Solution of Three-dimensional Heterogeneous Helmholtz Problems in Geophysics

Henri Calandra\textsuperscript{1}
Rafael Ferreira Lago\textsuperscript{2}
Serge Gratton\textsuperscript{3}
Xavier Pinel\textsuperscript{2}
Xavier Vasseur\textsuperscript{2}

\textsuperscript{1}Total \textsuperscript{2}CERFACS \textsuperscript{3}INPT-IRIT

Journée des Doctorant, September 07, 2010
1 Formulation
2 Method for 7-points Stencil
3 Method for 27-points Stencil
4 Conclusion
1 Formulation
2 Method for 7-points Stencil
3 Method for 27-points Stencil
4 Conclusion
Depth Migration in Geophysics

- Trigger a source on the surface
- Detect the discontinuities
- Estimate the position and thickness of reflective layers
Three-Dimensional Helmholtz Equation

Three-dimensional Wave Equation

\[
\frac{1}{v^2} \frac{\partial^2 w}{\partial t^2} = \Delta w, \quad w(x, y, z, t) = u(x, y, z) \exp(-i2\pi ft)
\]

Helmholtz equation in frequency domain

\[-\Delta u - k^2 u = s\]

- \(u(x, y, z) : \mathbb{C}^3 \rightarrow \mathbb{C}\) is the wave pressure function
- \(k(x, y, z) : \mathbb{R}^3 \rightarrow \mathbb{R}\) is the wavenumber, defined as
  \[k(x, y, z) = \frac{2\pi f}{v(x, y, z)}\]
  - \(v(x, y, z) : \mathbb{R}^3 \rightarrow \mathbb{R}\) is the velocity field
  - \(f\) is the frequency in \(Hz\)
- \(s\) is the source term
Velocity Field: SEG/EAGE Salt dome

**Size:** $13.5 \times 13.5 \times 4 \text{ km}^3$

**Velocity:**
- $v_{min} = 1500 \text{ m/s}$
- $v_{max} = 4481 \text{ m/s}$
Finite Differences 7-points Stencil

(a) Cartesian stencil

(b) Pattern of the matrix

Characteristics:

- Second order finite difference
- Sparse matrices
- Simple to implement
- Almost Matrix-free implementation
Finite Differences 27-points Stencil
[Operto et al., 2007]

**Characteristics:**
- Second order finite difference
- Less-sparse matrix
- Stencil designed to minimize the dispersion error for a given number of points per wavelength
Stability Condition

Stability Condition for second order (Cohen, 2002)

To have a physical significance, we must satisfy

\[ k(x, y, z) \leq \frac{2\pi}{n_{\lambda} h}, \quad \forall (x, y, z) \in \Omega \]

- \( h \) : distance between each point in the grid
- \( n_{\lambda} \) : number of points per wavelength

remembering that

\[ k(x, y, z) = \frac{2\pi f}{v(x, y, z)} \]

- We set
  \[ h = \frac{v_{min}}{n_{\lambda} f} \]

- When \( n_{\lambda} \) or \( f \) increases, the grid size increases as well
- For the 7-points stencil we fix \( n_{\lambda} = 12 \)
Solution for SEG/EAGE Salt dome

(a) Solution for $2.5 \ Hz$

(b) Solution for $5 \ Hz$

(c) Solution for $10 \ Hz$

(d) Solution for $20 \ Hz$
Outline

1  Formulation
2  Method for 7-points Stencil
3  Method for 27-points Stencil
4  Conclusion
Previous Works

Solution methods relying on Local or Global Gaussian Elimination

- Sparse multifrontal direct methods [Operto et al., 2007]
  - Part of Overthrust model, 10 Hz, $1.8 \times 10^7$: 450 GB of memory
- Multilevel $LDL^T$ factorization as a preconditioner [Bollhöfer et al., 2009]
  - Overthrust model, 5 Hz, $1.5 \times 10^7$: 32 GB of memory
- Domain Decomposition - Algebraic Additive Schwarz preconditioner [Haidar, 2009]
  - Overthrust model, 7 Hz, $5.6 \times 10^6$: 150 GB of memory.

