Massively Parallel PDE Solvers for Uncertainty Quantification
SIAM Conference on Uncertainty Quantification (UQ14)

Björn Gmeiner
w. U. Rüde, R. Scheichl, M. Huber, C. Waluga, B. Wohlmuth

April, 2014
Contents

Introduction

Simulation of the forward problem

Parallel examples

Multilevel Monte Carlo

Multilevel Monte Carlo: first experiments
Introduction
Simulation of seismic velocities (mantle convection)

"The only way to cool a planet is convection."

Fig. by Schuberth and Moder - Simulation with TERRA (LMU)
Boussinesq model for mantle convection problems

derived from the equations for balance of forces, conservation of mass and energy:

\[-\nabla \cdot (2\eta \epsilon(u)) + \nabla p = \rho(T)g,\]
\[\nabla \cdot u = 0,\]
\[\frac{\partial T}{\partial t} + u \cdot \nabla T - \nabla \cdot (\kappa \nabla T) = \gamma.\]

| \(u\)   | velocity               |
| \(p\)   | dynamic pressure       |
| \(T\)   | temperature            |
| \(\nu\) | viscosity of the material |
| \(\epsilon(u) = \frac{1}{2}(\nabla u + (\nabla u)^T)\) | strain rate tensor |
| \(\rho\) | density                |
| \(\kappa, \gamma, \mathbf{g}\) | thermal conductivity, heat sources, gravity vector |
Discretization

Temporal for the temperature (explicit):
- Modified Euler
- BDF-2 scheme

Spatial with FE for the Stokes system:
- Add an stabilization term
  - Choose a pair of FE spaces that satisfy the LBB condition:
    \[ Q_{q+1}^d \times Q_q, \ (q \geq 1) \] - Taylor-Hood elements
    \[ Q_{q+1}^d \times P_{-q}, \ (q \geq 0) \] - discontinuous elements

---

\(^1\) Hughes et al. A new finite element formulation for computational fluid dynamics: V. circumventing the Babuska-Brezzi condition...

Solution of the Stokes equations

- **Stokes equation:**
  \[-\text{div}(\nabla \mathbf{u} - p \mathbf{I}) = \mathbf{f},\]
  \[\text{div} \mathbf{u} = 0\]

- **FEM Discretization:**
  \[a(u_l, v_l) + b(v_l, p_l) = L(v_l) \quad \forall v_l \in V_l,\]
  \[b(u_l, q_l) - c(p_l, q_l) = 0 \quad \forall q_l \in Q_l,\]
  with:
  \[a(u, v) := \int_{\Omega} \nabla u : \nabla v \, dx, \quad b(u, q) := -\int_{\Omega} \text{div} u \cdot q \, dx\]

- **Schur-complement formulation:**
  \[
  \begin{bmatrix}
  A_l & B_l^T \\
  0 & C_l + B_l A_l^{-1} B_l^T 
  \end{bmatrix}
  \begin{bmatrix}
  u_l \\
  p_l 
  \end{bmatrix}
  =
  \begin{bmatrix}
  f_l \\
  B_l A_l^{-1} f_l 
  \end{bmatrix}
  \]
Simulation of the forward problem
HHG - Combining tetrahedral elements with multigrid

Semi-structured mesh

Tetrahedral refinement
Discretization with prismatic elements
Spherical refinement of the icosahedral mesh
Regular refinement of each block (non-curved boundaries)
Parallel examples
Weak scaling (Juqueen)

\[
\mu \Delta u - \nabla p = \text{Ra} \, T \, e_r, \\
\text{div}(u) = 0.
\]

<table>
<thead>
<tr>
<th>Nodes</th>
<th>Threads</th>
<th>Grid points</th>
<th>Resolution</th>
<th>Time-step</th>
<th>Stationary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>$2.1 \cdot 10^9$</td>
<td>32 km</td>
<td>30 s</td>
<td>89 s</td>
</tr>
<tr>
<td>4</td>
<td>240</td>
<td>$1.6 \cdot 10^8$</td>
<td>16 km</td>
<td>38 s</td>
<td>114 s</td>
</tr>
<tr>
<td>30</td>
<td>1 920</td>
<td>$1.3 \cdot 10^9$</td>
<td>8 km</td>
<td>40 s</td>
<td>121 s</td>
</tr>
<tr>
<td>240</td>
<td>15 360</td>
<td>$1.1 \cdot 10^{10}$</td>
<td>4 km</td>
<td>44 s</td>
<td>133 s</td>
</tr>
<tr>
<td>1 920</td>
<td>122 880</td>
<td>$8.5 \cdot 10^{10}$</td>
<td>2 km</td>
<td>48 s</td>
<td>153 s</td>
</tr>
<tr>
<td>15 360</td>
<td>983 040</td>
<td>$6.9 \cdot 10^{11}$</td>
<td>1 km</td>
<td>54 s</td>
<td>170 s</td>
</tr>
</tbody>
</table>
Run-times of strong scaling experiments of the spherical shell geometry with \(10^{10}\) grid points.
Time-dependent calculation

LSS-Cluster (256 cores, $\approx 50\,000$ core hours, $Ra=10^7$, $5 \cdot 10^9$ degree of freedoms)
Multilevel Monte Carlo
Standard Monte Carlo

\[ \nabla \cdot (k(x, \omega) \nabla p(x, \omega)) = f(x, \omega), \ \omega \in \Omega \]

- **Sampling** from \( k(x, \omega) \) by e.g.:
  - Truncated Karhunen-Loeve (KL) expansion
  - Circulant embedding (FFT)
  - PDE-based variants

\[ \text{Standard Monte Carlo estimator:} \]
\[ E[Q_{MC}] := \frac{1}{N} \sum_{i=1}^{N} Q(x_i) \]
where \( Q \) is the quantity of interest.

\[ \text{Here, the mean square error (MSE) is:} \]
\[ E[(Q_{MC} - E[Q])^2] = \text{sampling error} + \text{discretization error} \]
Standard Monte Carlo

\[ \nabla \cdot (k(x, \omega) \nabla p(x, \omega)) = f(x, \omega), \ \omega \in \Omega \]

- **Sampling** from \( k(x, \omega) \) by e.g.:
  - Truncated Karhunen-Loeve (KL) expansion
  - Circulant embedding (FFT)
  - PDE-based variants

- **Standard Monte Carlo** estimator:

\[ \mathbb{E}[Q_{hMC}] := \frac{1}{N} \sum_{i=1}^{N} Q_{h}^{(i)}, \]

where \( Q \) is the quantity of interest.

- Here, the **mean square error** (MSE) is:

\[ \mathbb{E}[(Q_{hMC} - \mathbb{E}[Q])^2] = \frac{\mathbb{V}[Q_h]}{N} + (\mathbb{E}[Q_h - Q])^2 \]

= sampling error + discretization error
Multilevel Monte Carlo

Because of the linearity of the expectation operator:

$$\mathbb{E}[Q_{L}^{ML}] = \mathbb{E}[Q_{0}] + \sum_{l=1}^{L} \mathbb{E}[Q_{l} - Q_{l-1}],$$

we can define the **multilevel Monte Carlo** estimator:

$$Q_{L}^{ML} = Q_{0}^{MC} + \sum_{l=1}^{L} Y_{l}^{MC}$$

with

$$Y_{l} := Q_{l} - Q_{l-1}.$$  

Important observation (compute corrections):

$$\nabla [Q_{l} - Q_{l-1}] \rightarrow 0 \text{ as } h_{l} \rightarrow 0$$

---


4 Cliffe, Giles, Scheichl, Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. Computing and Visualization in Science 14 (1),
Basic Multilevel Monte Carlo

Algorithm\(^5\):

1. Start with \( L = 0 \)
2. Estimate \( \mathbb{V}[Y_L] \) by an initial number of samples
3. Based on these estimate calculate the optimal \( N_l, l = 0, \ldots, L \)
4. Evaluate extra samples at each level
5. Check if the bias is small enough
6. If not, increase level \( L \) by one and go back to 2.

To increase concurrency, we would like to parallelize both loops over the levels.
- We know the costs of a sample on each level
- We do not know \( N_l \) and \( L \) in advance

\(^5\)Cliffe, Giles, Scheichl, Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. Computing and Visualization in Science 14 (1), 3-15
Basic Multilevel Monte Carlo

However, the parallelization of these loops leads to a load balancing problem:

How to distribute the different levels to the available processes?
Strong scaling MLMC: some possible ideas

**Bulk synchronous parallelization** (our favourite)
- Tasks have to be scheduled efficiently between two synchronization points
- We can split up the MPI communicator to be minimal intrusive to the application

**Master-client parallelization**
- Splitting up MPI communicators is hardly possible, since there are no synchronization points

**Asynchronous parallelization**
- MPI 2.0 functionality required (one-sided communication)
Parallelization over the samples for the coarser levels
Parallelization within the solver for the finer levels
Scheduling problem (sketch)

Objectives: a) minimize the required time and b) optimize the number of synchronization points
Flexibility: Strong scaling of the forward problem for certain runs.
Implementation

**Aim:** Design of a C layer/wrapper, which is minimal intrusive.

Possibly most intrusive changes for the solver:
- Allow to use a splitted MPI communicator for the forward simulation for all MPI calls
- No memory leaks at the end of a simulation run

Our current investigations to solve the scheduling problem:
- Genetic multi-objective optimization algorithm (NSGA-II or SPEA-II)
- Derivation of good heuristics
Multilevel Monte Carlo: first experiments
Our strategy to resolve the profiles:

- Resolve the large jumps by the coarsest mesh
- Stabilize the geometric multigrid as far as possible for the finer variations in viscosity
Table: Asymptotic MG convergence rates by applying the V(3,3)-cycle with $5.6 \cdot 10^7$ grid points.
Table: Asymptotic MG convergence rates by applying the V(3,3)-cycle with $1.7 \cdot 10^7$ grid points.
Coarse grid variations (randomly layered domain)

Asymptotic MG convergence rate of 0.2 for a V(3,3)-cycle
Fine-grid variations

2d coefficient field (left) and resulting streamlines for the gradient of $p$ (right) on a sample with lognormal distribution, exponential covariance, correlation length $\lambda = 0.2$ and variance $\sigma^2 = 3$. 
Performance plots

10^5 evaluations on each level, lognormal distribution, exponential covariance, correlation length $\lambda = 0.2$, variance $\sigma^2 = 3$, and $9^3$ grid points on level 1 ($257^3$ on level 6). Quantity of interest: $p$ at one point.
Summary and Outlook

Summary

► Introduction into the basic equations of mantle convection simulations
► Scalability of a multilevel Stokes solver
► An parallel implementation of a Multilevel Monte Carlo method

Current work

► Improved load balancing within the Multilevel Monte Carlo method
► Application to the Stokes problem
Thank you for your attention!