Multi-level Krylov: the next generation Helmholtz solver

Inverse Problems Seminar Series, UCL Centre for Inverse Problems, London, UK, July 17, 2015

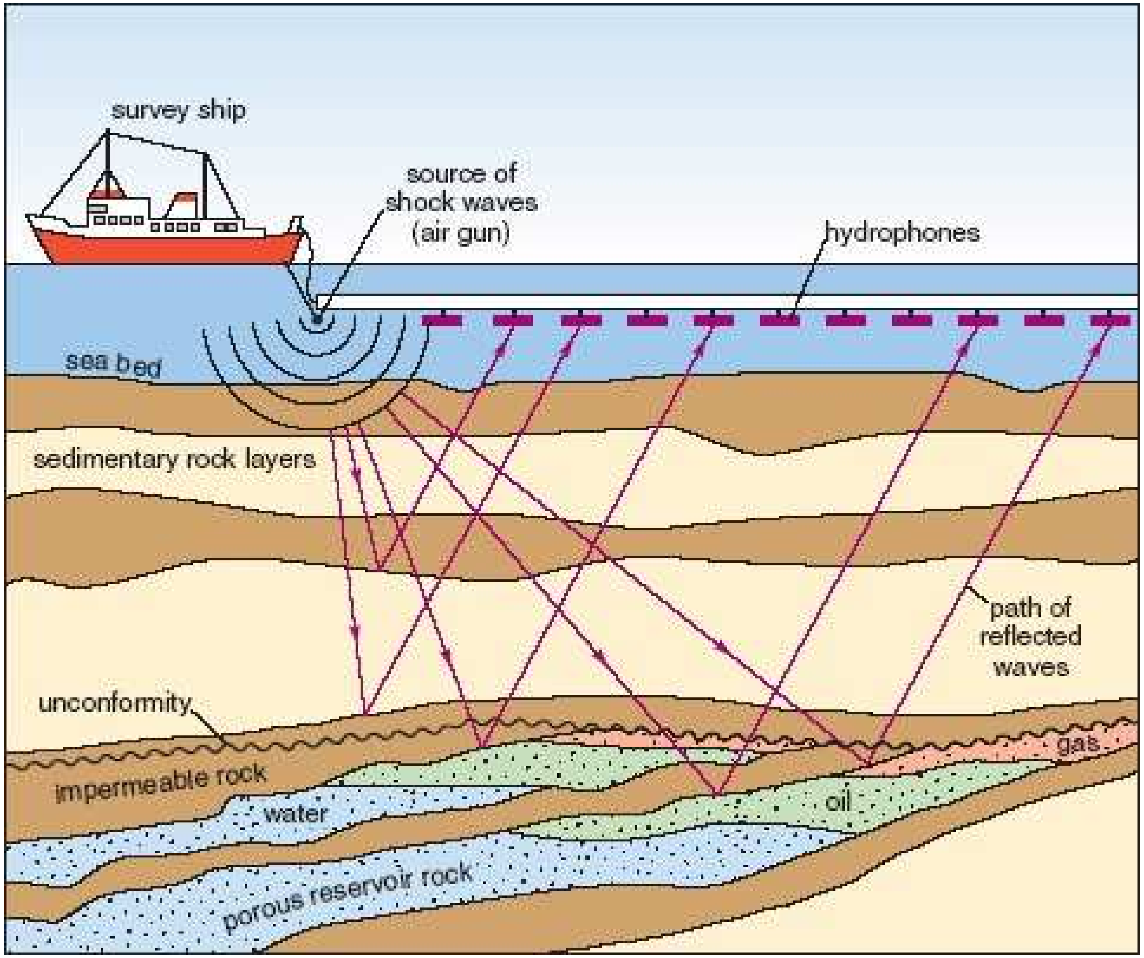
Kees Vuik, Abdul Sheikh and Domenico Lahaye

