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

Johann Rudi\textsuperscript{1}  Georg Stadler\textsuperscript{2}  Omar Ghattas\textsuperscript{1,3}

\textsuperscript{1}Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin, USA

\textsuperscript{2}Courant Institute of Mathematical Sciences, New York University, USA

\textsuperscript{3}Jackson School of Geosciences and Department of Mechanical Engineering, The University of Texas at Austin, USA
\textbf{\(\mu\)-BFBT: Key ideas and observations to be presented}

\[
\begin{bmatrix}
A_\mu & B^\top \\
B & 0
\end{bmatrix}
\begin{bmatrix}
A_\mu & B^\top \\
0 & \tilde{S}
\end{bmatrix}^{-1}
\begin{bmatrix}
u \\ p
\end{bmatrix} =
\begin{bmatrix}
f \\ 0
\end{bmatrix}
\]

\(\tilde{A}_\mu^{-1} \approx A_\mu^{-1}\)

\(\tilde{S}^{-1} \approx S^{-1} := (BA_\mu^{-1}B^\top)^{-1}\)

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

\(\tilde{S}^{-1} = (BD_\mu^{-1}B^\top)^{-1}(BD_\mu^{-1}A_\mu D_\mu^{-1}B^\top)(BD_\mu^{-1}B^\top)^{-1},\)

\(D_\mu = \tilde{M}_u(\sqrt{\mu})\)

---

![Graph showing the number of GMRES iterations vs. problem difficulty (number of sinkers)]
Outline

Driving scientific problem & computational challenges

Class of benchmark problems

$\mu$-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG+$\mu$-BFBT

Parallel scalability for HMG+$\mu$-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 $\mu$ after linearization.

Here, focus on preconditioning a linearized Stokes problem:

$$
- \nabla \cdot [\mu(x) (\nabla u + \nabla u^\top)] + \nabla p = f \\
- \nabla \cdot u = 0 \\
$$

sought: velocity $u$, pressure $p$

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

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

Iterative scheme with upper triangular block preconditioning:

$$
\begin{bmatrix}
A_{\mu} & B^\top \\
B & 0
\end{bmatrix}
\begin{bmatrix}
\tilde{A}_{\mu} & B^\top \\
0 & \tilde{S}
\end{bmatrix}^{-1}
\begin{bmatrix}
u \\
p
\end{bmatrix}
= 
\begin{bmatrix}
f \\
0
\end{bmatrix}
\tilde{A}_{\mu}^{-1} \approx A_{\mu}^{-1} \\
\tilde{S}^{-1} \approx S^{-1} := (BA_{\mu}^{-1}B^\top)^{-1}
$$
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.
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 $\tilde{A}_\mu^{-1}$ and inside $\mu$-BFBT via V-cycles.
- Inf-sup stable velocity-pressure discretization $\mathbb{Q}_k \times \mathbb{P}^{\text{disc}}_k$, 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

Driving scientific problem & computational challenges

Class of benchmark problems

\(\mu\)-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG + \(\mu\)-BFBT

Parallel scalability for HMG + \(\mu\)-BFBT
Class of multi-sinker benchmark problems

Vary 2 viscosity parameters to test robustness:

- Local param.: #sinkers \( n \) at random points \( c_i \)
- Global param.: \( \text{DR}(\mu) := \max(\mu)/\min(\mu) \)

\[
\mu(x) := (\mu_{\text{max}} - \mu_{\text{min}})(1 - \chi_n(x)) + \mu_{\text{min}}
\]

\[
\mu_{\text{min}} := \text{DR}(\mu)^{-\frac{1}{2}}, \quad \mu_{\text{max}} := \text{DR}(\mu)^{\frac{1}{2}}
\]

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

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

Vary 2 discretization parameters to test algorithmic scalability:

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

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

Driving scientific problem & computational challenges

Class of benchmark problems

$\mu$-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG+$\mu$-BFBT

Parallel scalability for HMG+$\mu$-BFBT
Propose: $\mu$-BFBT inverse Schur complement approx.

\[
\begin{bmatrix}
A_\mu & B^T \\
B & 0
\end{bmatrix}
\begin{bmatrix}
\tilde{A}_\mu & B^T \\
0 & \tilde{S}
\end{bmatrix}^{-1}
\begin{bmatrix}
u \\
p
\end{bmatrix} =
\begin{bmatrix}
f \\
0
\end{bmatrix}
\]

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

$\tilde{S}^{-1} \approx S^{-1} := (BA_\mu^{-1}B^T)^{-1}$
Propose: $\mu$-BFBT inverse Schur complement approx.

