μ-BFBT Preconditioner for Stokes Flow Problems with Highly Heterogeneous Viscosity

Johann Rudi[1] Georg Stadler[2] Omar Ghattas[1,3]

[1]Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin, USA

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

[3]Jackson School of Geosciences and Department of Mechanical Engineering, The University of Texas at Austin, USA

INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES

μ-BFBT: Key ideas and observations to be presented

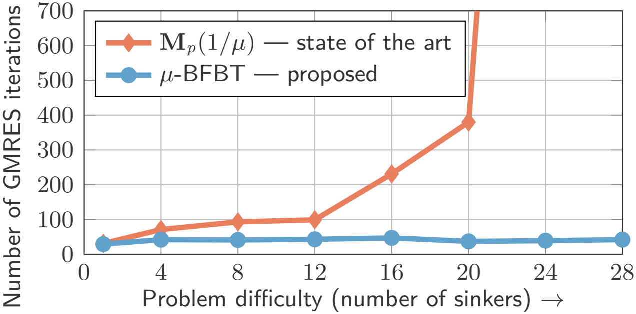
$$\underbrace{\begin{bmatrix} \mathbf{A}_\mu & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix}}_{\text{Stokes operator}} \underbrace{\begin{bmatrix} \tilde{\mathbf{A}}_\mu & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1}}_{\text{preconditioner}} \begin{bmatrix} \mathbf{u}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

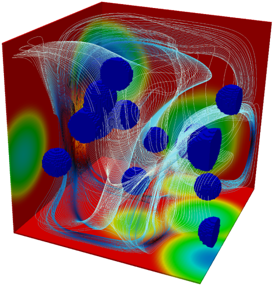
$$\tilde{\mathbf{A}}_\mu^{-1} \approx \mathbf{A}_\mu^{-1}$$

$$\tilde{\mathbf{S}}^{-1} \approx \mathbf{S}^{-1} := (\mathbf{B}\mathbf{A}_\mu^{-1}\mathbf{B}^\top)^{-1}$$

$$\tilde{\mathbf{S}}^{-1} = \tilde{\mathbf{M}}_p(1/\mu)^{-1} \quad \text{vs.}$$

$$\tilde{\mathbf{S}}^{-1} = (\mathbf{B}\mathbf{D}_\mu^{-1}\mathbf{B}^\top)^{-1}(\mathbf{B}\mathbf{D}_\mu^{-1}\mathbf{A}_\mu\mathbf{D}_\mu^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{D}_\mu^{-1}\mathbf{B}^\top)^{-1}, \quad \mathbf{D}_\mu = \tilde{\mathbf{M}}_{\boldsymbol{u}}(\sqrt{\mu})$$





Outline

Driving scientific problem & computational challenges

Class of benchmark problems

μ-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG+μ-BFBT

Parallel scalability for HMG+μ-BFBT

Incompressible Stokes flow with heterogeneous viscosity

Commonly occurring problem in CS&E:

Creeping non-Newtonian fluid modeled by incompressible Stokes equations with power-law rheology yields spatially-varying and highly heterogeneous viscosity μ after linearization.

Here, focus on preconditioning a linearized Stokes problem:

$$-\nabla \cdot \left[\mu(\boldsymbol{x})\left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\top\right)\right] + \nabla p = \boldsymbol{f} \qquad \text{heterogeneous viscosity } \mu$$

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

Discretization with inf-sub stable finite elements gives rise to the system:

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

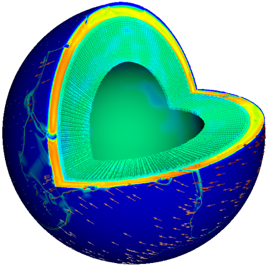
Iterative scheme with upper triangular block preconditioning:

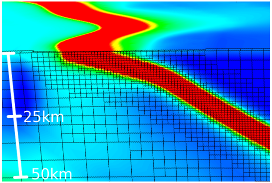
$$\begin{bmatrix} \mathbf{A}_\mu & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{A}}_\mu & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix} \qquad \begin{aligned} \tilde{\mathbf{A}}_\mu^{-1} &\approx \mathbf{A}_\mu^{-1} \\ \tilde{\mathbf{S}}^{-1} &\approx \mathbf{S}^{-1} := (\mathbf{B}\mathbf{A}_\mu^{-1}\mathbf{B}^\top)^{-1} \end{aligned}$$