http://ta.twi.tudelft.nl/users/vuik/

July 17, 2015

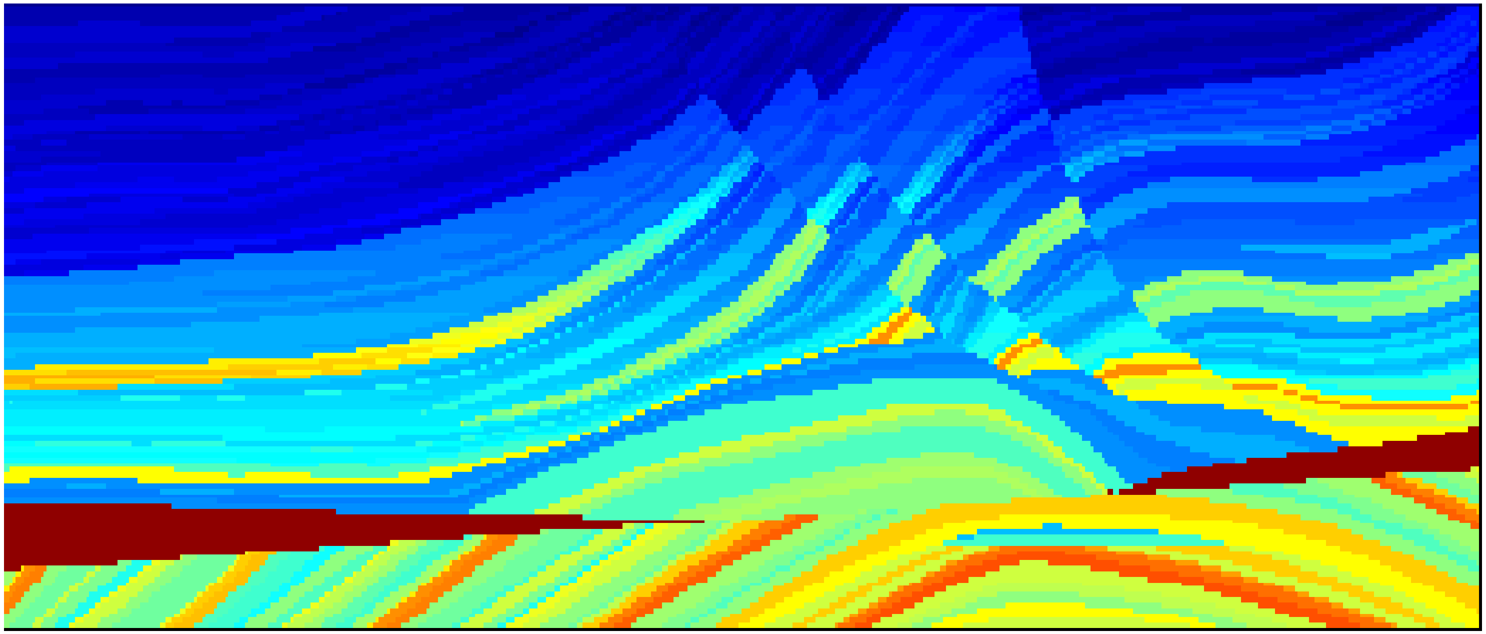
Application: geophysical survey

Marine Seismic

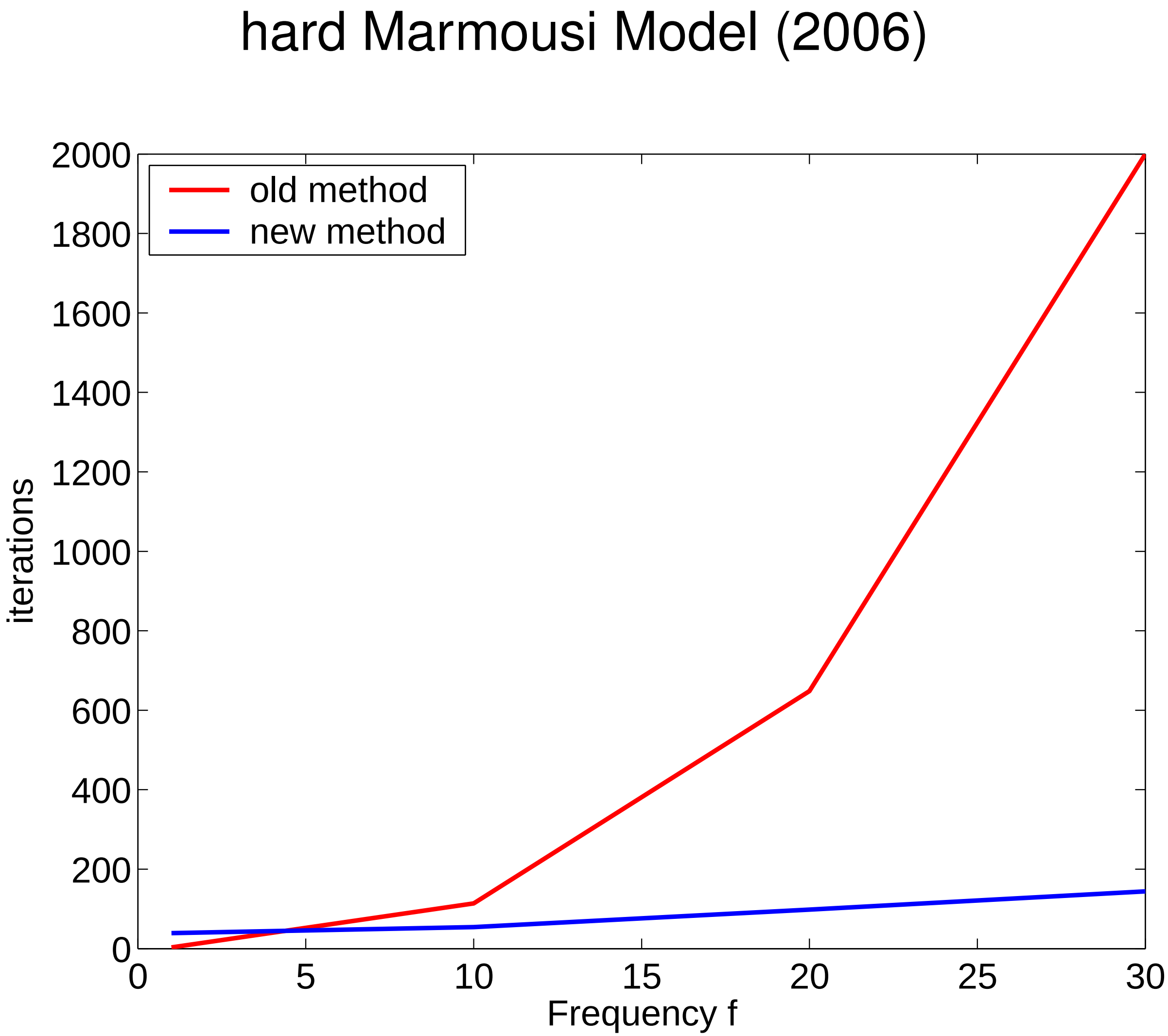


Application: geophysical survey

hard Marmousi Model



Application: geophysical survey



Contents

1. Introduction

The Helmholtz equation without damping

$$-\Delta \mathbf{u}(x,y) - k^2(x,y)\mathbf{u}(x,y) = \mathbf{g}(x,y) \ \text{in } \Omega$$

$\mathbf{u}(x,y)$ is the pressure field,

$\mathbf{k}(x,y)$ is the wave number,

$\mathbf{g}(x,y)$ is the point source function and

Ω is the domain. Absorbing boundary conditions are used on Γ.

$$\frac{\partial \mathbf{u}}{\partial n} - \iota \mathbf{u} = 0$$

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:

$$\begin{bmatrix} & -1 & \\ -1 & 4-k^2h^2 & -1 \\ & -1 & \end{bmatrix}$$

- 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 μ_i are the eigenvalues of the Laplace operator. Note that $\mu_1 - k^2$ may be negative.

Preconditioning (overview)

ILU	Meijerink and van der Vorst, 1977
ILU(tol)	Saad, 2003
SPAI	Grote and Huckle, 1997
Multigrid	Lahaye, 2001
	Elman, Ernst and O' Leary, 2001
AILU	Gander and Nataf, 2001
	analytic parabolic factorization
ILU-SV	Plessix and Mulder, 2003
	separation of variables

Preconditioning (Laplace type)