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

Underlying principle of BFBT / Least Squares Commutators (LSC):
find a commutator matrix $X$ s.t. (denote unit vectors by $e_j$)

$$A_\mu D^{-1}B^\top - B^\top X \approx 0 \quad \text{or} \quad \min_X \left\| A_\mu D^{-1}B^\top e_j - B^\top X e_j \right\|_{C^{-1}}^2 \quad \forall j$$

$$\Rightarrow \quad \tilde{S}_{\text{BFBT}}^{-1} := \left(BC^{-1}B^\top \right)^{-1} \left(BC^{-1}A_\mu D^{-1}B^\top \right) \left(BD^{-1}B^\top \right)^{-1}$$

Choice of matrices $C, D$ is critical for convergence and robustness.
Propose: $\mu$-BFBT inverse Schur complement approx.

\[
\begin{bmatrix}
A_\mu & B^T \\
B & 0
\end{bmatrix}
\begin{bmatrix}
\tilde{A}_\mu & B^T \\
0 & \tilde{S}
\end{bmatrix}^{-1}
\begin{bmatrix}
u \\
p
\end{bmatrix} =
\begin{bmatrix}
f \\
0
\end{bmatrix}
\tilde{A}_\mu^{-1} \approx A_\mu^{-1}
\tilde{S}^{-1} \approx S^{-1} := (B A_\mu^{-1} B^T)^{-1}
\]

Underlying principle of BFBT / Least Squares Commutators (LSC): find a commutator matrix $X$ s.t. (denote unit vectors by $e_j$)

\[
A_\mu D^{-1} B^T - B^T X \approx 0 \quad \text{or} \quad \min_X \| A_\mu D^{-1} B^T e_j - B^T X e_j \|_C^{-1} \quad \forall j
\]

\[
\Rightarrow \tilde{S}^{-1}_{\text{BFBT}} := \left( B C^{-1} B^T \right)^{-1} \left( B C^{-1} A_\mu D^{-1} B^T \right) \left( B D^{-1} B^T \right)^{-1}
\]

Choice of matrices $C, D$ is critical for convergence and robustness.

\[
\tilde{S}^{-1}_{\mu-\text{BFBT}} := \left( B C_\mu^{-1} B^T \right)^{-1} \left( B C_\mu^{-1} A_\mu D_\mu^{-1} B^T \right) \left( B D_\mu^{-1} B^T \right)^{-1}
\]

where $C_\mu = D_\mu := \tilde{M}_u(\sqrt{\mu})$ are responsible for efficacy and robustness.
Robustness of $\mu$-BFBT over established state of the art

![Graph showing comparison between $M_p(1/\mu)$ and $\mu$-BFBT for varying problem difficulty (number of sinkers)].

<table>
<thead>
<tr>
<th>Problem difficulty (number of sinkers)</th>
<th>$M_p(1/\mu)$ ($k = 2, \ell = 7$)</th>
<th>$\mu$-BFBT ($k = 2, \ell = 7$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-rand</td>
<td>$10^4$ 29 31 31 29</td>
<td>S1-rand 29 29 29 29 30</td>
</tr>
<tr>
<td>S8-rand</td>
<td>$10^6$ 64 79 93 165</td>
<td>S8-rand 38 40 41 44 44</td>
</tr>
<tr>
<td>S16-rand</td>
<td>$10^8$ 85 167 231 891</td>
<td>S16-rand 40 45 47 48 48</td>
</tr>
<tr>
<td>S24-rand</td>
<td>$10^{10}$ 117 286 3279 5983</td>
<td>S24-rand 31 32 39 55 55</td>
</tr>
<tr>
<td>S28-rand</td>
<td>$10^{10}$ 108 499 2472 &gt;10000</td>
<td>S28-rand 29 31 42 60 60</td>
</tr>
</tbody>
</table>
Eigenvalue/-vector analysis for system $Sp = g$ in 2D