Main references related to Multigrid

- Multigrid on a complex shifted Helmholtz operator [Erlangga et al., 2006]
- Multigrid with Krylov methods as a "smoother" and few grids in the hierarchy [Elman et al., 2001]
- Two-grid preconditioner in 2D using sparse direct solver for the coarse grid problems [Duff et al., 2007]
- Perturbed two-level preconditioner in 3D using iterative solver for the coarse grid problems [Pinel, 2010]
An Iterative Solver...

- **Outer Method:** FGMRES(5) [Saad, 1993]
- **Convergence Criterion:** \( \frac{||r_j||_2}{||b||_2} < 10^{-5} \)
- **Inner Method:** two-level geometric multigrid
- **Standard geometric coarsening in all directions**
- **Coarse and Smoother:** Krylov method - which?
Local Fourier Analysis

- Analysis made on the preconditioned matrix $AM^{-1}$
- **Coarse**: $\varepsilon_{2h} = 10^{-1}$ produces a spectrum similar to $\varepsilon_{2h} = 0$
- Very accurate solution in coarse level is not necessary
- We choose GMRES preconditioned by local symmetric Gauss Seidel to solve the coarse problem
Homogeneous problem

**Coarse:** Fixing the number of iterations to 100 can improve the performance over $\varepsilon_{2h} = 10^{-1}$

**Smother:** 2 iterations of GMRES preconditioned by local symmetric Gauss-Seidel is the most efficient

<table>
<thead>
<tr>
<th>$f$ (Hz)</th>
<th>CPU 128</th>
<th>CPU 1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{2h}$</td>
<td>$\text{lt}$</td>
<td>$T(s)$</td>
</tr>
<tr>
<td>1</td>
<td>529</td>
<td>1989</td>
</tr>
<tr>
<td>0.7</td>
<td>48</td>
<td>298</td>
</tr>
<tr>
<td>0.5</td>
<td>25</td>
<td>235</td>
</tr>
<tr>
<td>0.3</td>
<td>18</td>
<td>244</td>
</tr>
<tr>
<td>0.2</td>
<td>16</td>
<td>276</td>
</tr>
<tr>
<td>0.1</td>
<td>15</td>
<td>458</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>13</td>
<td>2639</td>
</tr>
<tr>
<td>100 it</td>
<td>21</td>
<td>234</td>
</tr>
</tbody>
</table>

Table: Comparison between different tolerances and (fixed) 100 iterations

- Test performed on IBM Blue Gene/L (CERFACS) using VN
- **Coarse Convergence:** $\frac{||r_{2h}||_2}{||b_{2h}||_2} < \varepsilon_{2h}$
**Perturbed Two-Level Preconditioner**

- **Outer Method:** FGMRES(5)
- **Convergence Criterion:** \( \frac{||r_j||_2}{||b||_2} < 10^{-5} \)
- **Inner Method:** two-level geometric multigrid
- Standard geometric coarsening in all directions

### Coarse Level
- **Method:** GMRES(5)
- **Precond.:** local symmetric Gauss-Seidel
- 100 iterations

### Smoother
- **Method:** GMRES(2)
- **Precond.:** local symmetric Gauss-Seidel
- 2 iterations is sufficient
Parallelization of the Method

(a) Example of a 24 processors Cartesian Grid

(b) Ghost Nodes

- Divide processors in \( p_x \times p_y \times p_z \)
- Attempt to divide the discretized domain equally among the processors
- Use of ghost nodes for updating bounds

- \( V \) matrix (Krylov vectors) splitted among processors
- \( \bar{H} \) Hessenberg matrix copied in each processor
- Least squares solved locally in each processor
IBM Blue Gene/L CERFACS

- **1024** nodes with **1GB**
- **1024** nodes with **512MB**

**Node characteristics:**
- 2 **PowerPC440** cores, 700Mhz
- custom 128-bits double FPU for each core
- 32-bits addressing
- L1=32KB (per core), L2=2KB (per core), L3=4MB (per node)
- **CO**: Peak of **2.8GFlop/s** per node
- **VN**: Peak of **5.6GFlop/s** per node

**Communication**
- 3D-Torus between nodes, **175 MB/s per link**, latency of 4-10 μs
- A tree network for global synchronization and interruptions, **350 MB/s per link**, latency of 1.3 μs
Performance of Perturbed Two-Level Preconditioner Method

7-points Stencil Results

- Method works very well for 7-points stencil
- Successfully obtained the solution up to $40\,\text{Hz}$ ($\approx 2.5 \times 10^9$ of unknowns) for SEG/EAGE Salt dome velocity field using 4232 GB of memory, using 2048 processors on an SGI Xenon Jade (CINES)
- Good performance in strong scalability tests

Storage Requirements for SEA/EAGE Overthrust:

<table>
<thead>
<tr>
<th>$f$</th>
<th>Grid</th>
<th>Grid Size</th>
<th>Mem.(GB)</th>
<th>#Processors</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$583 \times 583 \times 162$</td>
<td>$5.5 \times 10^7$</td>
<td>10</td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td>$805 \times 805 \times 215$</td>
<td>$1.4 \times 10^8$</td>
<td>25</td>
<td>256</td>
</tr>
<tr>
<td>10</td>
<td>$1134 \times 1134 \times 292$</td>
<td>$3.7 \times 10^8$</td>
<td>66</td>
<td>2048</td>
</tr>
<tr>
<td>30</td>
<td>$3356 \times 3356 \times 829$</td>
<td>$2.7 \times 10^9$</td>
<td>1604</td>
<td>16384</td>
</tr>
</tbody>
</table>
Outline

1  Formulation
2  Method for 7-points Stencil
3  Method for 27-points Stencil
4  Conclusion
27-points Stencil Overview

Advantages of 27-points Stencil

- Operto et al. set the weights for 4 points per wavelength
- Reduced size of the system

Difficulties

- Operto et al. use this 27-points stencil with direct and hybrid solvers
- We must use at least 4 points per wavelength in the coarse level, hence we must set $n_\lambda = 8$ on the fine level at least
**Increased $n_\lambda$ Strategy**

**Experiments**
- SEG/EAGE Salt dome model
- IBM Blue Gene/L (CERFACS) using VN
- Heterogeneous problem

**Comparison between different $n_\lambda$**

<table>
<thead>
<tr>
<th>$n_\lambda$</th>
<th>$h$</th>
<th>Grid</th>
<th>Grid Size</th>
<th>#It</th>
<th>$T$ (m)</th>
<th>M (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>37.5</td>
<td>$392 \times 392 \times 144$</td>
<td>$2.21 \times 10^7$</td>
<td>568</td>
<td>137.3</td>
<td>7.9</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>$482 \times 482 \times 172$</td>
<td>$3.99 \times 10^7$</td>
<td>32</td>
<td>13.5</td>
<td>14.4</td>
</tr>
<tr>
<td>12</td>
<td>25</td>
<td>$572 \times 572 \times 200$</td>
<td>$6.54 \times 10^7$</td>
<td>23</td>
<td>15.9</td>
<td>23.5</td>
</tr>
</tbody>
</table>

- $f = 5Hz$ for all tests
- Optimal performance for $n_\lambda = 10$
Increased $n_\lambda$ Residuals

**Figure:** FGMRES Residual - $\log_{10}\left(\frac{||r_j||_2}{||r_0||_2}\right)$

**Figure:** Coarse Residual after 100 coarse iterations - $\log_{10}\left(\frac{||r_{100}^h||_2}{||r_0^h||_2}\right)$
Strategy

- Use only one-level preconditioner and $n_\lambda = 4$
- Outer Method: FGMRES(5)
- Convergence Criterion: $\frac{||r_j||_2}{||b||_2} < 10^{-5}$
- Inner Method: 10 iterations of GMRES(5) preconditioned by local symmetric Gauss-Seidel
Comparison Between Strategies

- **7-points** stencil with \( n_\lambda = 12 \)
- FGMRES(5) with Perturbed Two-Level Preconditioner