Severe challenges for parallel scalable solvers

E.g., arising in Earth's mantle convection:

- Severe nonlinearity, heterogeneity, and anisotropy of the Earth's rheology
- Sharp viscosity gradients in narrow regions (6 orders of magnitude drop 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 mesh refinement is essential
- High-order finite elements with local mass conservation is crucial; yields a difficult to deal with discontinuous pressure approximation





Viscosity *(colors)*, surface velocity at sol. *(arrows)*, and locally refined mesh.

This talk's focus

Methods and preconditioners for the linearized Stokes problem:

- μ-BFBT inverse Schur complement approximation achieves robust convergence for Stokes problems with highly heterogeneous viscosity
- HMG: hybrid spectral-geometric-algebraic multigrid exhibits extreme parallel scalability & (nearly) optimal algorithmic scalability, used for preconditioning viscous block $\tilde{\mathbf{A}}_{\mu}^{-1}$ and inside μ-BFBT via V-cycles
- Inf-sup stable velocity-pressure discretization $\mathbb{Q}_k \times \mathbb{P}_{k-1}^{\mathrm{disc}}$, order $k \geq 2$
- Mass conservation at element level via discontinuous, modal pressure

Simplifications are made for the sake of clear analysis and wide applicability, but solver development targets Earth's M.C. as application

- Simple viscosity formulation vs. complicated nonlinear Earth rheology
- Undeformed cube domain vs. spherical shell
- Uniformly refined mesh vs. aggressively locally refined

Outline

Class of multi-sinker benchmark problems

Vary 2 viscosity parameters to test robustness:

- Local param.: #sinkers n at random points $\boldsymbol{c}_i$
- Global param.: $\mathrm{DR}(\mu) := \max(\mu)/\min(\mu)$

$$\mu(\boldsymbol{x}) := (\mu_{\max} - \mu_{\min})(1 - \chi_n(\boldsymbol{x})) + \mu_{\min}$$

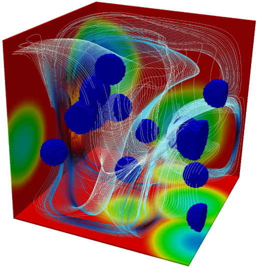
$$\mu_{\min} := \mathrm{DR}(\mu)^{-\frac{1}{2}}, \quad \mu_{\max} := \mathrm{DR}(\mu)^{\frac{1}{2}}$$

$$\chi_n(\boldsymbol{x}) := \prod_{i=1}^{n} 1 - \exp\left[-d \max\left(0, |\boldsymbol{c}_i - \boldsymbol{x}| - \frac{w}{2}\right)^2\right]$$

$$\boldsymbol{f}(\boldsymbol{x}) := b(1 - \chi_n(\boldsymbol{x})), \quad \text{(where } b, d, w \text{ const.)}$$

Vary 2 discretization parameters to test algorithmic scalability:

- Finite element order k (recall: $\mathbb{Q}_k \times \mathbb{P}_{k-1}^{\mathrm{disc}}$)
- Mesh refinement level ℓ



Viscosity *(colors)* with highest value *(blue)* assumed inside sinkers, and streamlines of nonlocal velocity field.

Outline

Propose: μ-BFBT inverse Schur complement approx.

$$\begin{bmatrix} \mathbf{A}_\mu & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{A}}_\mu & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}$$

$$\tilde{\mathbf{A}}_\mu^{-1} \approx \mathbf{A}_\mu^{-1}$$

$$\tilde{\mathbf{S}}^{-1} \approx \mathbf{S}^{-1} := (\mathbf{B}\mathbf{A}_\mu^{-1}\mathbf{B}^\top)^{-1}$$

Propose: μ-BFBT inverse Schur complement approx.

$$\begin{bmatrix} \mathbf{A}_\mu & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{A}}_\mu & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix} \qquad \begin{aligned} &\tilde{\mathbf{A}}_\mu^{-1} \approx \mathbf{A}_\mu^{-1} \\ &\tilde{\mathbf{S}}^{-1} \approx \mathbf{S}^{-1} := (\mathbf{B}\mathbf{A}_\mu^{-1}\mathbf{B}^\top)^{-1} \end{aligned}$$