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

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

Driving scientific problem & computational challenges

Class of benchmark problems

$\mu$-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG+$\mu$-BFBT

Parallel scalability for HMG+$\mu$-BFBT
Modifications for Dirichlet boundary conditions

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

$$AM_u^{-1}B^\top - B^\top X$$

vanishes in infinite dimensions:

$$0 = (\nabla \cdot \nabla)\nabla - \nabla(\nabla \cdot \nabla) =: A_uB^* - B^*A_p$$

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

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

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

Recall: \( \tilde{S}^{-1}_{\mu-\text{BFBT}} = \left( BC^{-1}_\mu B^T \right)^{-1} \left( BC^{-1}_\mu A_\mu D^{-1}_\mu B^T \right) \left( BD^{-1}_\mu B^T \right)^{-1} \)

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

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

Interpretation: Reduce weight of \( \Omega_D \) in commutator relationship.

\[
\begin{array}{ccccccc}
\hline
a_C \setminus a_D & 1 & 2 & 4 & 8 & 16 & 32 \\
\hline
1 & 33 & 33 & 34 & 34 & 34 & 35 \\
2 & 33 & 33 & 34 & 34 & 34 & 34 \\
4 & 33 & 34 & 34 & 36 & 38 & 39 \\
8 & 34 & 34 & 36 & 39 & 43 & 44 \\
16 & 34 & 34 & 38 & 43 & 46 & 49 \\
32 & 34 & 34 & 39 & 44 & 49 & 53 \\
\hline
\end{array}
\]

\[
\begin{array}{ccccccc}
\hline
a_C \setminus a_D & 1 & 2 & 4 & 8 & 16 & 32 \\
\hline
1 & 45 & 37 & 34 & 34 & 34 & 34 \\
2 & 37 & 36 & 35 & 36 & 36 & 36 \\
4 & 34 & 36 & 38 & 39 & 40 & 41 \\
8 & 34 & 36 & 39 & 42 & 44 & 44 \\
16 & 34 & 36 & 40 & 44 & 45 & 46 \\
32 & 34 & 36 & 41 & 44 & 46 & 47 \\
\hline
\end{array}
\]
"µ-BFBT Preconditioner for Stokes Flow Problems" by Johann Rudi

Modifications for Dirichlet boundary conditions

Recall: \( \tilde{S}_\mu^{-1} - BFBT = \left( BC_\mu^{-1} B^T \right)^{-1} \left( BC_\mu^{-1} A_\mu D_\mu^{-1} B^T \right) \left( BD_\mu^{-1} B^T \right)^{-1} \)

\[
w_{\mu,a}(x) := \begin{cases} \frac{a \sqrt{\mu(x)}}{\sqrt{\mu(x)}} & x \in \Omega_D, \\ \frac{\sqrt{\mu(x)}}{\mu(x)} & x \notin \Omega_D, \end{cases}
\]

\( \Omega_D = \) elems. touching Dirichlet bdr.

Choose \( a_C \geq 1 \) in \( C_\mu^{-1} = \tilde{M}_u(w_{\mu,a_C})^{-1} \), \( a_D \geq 1 \) in \( D_\mu^{-1} = \tilde{M}_u(w_{\mu,a_D})^{-1} \)

Interpretation: Reduce weight of \( \Omega_D \) in commutator relationship.