- **27-points** stencil with \( n_\lambda = 4 \)
- FGMRES(5) with One-Level Preconditioner

- **27-points** stencil with \( n_\lambda = 10 \)
- FGMRES(5) with Perturbed Two-Level Preconditioner

**Experiments**

- **SEG/EAGE Salt dome model**
- IBM Blue Gene/L (CERFACS) using VN
- Heterogeneous problem
Comparison Between Strategies

[2-Levels | $n_\lambda = 10$ | 27-points] vs. [2-Levels | $n_\lambda = 12$ | 7-points]

### Parameters

- **For $n_\lambda = 10$**
  - $h = 15m$
  - $f = 10Hz$
  - Grid=$932 \times 932 \times 311$
    ($\approx 2.7 \times 10^8$)
  - **105.2 GB** of memory

- **For $n_\lambda = 12$**
  - $h = 12.5m$
  - $f = 10Hz$
  - Grid=$1112 \times 1112 \times 367$
    ($\approx 4.53 \times 10^8$)
  - **75.2 GB** of memory

<table>
<thead>
<tr>
<th>#Cores</th>
<th>Partition</th>
<th>$T(m)_{10}$</th>
<th>#It$_{10}$</th>
<th>$T(m)_{12}$</th>
<th>#It$_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024</td>
<td>$16 \times 8 \times 8$</td>
<td>49.7</td>
<td>68</td>
<td>62.5</td>
<td>50</td>
</tr>
<tr>
<td>2048</td>
<td>$16 \times 16 \times 8$</td>
<td>24.9</td>
<td>67</td>
<td>33.0</td>
<td>53</td>
</tr>
<tr>
<td>4096</td>
<td>$16 \times 16 \times 16$</td>
<td>15.3</td>
<td>85</td>
<td>23.3</td>
<td>75</td>
</tr>
</tbody>
</table>
Comparison Between Strategies

[2-Levels | \( n_\lambda = 10 \) | 27-points] vs. [1-Level | \( n_\lambda = 4 \) | 27-points]

### Parameters

- **For \( n_\lambda = 10 \)**
  - \( h = 15m \)
  - \( f = 10Hz \)
  - Grid=932 \times 932 \times 311 \( \approx 2.7 \times 10^8 \)
  - **105.2 GB** of memory

- **For \( n_\lambda = 4 \)**
  - \( h = 25m \)
  - \( f = 10Hz \)
  - Grid=572 \times 572 \times 200 \( \approx 6.54 \times 10^7 \)
  - **24.8 GB** of memory

<table>
<thead>
<tr>
<th>#Cores</th>
<th>Partition</th>
<th>( T(m)_{10} )</th>
<th>#It(_{10} )</th>
<th>( T(m)_{4} )</th>
<th>#It(_{4} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>8 \times 8 \times 8</td>
<td>99.6</td>
<td>68</td>
<td>119.9</td>
<td>572</td>
</tr>
<tr>
<td>1024</td>
<td>16 \times 8 \times 8</td>
<td>49.7</td>
<td>68</td>
<td>64.8</td>
<td>641</td>
</tr>
<tr>
<td>2048</td>
<td>16 \times 16 \times 8</td>
<td>24.9</td>
<td>67</td>
<td>37.5</td>
<td>724</td>
</tr>
<tr>
<td>4096</td>
<td>16 \times 16 \times 16</td>
<td>15.3</td>
<td>85</td>
<td>19.8</td>
<td>786</td>
</tr>
</tbody>
</table>
1 Formulation
2 Method for 7-points Stencil
3 Method for 27-points Stencil
4 Conclusion
Conclusion and Future Work

**Numerical Experiments**

- Satisfactory performance of the method for 27-points
  - Approximately 40% more storage requirements
  - Approximately 30% faster for the same frequency
  - Good **strong scalability**
- Satisfactory performance of one-level alternative
  - Almost the same time performance
  - Approximately 1/3 of the storage requirements
  - Reasonably **strong scalability**
- Compare the quality and accuracy of the solution obtained by each stencil

**Theoretical Work**

- Extend Local Fourier Analysis to 27-points stencil and to higher-order discretization methods