Underlying principle of BFBT / Least Squares Commutators (LSC): find a commutator matrix $\mathbf{X}$ s.t. (denote unit vectors by $\mathbf{e}_j$)

$$\mathbf{A}_\mu\mathbf{D}^{-1}\mathbf{B}^\top - \mathbf{B}^\top\mathbf{X} \approx \mathbf{0} \quad \text{or} \quad \min_{\mathbf{X}} \left\| \mathbf{A}_\mu\mathbf{D}^{-1}\mathbf{B}^\top\mathbf{e}_j - \mathbf{B}^\top\mathbf{X}\mathbf{e}_j \right\|^2_{\mathbf{C}^{-1}} \quad \forall j$$

$$\Rightarrow \quad \tilde{\mathbf{S}}^{-1}_{\text{BFBT}} := \left(\mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top\right)^{-1} \left(\mathbf{B}\mathbf{C}^{-1}\mathbf{A}_\mu\mathbf{D}^{-1}\mathbf{B}^\top\right) \left(\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^\top\right)^{-1}.$$

Choice of matrices $\mathbf{C}, \mathbf{D}$ is critical for convergence and robustness.

Propose: μ-BFBT inverse Schur complement approx.

$$\begin{bmatrix} \mathbf{A}_\mu & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{A}}_\mu & \mathbf{B}^\top \\ \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix} \qquad \begin{aligned} &\tilde{\mathbf{A}}_\mu^{-1} \approx \mathbf{A}_\mu^{-1} \\ &\tilde{\mathbf{S}}^{-1} \approx \mathbf{S}^{-1} := (\mathbf{B}\mathbf{A}_\mu^{-1}\mathbf{B}^\top)^{-1} \end{aligned}$$

Underlying principle of BFBT / Least Squares Commutators (LSC): find a commutator matrix $\mathbf{X}$ s.t. (denote unit vectors by $\mathbf{e}_j$)

$$\mathbf{A}_\mu \mathbf{D}^{-1}\mathbf{B}^\top - \mathbf{B}^\top \mathbf{X} \approx \mathbf{0} \quad \text{or} \quad \min_{\mathbf{X}} \left\| \mathbf{A}_\mu \mathbf{D}^{-1}\mathbf{B}^\top \mathbf{e}_j - \mathbf{B}^\top \mathbf{X} \mathbf{e}_j \right\|^2_{\mathbf{C}^{-1}} \quad \forall j$$

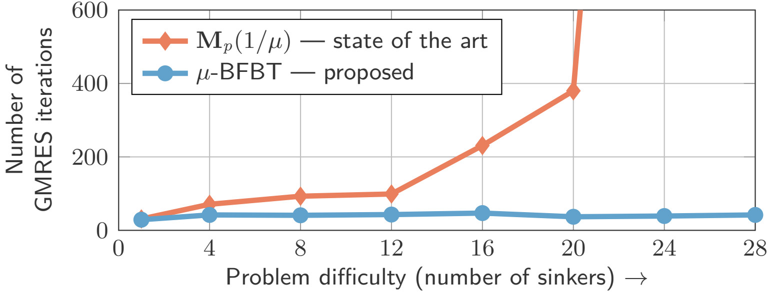
$$\Rightarrow \quad \tilde{\mathbf{S}}^{-1}_{\mathsf{BFBT}} := \left(\mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top\right)^{-1} \left(\mathbf{B}\mathbf{C}^{-1}\mathbf{A}_\mu\mathbf{D}^{-1}\mathbf{B}^\top\right) \left(\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^\top\right)^{-1}.$$

Choice of matrices $\mathbf{C}, \mathbf{D}$ is critical for convergence and robustness.

$$\tilde{\mathbf{S}}^{-1}_{\mu\text{-}\mathsf{BFBT}} := \left(\mathbf{B}\mathbf{C}_\mu^{-1}\mathbf{B}^\top\right)^{-1} \left(\mathbf{B}\mathbf{C}_\mu^{-1}\mathbf{A}_\mu\mathbf{D}_\mu^{-1}\mathbf{B}^\top\right) \left(\mathbf{B}\mathbf{D}_\mu^{-1}\mathbf{B}^\top\right)^{-1}$$

where $\mathbf{C}_\mu = \mathbf{D}_\mu := \tilde{\mathbf{M}}_{\boldsymbol{u}}(\sqrt{\mu})$ are responsible for efficacy and robustness.

Robustness of μ-BFBT over established state of the art



$\mathbf{M}_p(1/\mu)$ $(k=2, \ell=7)$

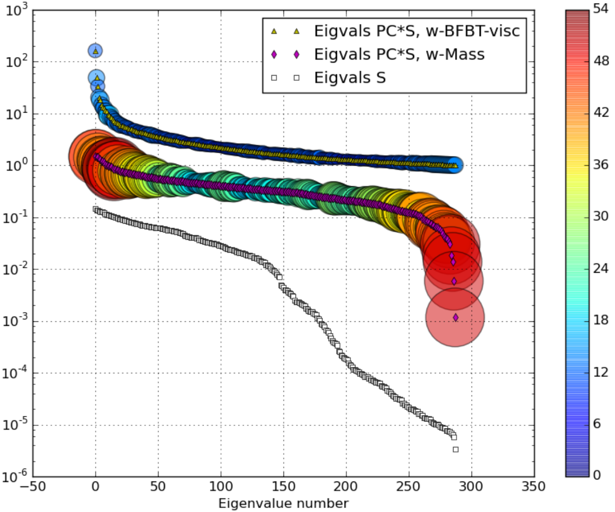
$\mathrm{DR}(\mu) = \ldots$	10^4	10^6	10^8	10^{10}
S1-rand	29	31	31	29
S8-rand	64	79	93	165
S16-rand	85	167	231	891
S24-rand	117	286	3279	5983
S28-rand	108	499	2472	>10000

μ-BFBT $(k=2, \ell=7)$

$\mathrm{DR}(\mu) = \ldots$	10^4	10^6	10^8	10^{10}
S1-rand	29	29	29	30
S8-rand	38	40	41	44
S16-rand	40	45	47	48
S24-rand	31	32	39	55
S28-rand	29	31	42	60

Eigenvalue/-vector analysis for system $\mathbf{Sp} = \mathbf{g}$ in 2D

Spectrum of exact and preconditioned Schur complement *(markers)*,
$\#$GMRES iter. with eigenvector components of rel. residual $> 10^{-2}$ *(circles/colors)*



$\#$sinkers $= 4$,
$\mathrm{DR}(\mu) = 10^4$,
$k = 2, \ell = 4$

Outline

Modifications for Dirichlet boundary conditions

Consider $\Omega = \mathbb{R}^3$, $\mu \equiv 1$, then the discrete commutator

$$\mathbf{A}\mathbf{M}_{\boldsymbol{u}}^{-1}\mathbf{B}^\top - \mathbf{B}^\top\mathbf{X}$$

vanishes in infinite dimensions:

$$0 = (\nabla \cdot \nabla)\nabla - \nabla(\nabla \cdot \nabla) =: A_{\boldsymbol{u}}B^* - B^*A_p$$

However, if Ω is bounded and Dirichlet BC's are enforced on $\partial\Omega$, then in general

$$A_{\boldsymbol{u}}B^* - B^*A_p \neq 0 \quad \text{on } \partial\Omega$$

This poses a problem for algorithmic scalability, i.e., maintained convergence rate for increasing k and ℓ; similar observations are made in [Elman, Tuminaro, 2009] for Navier-Stokes equations.

Modifications for Dirichlet boundary conditions

Recall: $\tilde{\mathbf{S}}^{-1}_{\mu\text{-BFBT}} = \left(\mathbf{B}\mathbf{C}_\mu^{-1}\mathbf{B}^\top\right)^{-1}\left(\mathbf{B}\mathbf{C}_\mu^{-1}\mathbf{A}_\mu\mathbf{D}_\mu^{-1}\mathbf{B}^\top\right)\left(\mathbf{B}\mathbf{D}_\mu^{-1}\mathbf{B}^\top\right)^{-1}$

$$w_{\mu,a}(\boldsymbol{x}) := \begin{cases} a\sqrt{\mu(\boldsymbol{x})} & \boldsymbol{x} \in \Omega_D, \\ \sqrt{\mu(\boldsymbol{x})} & \boldsymbol{x} \notin \Omega_D, \end{cases} \qquad \Omega_D = \text{elems. touching Dirichlet bdr.}$$

Choose $a_C \geq 1$ in $\mathbf{C}_\mu^{-1} = \tilde{\mathbf{M}}_{\boldsymbol{u}}(w_{\mu,a_C})^{-1}$, $a_D \geq 1$ in $\mathbf{D}_\mu^{-1} = \tilde{\mathbf{M}}_{\boldsymbol{u}}(w_{\mu,a_D})^{-1}$

Interpretation: Reduce weight of Ω_D in commutator relationship.

$k = 2,\ \ell = 5$

$a_C \setminus a_D$	1	2	4	8	16	32
1	33	33	34	34	34	35
2	33	33	34	34	34	34
4	33	34	34	36	38	39
8	34	34	36	39	43	44
16	34	34	38	43	46	49
32	34	34	39	44	49	53

$k = 2,\ \ell = 7$

$a_C \setminus a_D$	1	2	4	8	16	32
1	45	37	34	34	34	34
2	37	36	35	36	36	36
4	34	36	38	39	40	41
8	34	36	39	42	44	44
16	34	36	40	44	45	46
32	34	36	41	44	46	47

Modifications for Dirichlet boundary conditions

Recall: $\tilde{\mathbf{S}}^{-1}_{\mu\text{-BFBT}} = \left(\mathbf{B}\mathbf{C}_\mu^{-1}\mathbf{B}^\top\right)^{-1} \left(\mathbf{B}\mathbf{C}_\mu^{-1}\mathbf{A}_\mu\mathbf{D}_\mu^{-1}\mathbf{B}^\top\right) \left(\mathbf{B}\mathbf{D}_\mu^{-1}\mathbf{B}^\top\right)^{-1}$

$$w_{\mu,a}(\boldsymbol{x}) := \begin{cases} a\sqrt{\mu(\boldsymbol{x})} & \boldsymbol{x} \in \Omega_D, \\ \sqrt{\mu(\boldsymbol{x})} & \boldsymbol{x} \notin \Omega_D, \end{cases} \qquad \Omega_D = \text{elems. touching Dirichlet bdr.}$$

Choose $a_C \geq 1$ in $\mathbf{C}_\mu^{-1} = \tilde{\mathbf{M}}_{\boldsymbol{u}}(w_{\mu,a_C})^{-1}$, $a_D \geq 1$ in $\mathbf{D}_\mu^{-1} = \tilde{\mathbf{M}}_{\boldsymbol{u}}(w_{\mu,a_D})^{-1}$

Interpretation: Reduce weight of Ω_D in commutator relationship.

$k = 2,\ \ell = 5$

$a_C \setminus a_D$	1	2	4	8	16	32
1	33	33	34	34	34	35
2	33	33	34	34	34	34
4	33	34	34	36	38	39
8	34	34	36	39	43	44
16	34	34	38	43	46	49
32	34	34	39	44	49	53

$k = 5,\ \ell = 5$

$a_C \setminus a_D$	1	2	4	8	16	32
1	63	53	46	43	43	44
2	53	51	51	51	52	53
4	47	51	55	59	62	64
8	44	51	59	65	69	72
16	43	52	62	69	75	78
32	44	53	64	72	78	82

Outline

Driving scientific problem & computational challenges

Class of benchmark problems

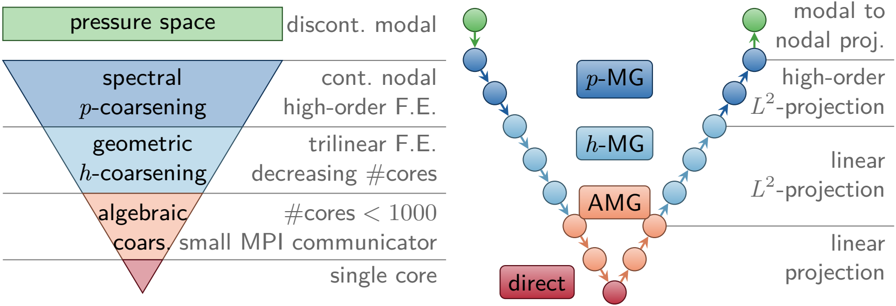
μ-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG+μ-BFBT

Parallel scalability for HMG+μ-BFBT

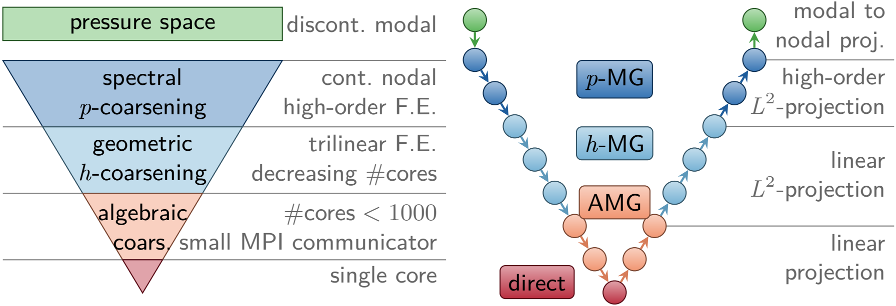
Algorithmic scalability for HMG+μ-BFBT



HMG: hybrid spectral-geometric-algebraic multigrid

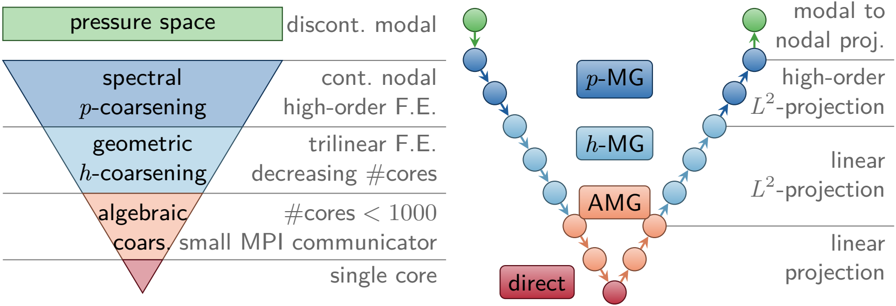
- Parallel repartitioning of coarser meshes for load-balancing (crucial for AMR); sufficiently coarse meshes occupy only subsets of cores
- High-order L^2-projection onto coarser levels; restriction & interpolation are adjoints of each other in L^2-sense
- Chebyshev accelerated Jacobi smoother (Cheb. from PETSc) with tensorized matrix-free high-order stiffness apply; assembly of high-order diagonal only

Algorithmic scalability for HMG+μ-BFBT



ℓ	a_D	$\boldsymbol{u}$-DOF $[\times 10^6]$	It. $\mathbf{A}_\mu$	p-DOF $[\times 10^6]$	It. $\mathbf{K}_\mathrm{d}$	DOF $[\times 10^6]$	It. **Stokes**
4	1	0.11	18	0.02	8	0.12	40
5	2	0.82	18	0.13	7	0.95	33
6	4	6.44	18	1.05	6	7.49	33
7	8	50.92	18	8.39	6	59.31	34
8	16	405.02	18	67.11	6	472.12	34
9	32	3230.67	18	536.87	6	3767.54	34
10	64	25807.57	18	4294.97	6	30102.53	34