<table>
<thead>
<tr>
<th>( a_C \setminus a_D )</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33</td>
<td>33</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>35</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>33</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>4</td>
<td>33</td>
<td>34</td>
<td>34</td>
<td>36</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>34</td>
<td>34</td>
<td>36</td>
<td>39</td>
<td>43</td>
<td>44</td>
</tr>
<tr>
<td>16</td>
<td>34</td>
<td>34</td>
<td>38</td>
<td>43</td>
<td>46</td>
<td>49</td>
</tr>
<tr>
<td>32</td>
<td>34</td>
<td>34</td>
<td>39</td>
<td>44</td>
<td>49</td>
<td>53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( a_C \setminus a_D )</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>63</td>
<td>53</td>
<td>46</td>
<td>43</td>
<td>43</td>
<td>44</td>
</tr>
<tr>
<td>2</td>
<td>53</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>52</td>
<td>53</td>
</tr>
<tr>
<td>4</td>
<td>47</td>
<td>51</td>
<td>55</td>
<td>59</td>
<td>62</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>44</td>
<td>51</td>
<td>65</td>
<td>69</td>
<td>75</td>
<td>78</td>
</tr>
<tr>
<td>16</td>
<td>43</td>
<td>52</td>
<td>62</td>
<td>69</td>
<td>75</td>
<td>78</td>
</tr>
<tr>
<td>32</td>
<td>44</td>
<td>53</td>
<td>64</td>
<td>72</td>
<td>78</td>
<td>82</td>
</tr>
</tbody>
</table>
Outline

Driving scientific problem & computational challenges

Class of benchmark problems

$\mu$-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG + $\mu$-BFBT

Parallel scalability for HMG + $\mu$-BFBT
Algorithmic scalability for HMG + $\mu$-BFBT

**pressure space**
- **discont. modal**
- **cont. nodal**
  - high-order F.E.
  - trilinear F.E.
  - decreasing #cores
- #cores < 1000
- small MPI communicator
- single core

**spectral $p$-coarsening**
- geometric $h$-coarsening
- algebraic coars.

**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 + $\mu$-BFBT

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>0.11</td>
<td>18</td>
<td>0.02</td>
<td>8</td>
<td>0.12</td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.82</td>
<td>18</td>
<td>0.13</td>
<td>7</td>
<td>0.95</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>6.44</td>
<td>18</td>
<td>1.05</td>
<td>6</td>
<td>7.49</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>50.92</td>
<td>18</td>
<td>8.39</td>
<td>6</td>
<td>59.31</td>
<td>34</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>405.02</td>
<td>18</td>
<td>67.11</td>
<td>6</td>
<td>472.12</td>
<td>34</td>
</tr>
<tr>
<td>9</td>
<td>32</td>
<td>3230.67</td>
<td>18</td>
<td>536.87</td>
<td>6</td>
<td>3767.54</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>64</td>
<td>25807.57</td>
<td>18</td>
<td>4294.97</td>
<td>6</td>
<td>30102.53</td>
<td>34</td>
</tr>
</tbody>
</table>
Algorithmic scalability for HMG + $\mu$-BFBT

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>0.82</td>
<td>18</td>
<td>0.13</td>
<td>7</td>
<td>0.95</td>
<td>33</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2.74</td>
<td>20</td>
<td>0.32</td>
<td>8</td>
<td>3.07</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>6.44</td>
<td>20</td>
<td>0.66</td>
<td>7</td>
<td>7.10</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>12.52</td>
<td>23</td>
<td>1.15</td>
<td>12</td>
<td>13.67</td>
<td>43</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>21.56</td>
<td>23</td>
<td>1.84</td>
<td>12</td>
<td>23.40</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>64</td>
<td>34.17</td>
<td>22</td>
<td>2.75</td>
<td>10</td>
<td>36.92</td>
<td>54</td>
</tr>
<tr>
<td>8</td>
<td>128</td>
<td>50.92</td>
<td>22</td>
<td>3.93</td>
<td>10</td>
<td>54.86</td>
<td>67</td>
</tr>
</tbody>
</table>
Outline

Driving scientific problem & computational challenges

Class of benchmark problems

$\mu$-BFBT and improved robustness of over established state of the art

Modifications for Dirichlet boundary conditions

Algorithmic scalability for HMG+$\mu$-BFBT

Parallel scalability for HMG+$\mu$-BFBT
Weekly scalability for HMG + $\mu$-BFBT

Perform 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]
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]
"μ-BFBT Preconditioner for Stokes Flow Problems" by Johann Rudi

References

Viscosity-weighted pressure mass matrix for Stokes:


BFBT for Navier-Stokes:


BFBT for Stokes: