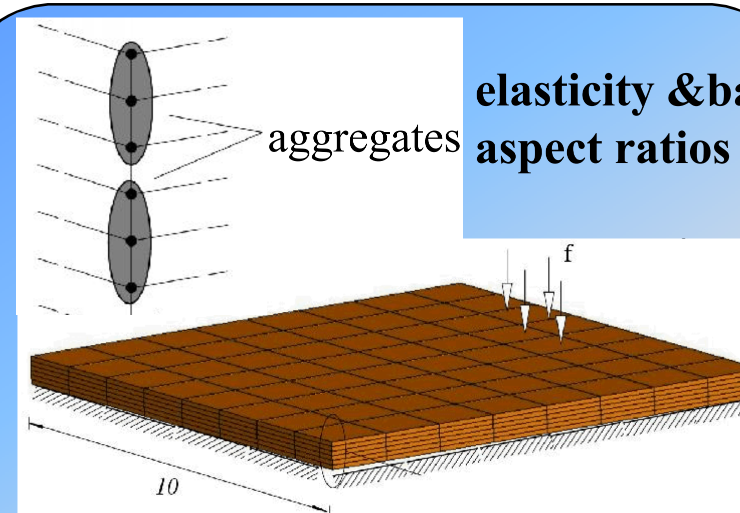
Exploiting Energy Minimization's Flexibility