Laplace operator	Bayliss and Turkel, 1983
Definite Helmholtz	Laird, 2000
Shifted Laplace	Y.A. Erlangga, C. Vuik and C.W.Oosterlee, 2003

Shifted Laplace preconditioner (SLP)

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

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 Ω

Iterative solver: Bi-CGSTAB

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

k	ILU(0.01)	M_0	M_{-1}	M_i
5	9	13	13	13
10	25	29	28	22
15	47	114	45	26
20	82	354	85	34
30	139	> 1000	150	52

Spectrum of SLP

References: Manteuffel, Parter, 1990; Yserentant, 1988

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

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

The eigenvalues σ_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

$$\sigma_j = \frac{\lambda_j - z_1}{\lambda_j - z_2} \text{ holds.}$$

Spectrum of SLP

Theorem 2

If $\beta_2 = 0$, the eigenvalues $\sigma_r + \mathrm{i}\sigma_i$ 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 $\sigma_r + \mathrm{i}\sigma_i$ are on the circle in the complex plane with center c and radius R:

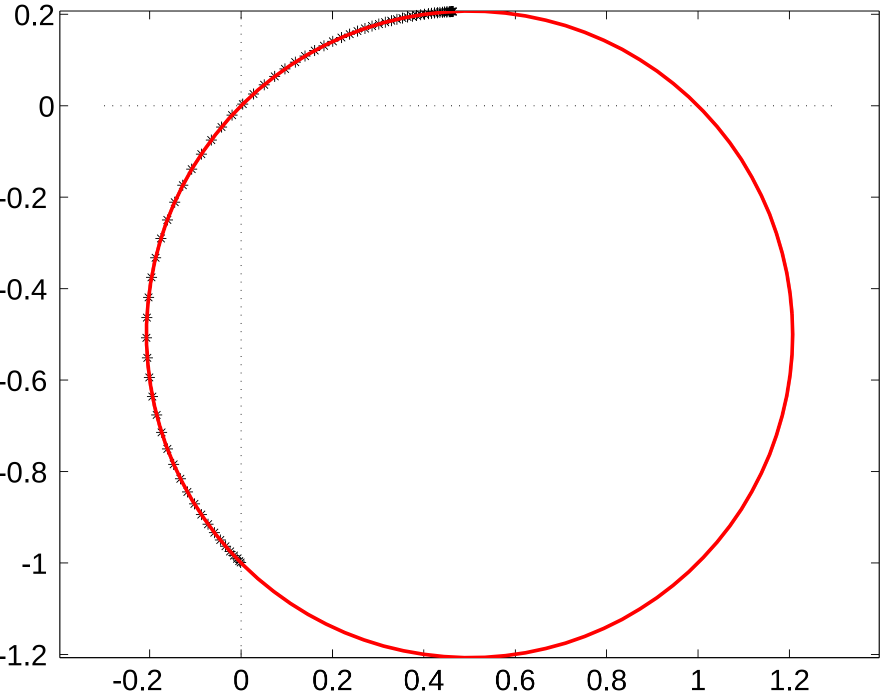
$$c = \frac{z_1 - \bar{z}_2}{z_2 - \bar{z}_2}, \quad R = \left| \frac{z_2 - z_1}{z_2 - \bar{z}_2} \right|.$$

Note that if $\beta_1 \beta_2 > 0$ the origin is not enclosed in the circle.