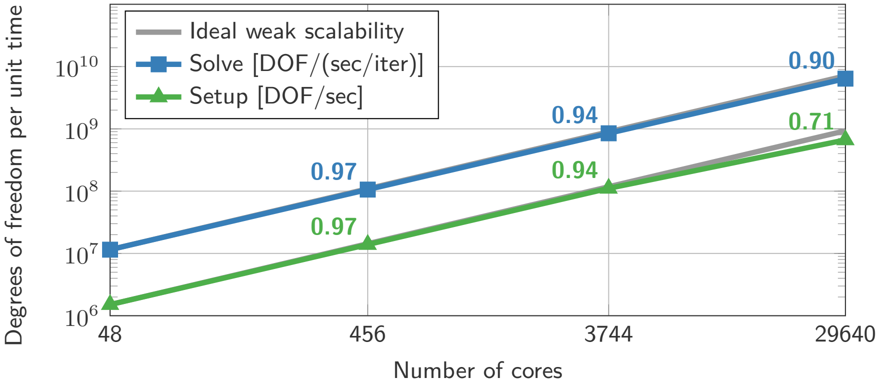
Algorithmic scalability for HMG+μ-BFBT



k	a_D	$\boldsymbol{u}$-DOF $[\times 10^6]$	It. $\mathbf{A}_\mu$	p-DOF $[\times 10^6]$	It. $\mathbf{K}_\mathrm{d}$	DOF $[\times 10^6]$	It. **Stokes**
2	2	0.82	18	0.13	7	0.95	33
3	4	2.74	20	0.32	8	3.07	37
4	8	6.44	20	0.66	7	7.10	36
5	16	12.52	23	1.15	12	13.67	43
6	32	21.56	23	1.84	12	23.40	50
7	64	34.17	22	2.75	10	36.92	54
8	128	50.92	22	3.93	10	54.86	67

Outline

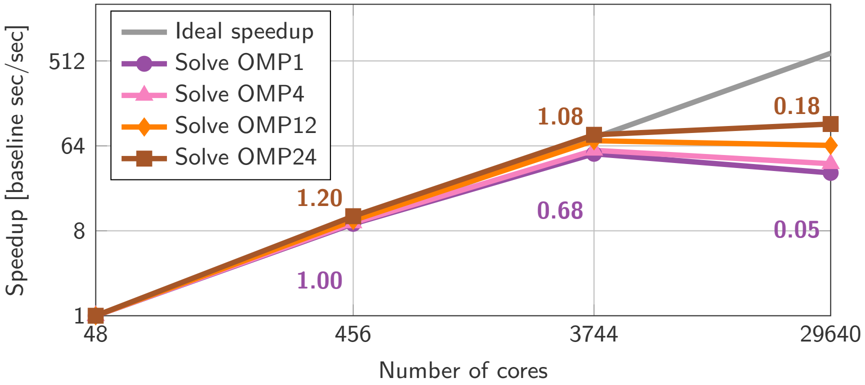
Weak scalability for HMG+μ-BFBT



Performed on TACC's Lonestar 5: Cray XC40 with 1252 compute nodes, each has 2 Intel Haswell 12-core processors and 64 GBytes of memory.

Extreme scalability for Earth's M.C. on up to 1.6 million cores of IBM's BG/Q: 97 % weak efficiency [SC'15 Gordon Bell paper: Rudi, Malossi, Isaac et al., 2015]

Strong scalability for HMG+μ-BFBT



Performed on TACC's Lonestar 5: Cray XC40 with 1252 compute nodes, each has 2 Intel Haswell 12-core processors and 64 GBytes of memory.

Extreme scalability for Earth's M.C. on up to 1.6 million cores of IBM's BG/Q: 32 % strong efficiency [SC'15 Gordon Bell paper: Rudi, Malossi, Isaac et al., 2015]

References

Viscosity-weighted pressure mass matrix for Stokes:

- Burstedde, Ghattas, Stadler, Tu, and Wilcox, Comput. Method. Appl. M., 198 (2009).
- Grinevich and Olshanskii, SIAM J. Sci. Comput., 31 (2009).

BFBT for Navier-Stokes:

- Elman, SIAM J. Sci. Comput., 20 (1999).
- Silvester, Elman, Kay, and Wathen, J. Comput. Appl. Math., 128 (2001).
- Kay, Loghin, and Wathen, SIAM J. Sci. Comput., 24 (2002).
- Elman, Howle, Shadid, Shuttleworth, and Tuminaro, SIAM J. Sci. Comput., 27 (2006).
- Elman and Tuminaro, Trans. Numer. Anal., 35 (2009).

BFBT for Stokes:

- May and Moresi, Phys. Earth Planet. In., 171 (2008).
- Rudi, Malossi, Isaac, Stadler, Gurnis, Ineichen, Bekas, Curioni, and Ghattas, Proceedings of SC15 (2015), Winner of Gordon Bell Prize.