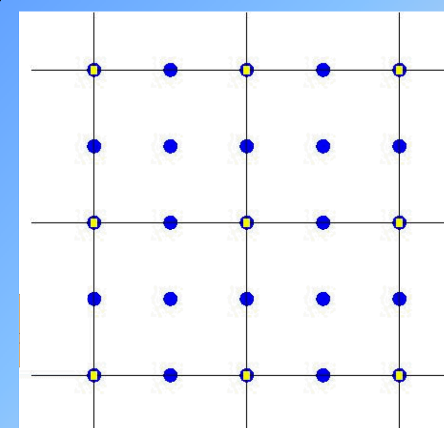
extended FE



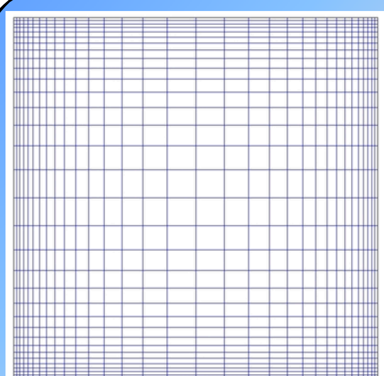


elasticity &bad aspect ratios

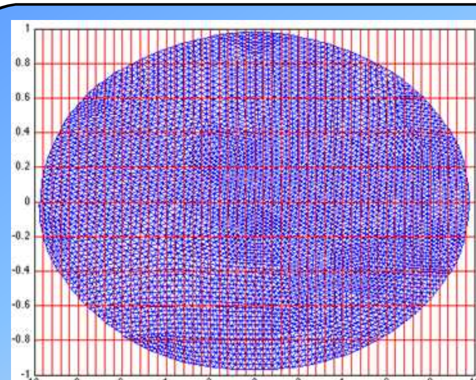




stretched meshes



unstructured to structured

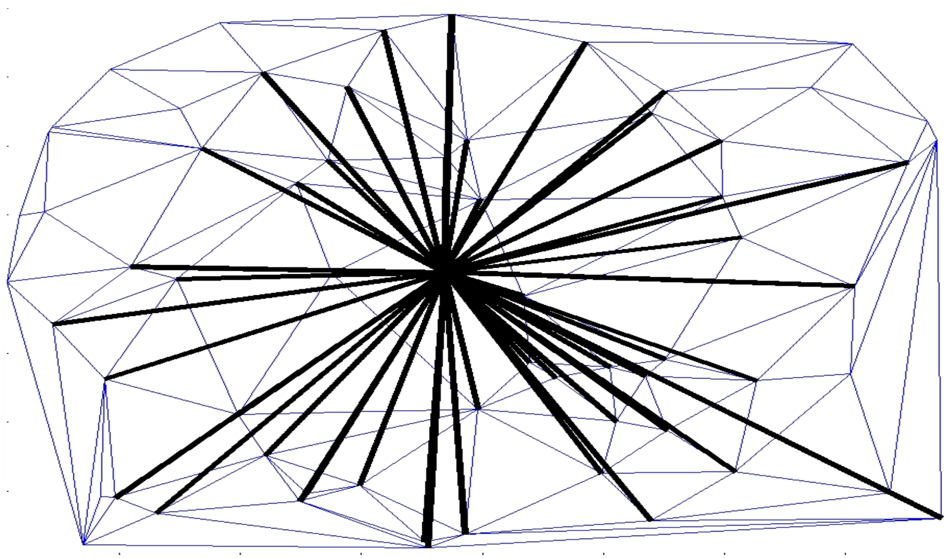


mixed finite elements

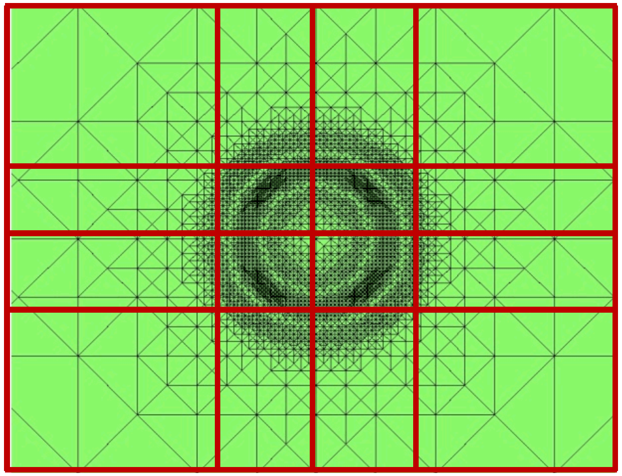
Unstructured ⇒ Structured

Why?

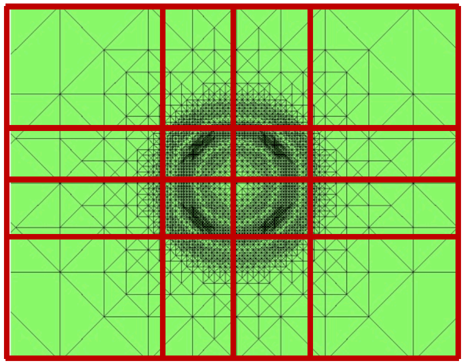
computational efficiency … on large parallel systems



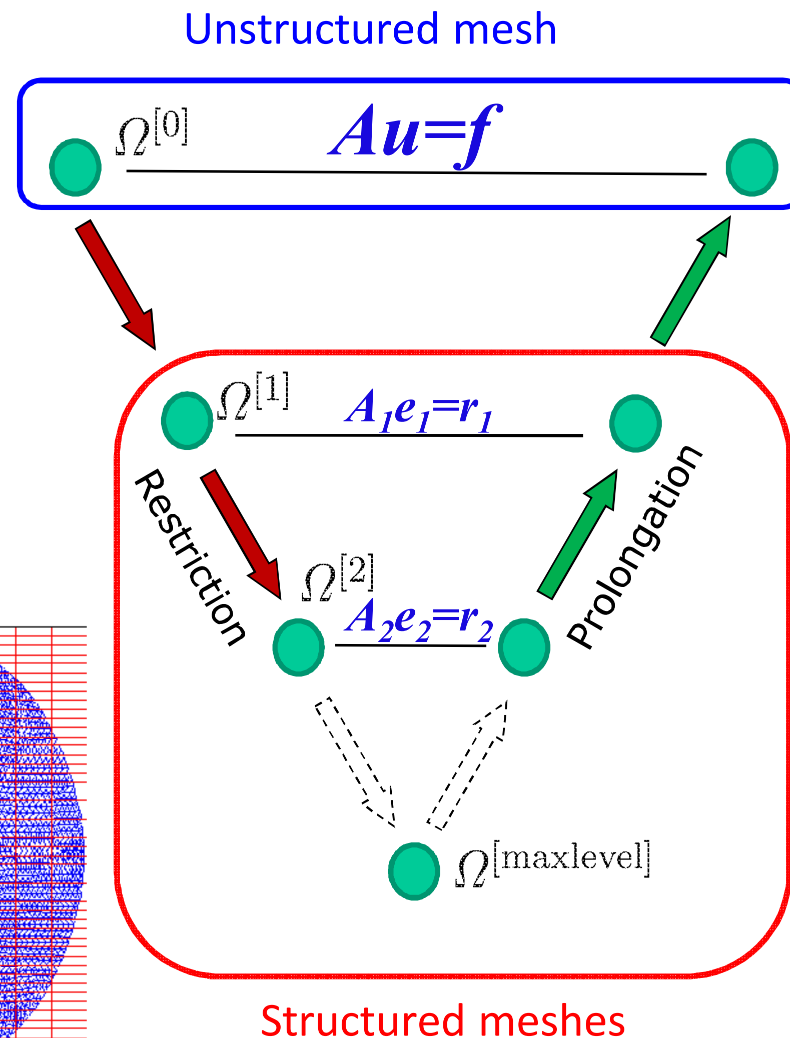
- **Overlay unstructured grid with structured grid.**
- **Coarse DOFs on structured mesh, interpolate from fine DOFs within rectangles.**
- **Interpolation weights found with energy minimization.**