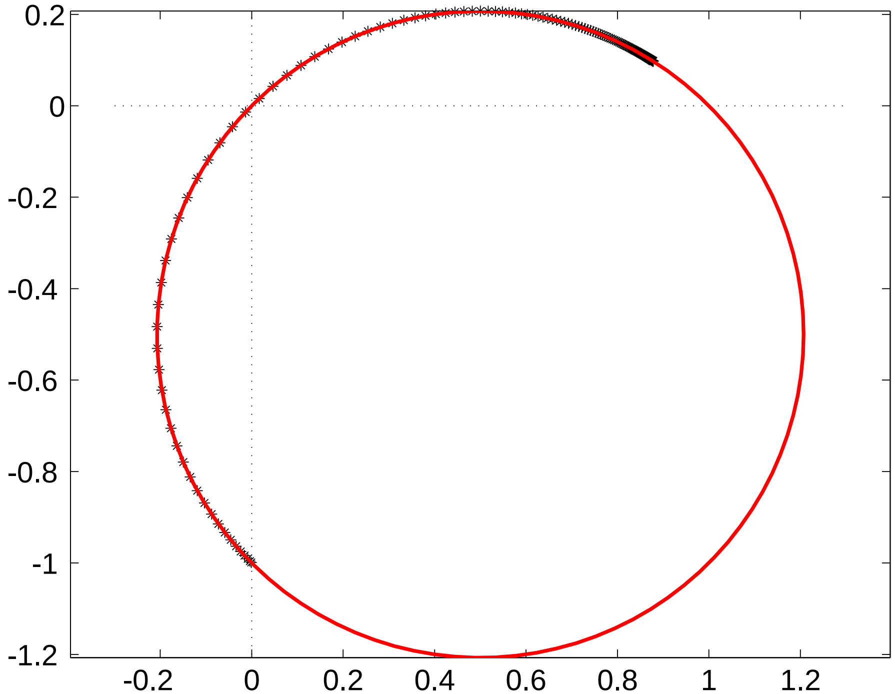
Eigenvalues for Complex preco $k = 100$

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

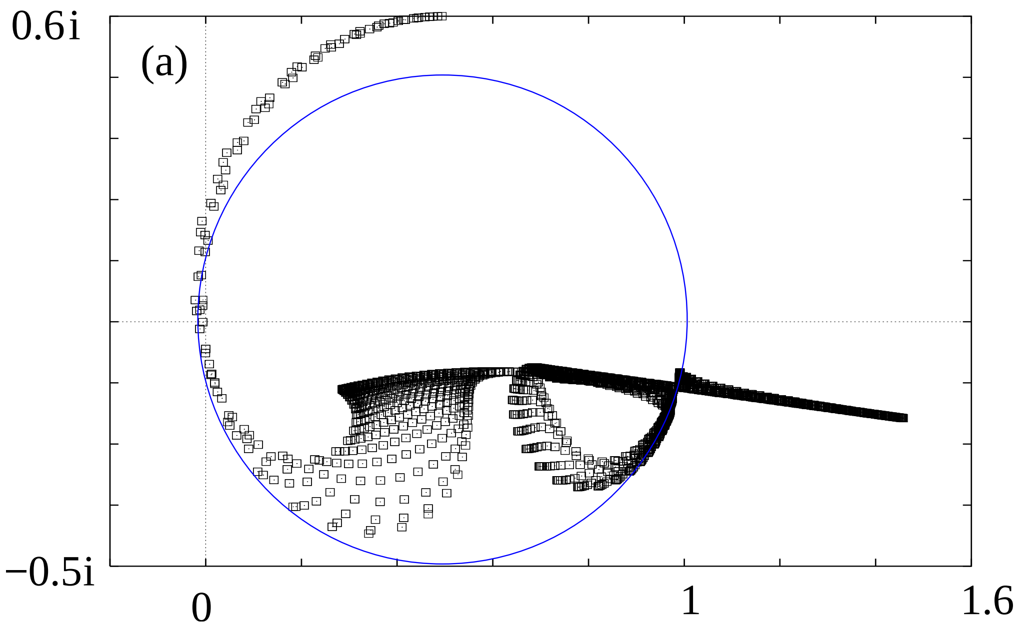
- Galerkin coarse grid approximation

Numerical results for a wedge problem

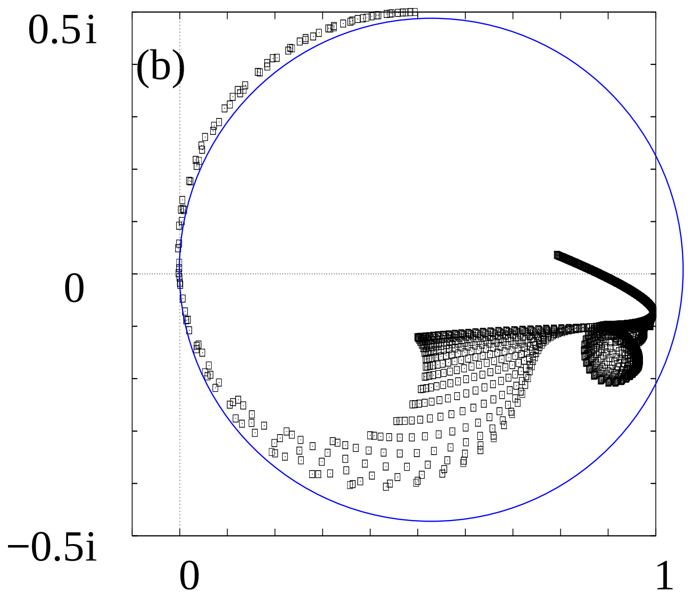
k_2	10	20	40	50	100
grid	32^2	64^2	128^2	192^2	384^2
No-Prec	201(0.56)	1028(12)	5170(316)	–	–
ILU(A,0)	55(0.36)	348(9)	1484(131)	2344(498)	–
ILU(A,1)	26(0.14)	126(4)	577(62)	894(207)	–
ILU(M,0)	57(0.29)	213(8)	1289(122)	2072(451)	–
ILU(M,1)	28(0.28)	116(4)	443(48)	763(191)	2021(1875)
MG(V(1,1))	13(0.21)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

Spectrum with inner iteration

1 MG iteration



2 MG iterations



4. Second Level Precond. (2008-2014)

Summary so far

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

$$M := -\Delta \mathbf{u} - (\beta_1 - \iota\beta_2)k^2\mathbf{u}$$

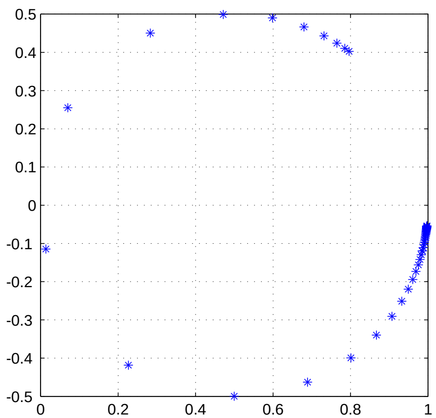
- Results show: $(\beta_1, \beta_2) = (1, 0.5)$ is the shift of choice
- Properties of SLP?

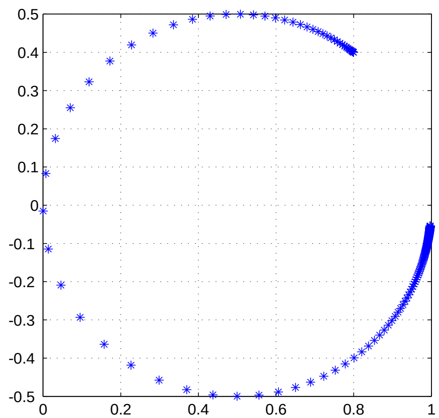
Shifted Laplace Preconditioner (SLP)

- 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.

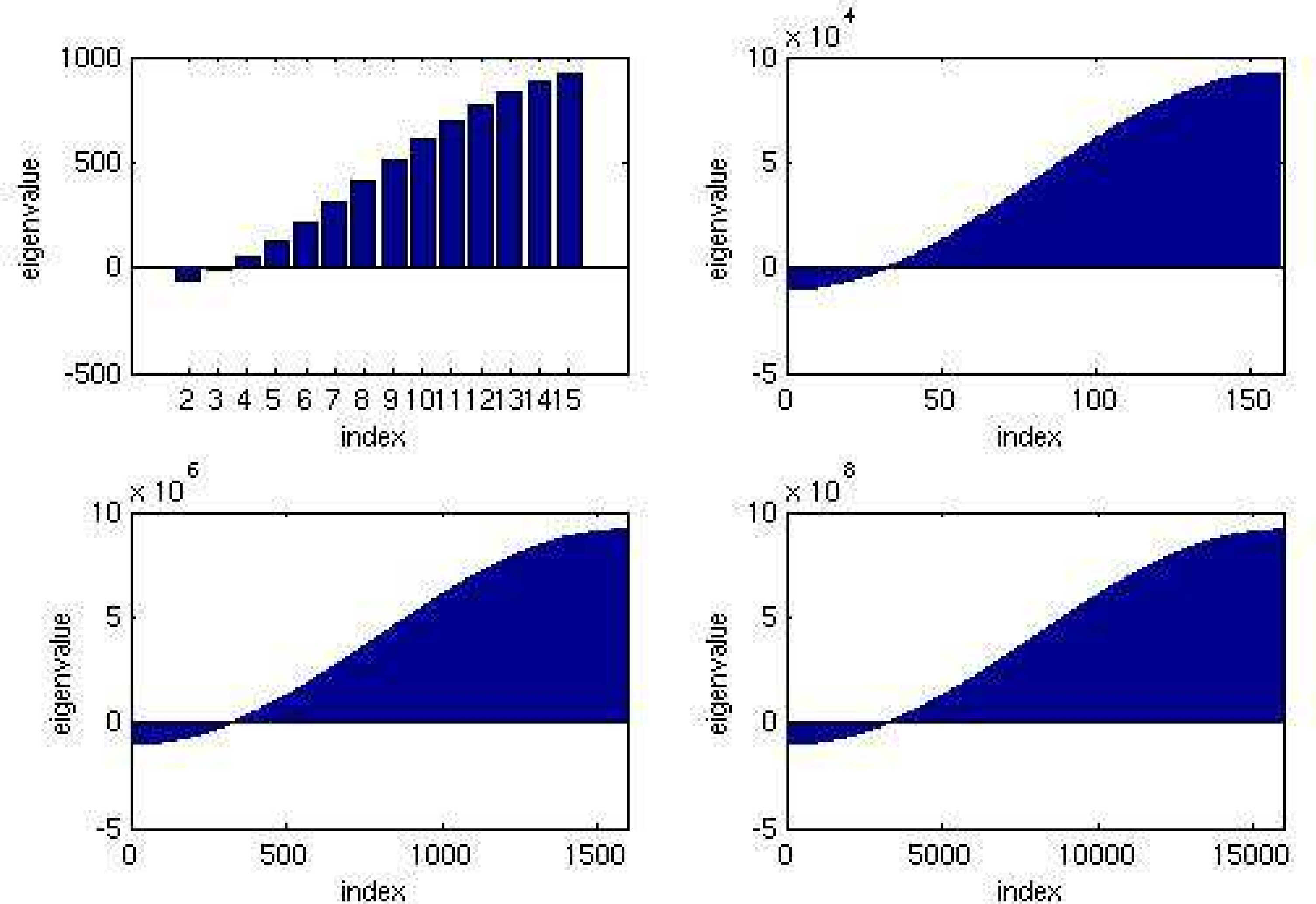
Spectrum of $M^{-1}(1, 0.5)A$ for

$k = 30$ and $k = 120$





Spectrum as function of k



Deflation: or two-grid method

Deflation, a projection preconditioner

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

where,

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

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. $Z = I_h^{2h}$ and $Z^T = I_{2h}^h$ then

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

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 λ_h^{max} by adding a shift in the two level preconditioner. This leads to the **ADEF1** preconditioner

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

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

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

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

Approximate A_{2h}

Ideal	**Practical**
$\hat{A}_{2h} = I_{2h}^h (M_h^{-1} A_h) I_h^{2h}$	$\hat{A}_{2h} = I_{2h}^h (M_h^{-1} A_h) I_h^{2h}$
	$\hat{A}_{2h} \approx I_{2h}^h I_h^{2h} M_{2h}^{-1} A_{2h}$

[a] Erlangga, Y.A and Nabben R., ETNA 2008

5. Fourier Analysis of two-level methods

Dirichlet boundary conditions for analysis.
With above deflation,

$$\textbf{spec}(PM^{-1}A) = f(\beta_1, \beta_2, k, h)$$

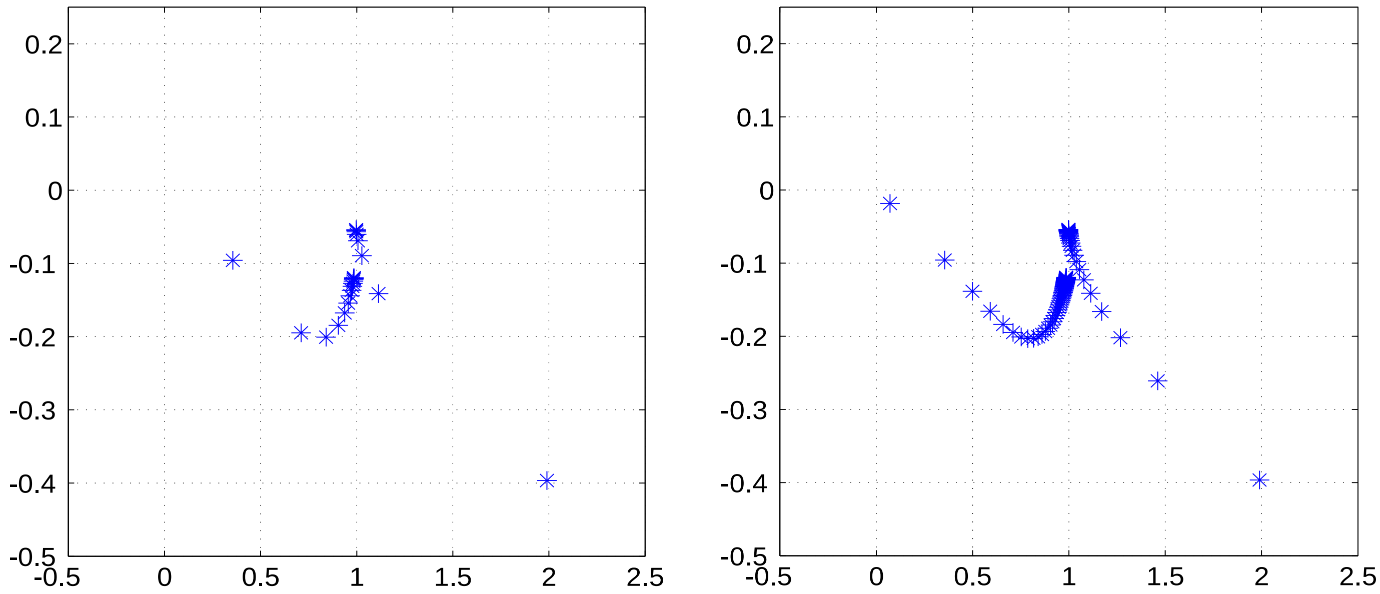
is a complex valued function.
Setting $kh = 0.625$,

- 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.

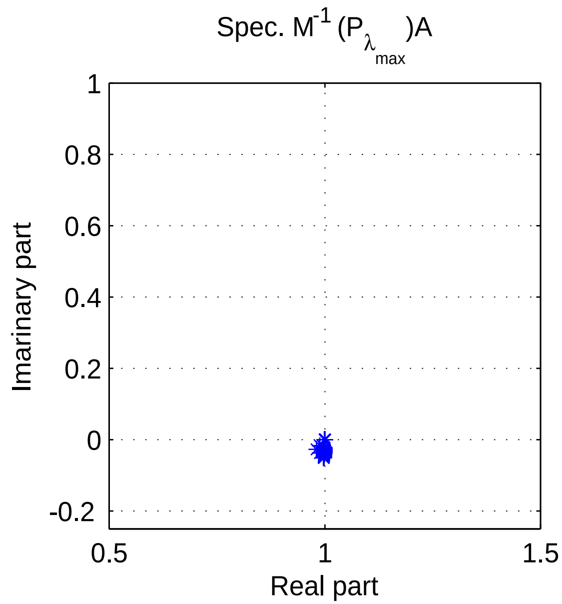
$k = 30$ $\qquad$ $k = 120$



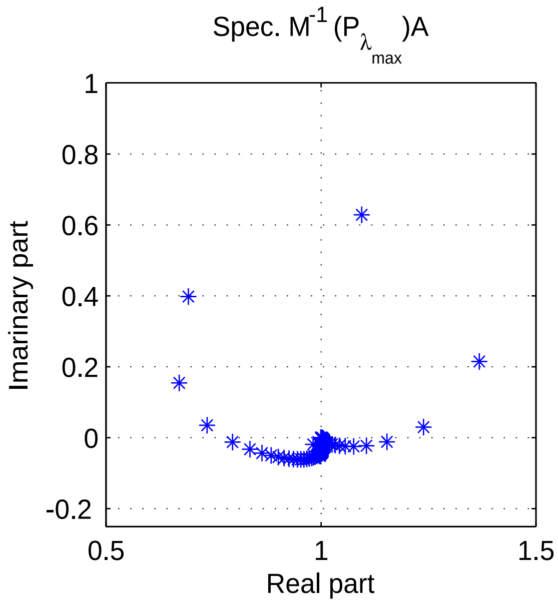
Fourier Analysis

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

Ideal



Practical



[b] Two-level

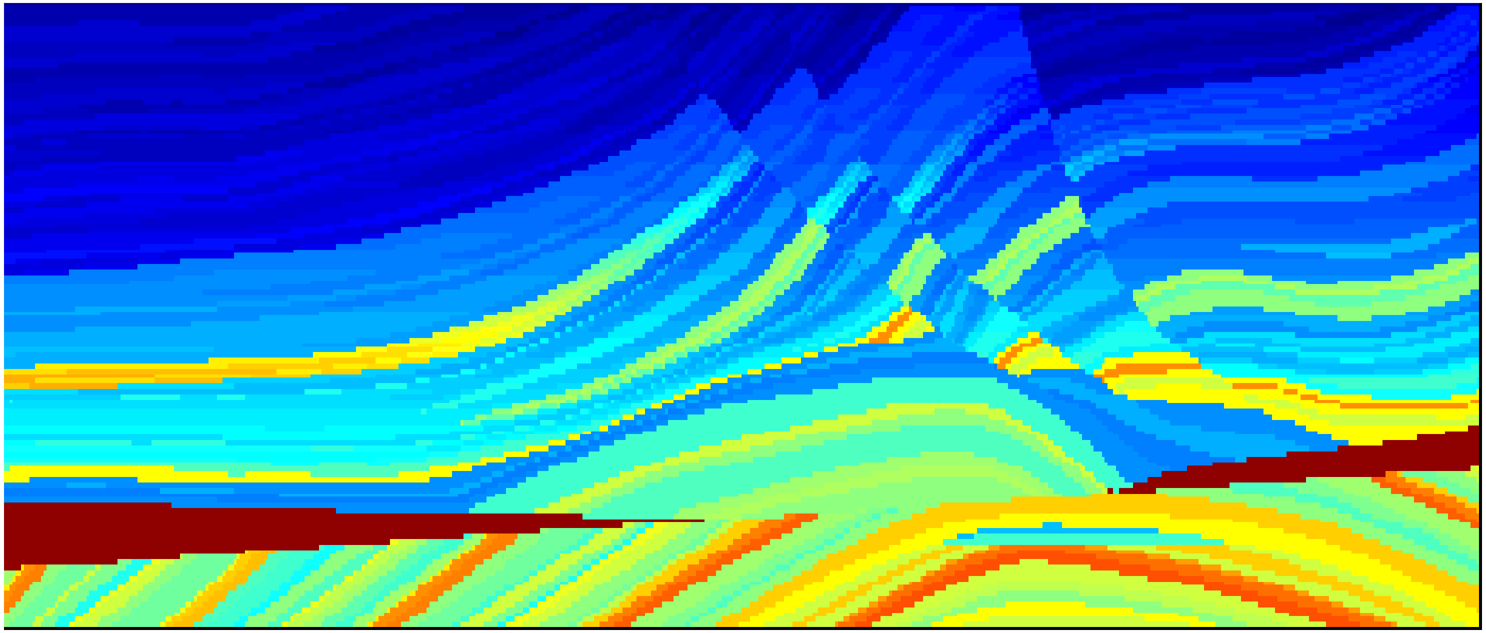
6. Numerical results (no pollution)

