Multi-level Krylov: the next generation Helmholtz solver


Kees Vuik, Abdul Sheikh and Domenico Lahaye
http://ta.twi.tudelft.nl/users/vuik/
July 17, 2015
Application: geophysical survey

Marine Seismic
Application: geophysical survey

hard Marmousi Model
Application: geophysical survey

hard Marmousi Model (2006)

![Graph showing old method and new method iterations vs. frequency](image-url)
Contents

1. Introduction
3. Numerical experiments
5. Fourier Analysis of two-level methods
6. Numerical experiments
7. Conclusions
1. Introduction

The Helmholtz equation without damping

\[-\Delta u(x, y) - k^2(x, y)u(x, y) = g(x, y) \quad \text{in } \Omega\]

\(u(x, y)\) is the pressure field,
\(k(x, y)\) is the wave number,
\(g(x, y)\) is the point source function and
\(\Omega\) is the domain. Absorbing boundary conditions are used on \(\Gamma\).

\[\frac{\partial u}{\partial n} - \nu u = 0\]

\(n\) is the unit normal vector pointing outwards on the boundary.

Perfectly Matched Layer (PML) and Absorbing Boundary Layer (ABL)
Problem description

• Second order Finite Difference stencil:

\[
\begin{pmatrix}
-1 \\
-1 & 4 - k^2 h^2 & -1 \\
-1
\end{pmatrix}
\]

• Linear system \( Au = g \): properties
  
  Sparse & complex valued
  Symmetric & Indefinite for large \( k \)

• For high resolution a very fine grid is required: 10 – 20 gridpoints per wavelength \( \rightarrow A \) is extremely large!

• Is traditionally solved by a Krylov subspace method, which exploits the sparsity.
2. Preconditioning

Equivalent linear system \( M_1^{-1} A M_2^{-1} \tilde{x} = \tilde{b} \), where \( M = M_1 \cdot M_2 \) is the preconditioning matrix and

\[
\tilde{x} = M_2 x, \quad \tilde{b} = M_1 b.
\]

Requirements for a preconditioner

- better spectral properties of \( M^{-1} A \)
- cheap to perform \( M^{-1} r \).

Spectrum of \( A \) is \( \{\mu_i - k^2\} \), with \( k \) a given constant and \( \mu_i \) are the eigenvalues of the Laplace operator. Note that \( \mu_1 - k^2 \) may be negative.
## Preconditioning (overview)

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILU</td>
<td>Meijerink and van der Vorst, 1977</td>
</tr>
<tr>
<td>ILU(tol)</td>
<td>Saad, 2003</td>
</tr>
<tr>
<td>SPAI</td>
<td>Grote and Huckle, 1997</td>
</tr>
<tr>
<td>Multigrid</td>
<td>Lahaye, 2001</td>
</tr>
<tr>
<td></td>
<td>Elman, Ernst and O’ Leary, 2001</td>
</tr>
<tr>
<td>AILU</td>
<td>Gander and Nataf, 2001</td>
</tr>
<tr>
<td></td>
<td>analytic parabolic factorization</td>
</tr>
<tr>
<td>ILU-SV</td>
<td>Plessix and Mulder, 2003</td>
</tr>
<tr>
<td></td>
<td>separation of variables</td>
</tr>
</tbody>
</table>
Preconditioning (Laplace type)

Laplace operator            Bayliss and Turkel, 1983
Definite Helmholtz          Laird, 2000

Shifted Laplace preconditioner (SLP)

\[ M \equiv -\Delta - (\beta_1 - i\beta_2)k^2, \quad \beta_1, \beta_2 \in \mathbb{R}. \]

Condition \( \beta_1 \leq 0 \) is used to ensure that \( M \) is a (semi) definite operator.

- \( \rightarrow \beta_1, \beta_2 = 0 \) : Bayliss and Turkel
- \( \rightarrow \beta_1 = -1, \beta_2 = 0 \) : Laird
- \( \rightarrow \beta_1 = 1, \beta_2 = 0.5 \) : Y.A. Erlangga, C. Vuik and C.W.Oosterlee
3. Numerical experiments

Example with constant \( k \) in \( \Omega \)

Iterative solver: Bi-CGSTAB

Preconditioner: Shifted-Laplace operator, discretized using the same method as the Helmholtz operator.

<table>
<thead>
<tr>
<th>( k )</th>
<th>ILU(0.01)</th>
<th>( M_0 )</th>
<th>( M_{-1} )</th>
<th>( M_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>9</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>29</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>15</td>
<td>47</td>
<td>114</td>
<td>45</td>
<td>26</td>
</tr>
<tr>
<td>20</td>
<td>82</td>
<td>354</td>
<td>85</td>
<td>34</td>
</tr>
<tr>
<td>30</td>
<td>139</td>
<td>&gt; 1000</td>
<td>150</td>
<td>52</td>
</tr>
</tbody>
</table>
Spectrum of SLP

References: Manteuffel, Parter, 1990; Yserentant, 1988

Since $L \equiv -\Delta$ is SPD we have the following eigenpairs

$$Lv_j = \lambda_j v_j, \text{ where, } \lambda_j \in \mathbb{R}^+$$

The eigenvalues $\sigma_j$ of the preconditioned matrix satisfy

$$(L - z_1 I)v_j = \sigma_j (L - z_2 I)v_j.$$ 

Theorem 1
Provided that $z_2 \neq \lambda_j$, the relation

$$\sigma_j = \frac{\lambda_j - z_1}{\lambda_j - z_2}$$

holds.
Spectrum of SLP

**Theorem 2**
If $\beta_2 = 0$, the eigenvalues $\sigma_r + i\sigma_i$ are located on the straight line in the complex plane given by

$$\beta_1 \sigma_r - (\alpha_1 - \alpha_2)\sigma_i = \beta_1.$$

**Theorem 3**
If $\beta_2 \neq 0$, the eigenvalues $\sigma_r + i\sigma_i$ are on the circle in the complex plane with center $c$ and radius $R$:

$$c = \frac{z_1 - \bar{z}_2}{z_2 - \bar{z}_2}, \quad R = \left| \frac{z_2 - \bar{z}_1}{z_2 - \bar{z}_2} \right|.$$

Note that if $\beta_1 \beta_2 > 0$ the origin is not enclosed in the circle.
Eigenvalues for Complex preco $k = 100$

spectrum is independent of the grid size

75 grid points

150 grid points
**Inner iteration**

Possible solvers for solution of $Mz = r$:

- ILU approximation of $M$
- inner iteration with ILU as preconditioner
- Multigrid

Multigrid components
- geometric multigrid
- Gauss-Seidel with red-black ordering
- matrix dependent interpolation, full weighting restriction
- Galerkin coarse grid approximation
Numerical results for a wedge problem

<table>
<thead>
<tr>
<th>( k_2 )</th>
<th>( 10 )</th>
<th>( 20 )</th>
<th>( 40 )</th>
<th>( 50 )</th>
<th>( 100 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>grid</td>
<td>( 32^2 )</td>
<td>( 64^2 )</td>
<td>( 128^2 )</td>
<td>( 192^2 )</td>
<td>( 384^2 )</td>
</tr>
<tr>
<td>No-Prec</td>
<td>201(0.56)</td>
<td>1028(12)</td>
<td>5170(316)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ILU(( A,0 ))</td>
<td>55(0.36)</td>
<td>348(9)</td>
<td>1484(131)</td>
<td>2344(498)</td>
<td>–</td>
</tr>
<tr>
<td>ILU(( A,1 ))</td>
<td>26(0.14)</td>
<td>126(4)</td>
<td>577(62)</td>
<td>894(207)</td>
<td>–</td>
</tr>
<tr>
<td>ILU(( M,0 ))</td>
<td>57(0.29)</td>
<td>213(8)</td>
<td>1289(122)</td>
<td>2072(451)</td>
<td>–</td>
</tr>
<tr>
<td>ILU(( M,1 ))</td>
<td>28(0.28)</td>
<td>116(4)</td>
<td>443(48)</td>
<td>763(191)</td>
<td>2021(1875)</td>
</tr>
<tr>
<td>MG(V(1,1))</td>
<td>13(0.21)</td>
<td>38(3)</td>
<td>94(28)</td>
<td>115(82)</td>
<td>252(850)</td>
</tr>
</tbody>
</table>
Spectrum with inner iteration

1 MG iteration

2 MG iterations

Summary so far

- ILU and variants
- From Laplace to complex Shifted Laplace Preconditioner (2005)
- Shifted Laplace Preconditioner (SLP)

\[ M := -\Delta u - (\beta_1 - i\beta_2)k^2 u \]

- Results show: \((\beta_1, \beta_2) = (1, 0.5)\) is the shift of choice
- Properties of SLP?
Shifted Laplace Preconditioner (SLP)

- Introduces damping, Multi-grid approximation is possible
- The modulus of all eigenvalues of the preconditioned operator is bounded by 1
- Small eigenvalues move to zero, as $k$ increases.

Spectrum of $M^{-1}(1, 0.5)A$ for $k = 30$ and $k = 120$
Spectrum as function of $k$
Deflation: or two-grid method

Deflation, a projection preconditioner

\[ P = I - AQ, \quad \text{with} \quad Q = ZE^{-1}Z^T \quad \text{and} \quad E = Z^T AZ \]

where,

\[ Z \in \mathbb{R}^{n \times r}, \quad \text{with deflation vectors} \quad Z = [z_1, ..., z_r], \quad \text{rank}(Z) = r \leq n \]

Along with a traditional preconditioner \( M \), deflated preconditioned system reads

\[ PM^{-1}Au = PM^{-1}g. \]

Deflation vectors shifted the eigenvalues to zero.
Deflation for Helmholtz

With choice of multigrid inter-grid transfer operator (Prolongation) as deflation matrix, i.e. \( Z = I_{2h}^h \) and \( Z^T = I_{2h}^h \) then

\[
P_h = I_h - A_h Q_h, \quad \text{with} \quad Q_h = I_{2h}^h A_{2h}^{-1} I_{2h}^h \quad \text{and} \quad A_{2h} = I_{2h}^h A_h I_{2h}^h
\]

where

\( P_h \) can be interpreted as a coarse grid correction and

\( Q_h \) as the coarse grid operator
Deflation: ADEF1

Deflation can be implemented combined with SLP $M_h$,\[ M_h^{-1} P_h A_h u_h = M_h^{-1} P_h g_h \]

$A_h u_h = g_h$ is preconditioned by the two-level preconditioner $M_h^{-1} P_h$.

For large problems, $A_{2h}$ is too large to invert exactly. Inversion of $A_{2h}$ is sensitive, since $P_h$ deflates the spectrum to zero.

To do: Solve $A_{2h}$ iteratively to a required accuracy on certain levels, and shift the deflated spectrum to $\lambda_h^{\text{max}}$ by adding a shift in the two level preconditioner. This leads to the ADEF1 preconditioner

$$P_{(h,\text{ADEF1})} = M_h^{-1} P_h + \lambda_h^{\text{max}} Q_h$$
Deflation: MLKM

Multi Level Krylov Method $^a$, take $\hat{A}_h = M_h^{-1} A_h$, and define $\hat{P}_h$ by using $\hat{A}_h$ (instead of $A_h$) will be

$$\hat{P}_h = I_h - \hat{A}_h \hat{Q}_h,$$

where

$$\hat{Q}_h = I_{2h} \hat{A}_{2h}^{-1} I_{2h}^h \quad \text{and} \quad \hat{A}_{2h} = I_{2h} \hat{A}_h I_{2h}^h = I_{2h} (M_h^{-1} A_h) I_{2h}^h.$$

Construction of coarse matrix $A_{2h}$ at level $2h$ costs inversion of preconditioner at level $h$.

Approximate $A_{2h}$

<table>
<thead>
<tr>
<th>Ideal</th>
<th>Practical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{A}<em>{2h} = I</em>{2h}^h (M_h^{-1} A_h) I_{2h}^2$</td>
<td>$\hat{A}<em>{2h} = I</em>{2h}^h (M_h^{-1} A_h) I_{2h}^2$</td>
</tr>
<tr>
<td>$\hat{A}<em>{2h} \approx I</em>{2h}^h I_{h}^2 (M_{2h}^{-1} A_{2h})$</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Erlangga, Y.A and Nabben R., ETNA 2008
5. Fourier Analysis of two-level methods

Dirichlet boundary conditions for analysis.
With above deflation,

\[
\text{spec}(PM^{-1}A) = f(\beta_1, \beta_2, k, h)
\]

is a complex valued function.

Setting \(kh = 0.625\),

- Spectrum of \(PM^{-1}A\) with shifts \((1, 0.5)\) is clustered around 1 with a few outliers.
- Spectrum remains almost the same, when the imaginary shift for the preconditioner is varied from 0.5 to 1.
Fourier Analysis

**ADEF1**: Analysis shows spectrum clustered around 1 with few outliers.

$k = 30$

$k = 120$
Fourier Analysis

Spectrum of Helmholtz preconditioned by MLKM, $k = 160$ and 20 gp/wl

Ideal

![Ideal Spectrum](image)

Practical

![Practical Spectrum](image)

$^b$Two-level
6. Numerical results (no pollution)

<table>
<thead>
<tr>
<th>k</th>
<th>(N(k^3h^2 \leq 0.625))</th>
<th>iter</th>
<th>(N(kh \leq 0.625))</th>
<th>iter</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>44</td>
<td>4</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>116</td>
<td>4</td>
<td>32</td>
<td>5</td>
</tr>
<tr>
<td>40</td>
<td>320</td>
<td>4</td>
<td>64</td>
<td>5</td>
</tr>
<tr>
<td>80</td>
<td>800</td>
<td>4</td>
<td>128</td>
<td>6</td>
</tr>
<tr>
<td>100</td>
<td>1268</td>
<td>4</td>
<td>160</td>
<td>7</td>
</tr>
<tr>
<td>200</td>
<td>3572</td>
<td>4</td>
<td>320</td>
<td>8</td>
</tr>
<tr>
<td>400</td>
<td>10124</td>
<td>4</td>
<td>340</td>
<td>10</td>
</tr>
<tr>
<td>500</td>
<td>14144</td>
<td>4</td>
<td>800</td>
<td>12</td>
</tr>
<tr>
<td>800</td>
<td>28628</td>
<td>4</td>
<td>1280</td>
<td>15</td>
</tr>
<tr>
<td>1000</td>
<td>400004</td>
<td>4</td>
<td>1600</td>
<td>17</td>
</tr>
</tbody>
</table>
Application: geophysical survey

hard Marmousi Model
Application: geophysical survey

hard Marmousi Model, PETSc solver

$kh = 0.39$, Bi-CGSTAB for SLP, FGMRES(20) for ADEF1(8,2,1)

<table>
<thead>
<tr>
<th>Frequency $f$</th>
<th>Solve Time</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SLP-F</td>
<td>ADEF1-F</td>
</tr>
<tr>
<td>1</td>
<td>1.22</td>
<td>5.07</td>
</tr>
<tr>
<td>10</td>
<td>10.18</td>
<td>9.43</td>
</tr>
<tr>
<td>20</td>
<td>72.16</td>
<td>60.32</td>
</tr>
<tr>
<td>40</td>
<td>550.20</td>
<td>426.79</td>
</tr>
</tbody>
</table>
Application: geophysical survey

Cube with constant $k$
Application: geophysical survey

Cube with constant $k$

<table>
<thead>
<tr>
<th>Wave number</th>
<th>Solve Time</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>SLP-F</td>
<td>ADEF1-F</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>0.32</td>
</tr>
<tr>
<td>10</td>
<td>0.48</td>
<td>2.32</td>
</tr>
<tr>
<td>20</td>
<td>8.14</td>
<td>17.28</td>
</tr>
<tr>
<td>40</td>
<td>228.29</td>
<td>155.52</td>
</tr>
<tr>
<td>60</td>
<td>1079.99</td>
<td>607.45</td>
</tr>
</tbody>
</table>
Application: geophysical survey

Cube with constant $k$

![Graph showing solve time/grid point versus wave number $k$. The graph includes two lines: red for SLP and blue for ADEF-1. The y-axis is labeled as solve time/grid point and ranges from 0 to $1.6 \times 10^{-4}$. The x-axis is labeled as wave number $k$ and ranges from 0 to 60. The graph compares the performance of SLP and ADEF-1.]
Application: geophysical survey

Cube with variable $k$

Grid size $h$ is such that $kh \approx 0.625$

<table>
<thead>
<tr>
<th>$k$</th>
<th>CLSP(time)</th>
<th>ADEF1(time)</th>
<th>CLSP</th>
<th>ADEF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.09</td>
<td>0.24</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>1.07</td>
<td>1.94</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>20</td>
<td>16.7</td>
<td>18.9</td>
<td>32</td>
<td>16</td>
</tr>
<tr>
<td>40</td>
<td>1304</td>
<td>214</td>
<td>331</td>
<td>24</td>
</tr>
</tbody>
</table>
**Application: geophysical survey**

Cube with variable \( k \)

Grid size \( h \) is such that \( kh \approx 0.3125 \)

<table>
<thead>
<tr>
<th>( k )</th>
<th>CLSP(time)</th>
<th>ADEF1(time)</th>
<th>CLSP</th>
<th>ADEF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.6</td>
<td>1.4</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>7.5</td>
<td>10.04</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>324.1</td>
<td>79.2</td>
<td>72</td>
<td>9</td>
</tr>
<tr>
<td>30</td>
<td>3810.9</td>
<td>361.7</td>
<td>285</td>
<td>11</td>
</tr>
</tbody>
</table>
7. Conclusions

- Without deflation, when imaginary shift is increased in SLP, spectrum remains bounded above 1, but lower part moves to zero.
- With deflation the convergence is nearly independent of the imaginary shift.
- With deflation the convergence is initially weakly depending on $k$. For large $k$ it scales again linearly.
- With deflation the CPU time is less than without deflation.
- The convergence of ADEF1 and the practical variant of MLKM are similar.
References

- [http://ta.twi.tudelft.nl/nw/users/vuik/pub_it_helmholtz.html](http://ta.twi.tudelft.nl/nw/users/vuik/pub_it_helmholtz.html)