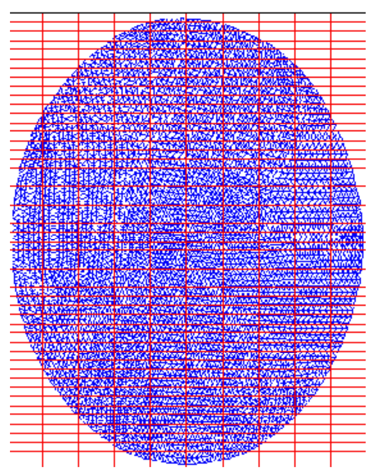
Unstructured ⇒ Structured



domains	*Emin its*
4 × 4	*12*
8× 8	*11*
12 × 12	*10*
16 × 16	*10*



#DOFs	Unstructured AMG		Unstruct /struct AMG
	SA	Emin	Emin
69185	77 (1.95)	59 (1.95)	31 (1.59)
277633	112 (1.93)	84 (1.93)	38 (1.61)



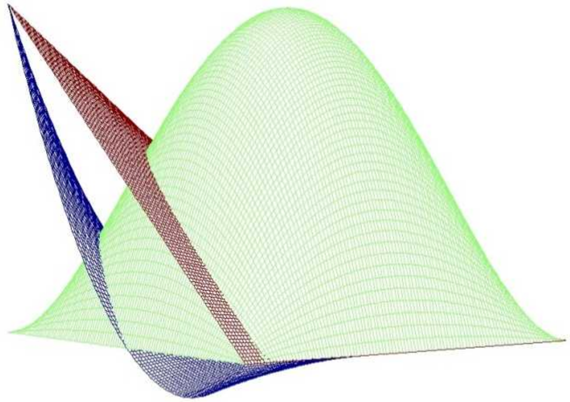
- caveat: unstr. is 3-level, unstr. to struct. is 2-level

Mixed Finite Elements

Stokes flow

$$\begin{pmatrix} \mu\Delta & \nabla \\ \nabla\bullet & \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}$$

Q2-Q1 elements due to Inf-Sup conditions



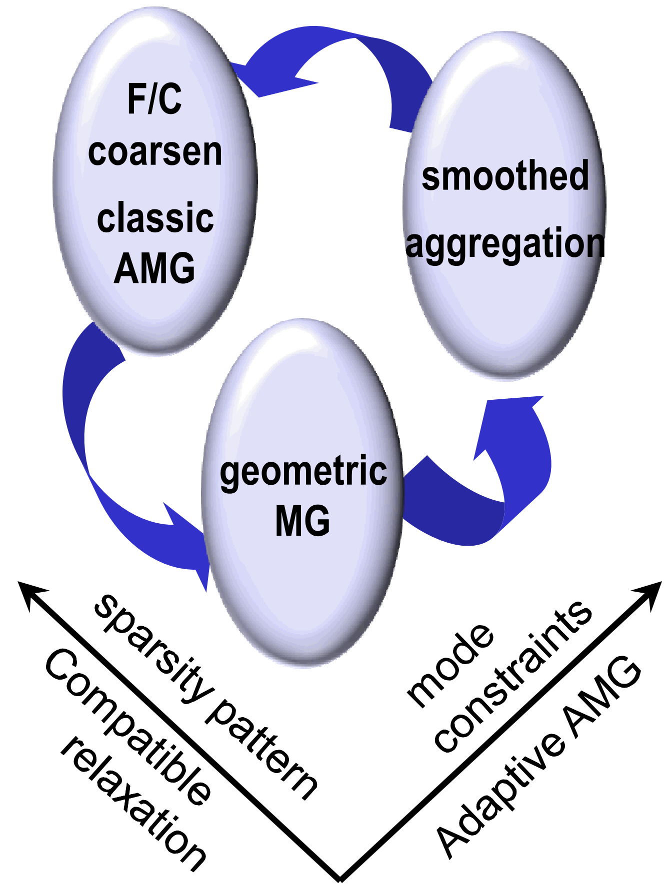
Mesh	Iters	complexity
9 x 9	20	2.01
17 x 17	25	1.81
65 x 65	Xx	Xx
257 x 257	29	1.91

Sandia National Laboratories

Concluding Remarks

- **Energy Minimizaiton AMG provides great flexibility**

- **Parallel practical variants definitely possible**
  - **a couple of Krylov sweeps**
  - **careful implementation of constraints**
  - **amortization**



- **Many situations arise where flexibility is useful**
  - **large scale parallel computations**
  - **advanced discretizations**
  - **stretched meshes**

- **Research needed to steer through some of the choices**

Related & Future Work

Helmholtz solvers

- Implemented parallel shifted-Laplacian AMG solver
    - tested on acoustic weapons problem
- Analyzed strengths/weaknesses of shifted-Laplacian
- Developed 2nd projection to accelerate shifted-Laplacian

Smoothers for MHD based on physics-based preconditioning

- Trilinos code under development
- Will be combined with Q2-Q1 AMG work

Coarsening/sparsity pattern scheme mimicking compatible relaxation idea

- Analyze intermediate *Emin* prolongators & correct deficiencies.

Sandia National Laboratories