-

k	N($k^3h^2 \leq 0.625$)	iter	N($kh \leq 0.625$)	iter
10	44	4	16	4
20	116	4	32	5
40	320	4	64	5
80	800	4	128	6
100	1268	4	160	7
200	3572	4	320	8
400	10124	4	340	10
500	14144	4	800	12
800	28628	4	1280	15
1000	40004	4	1600	17

Application: geophysical survey

hard Marmousi Model



Application: geophysical survey

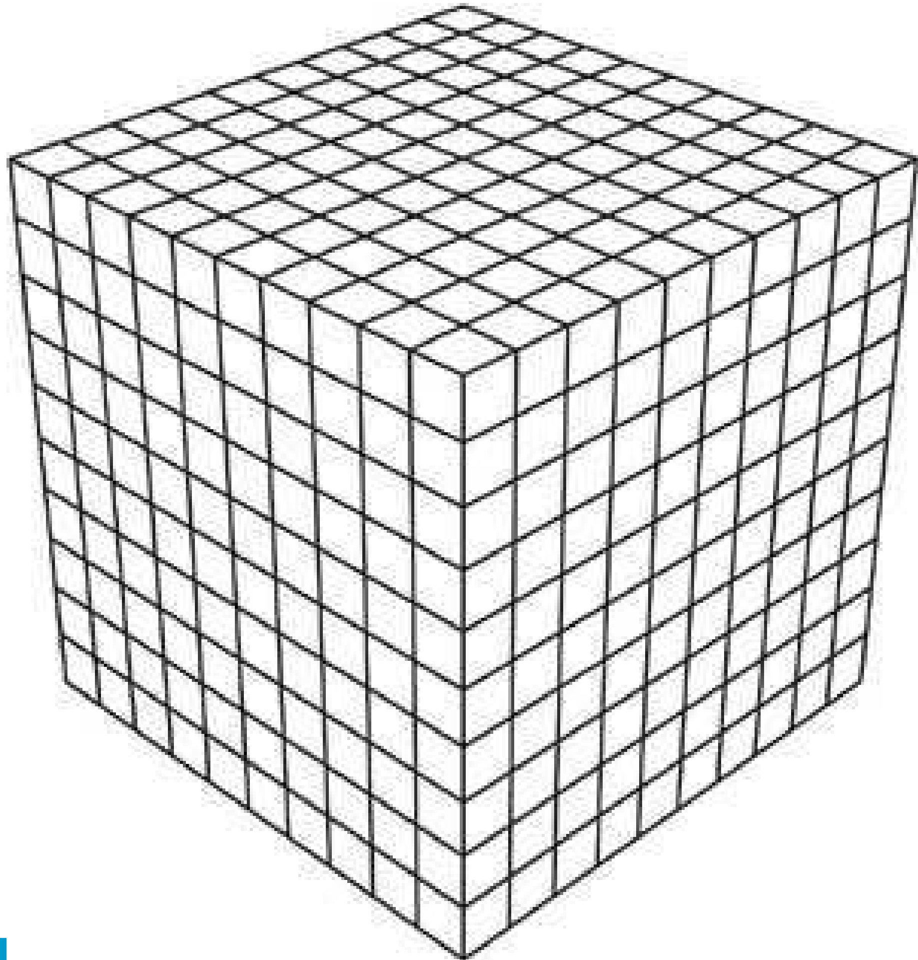
hard Marmousi Model, PETSc solver

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

Frequency f	Solve Time		Iterations	
	SLP-F	ADEF1-F	SLP-F	ADEF1-F
1	1.22	5.07	13	7
10	10.18	9.43	112	13
20	72.16	60.32	189	22
40	550.20	426.79	354	39

Application: geophysical survey

Cube with constant k



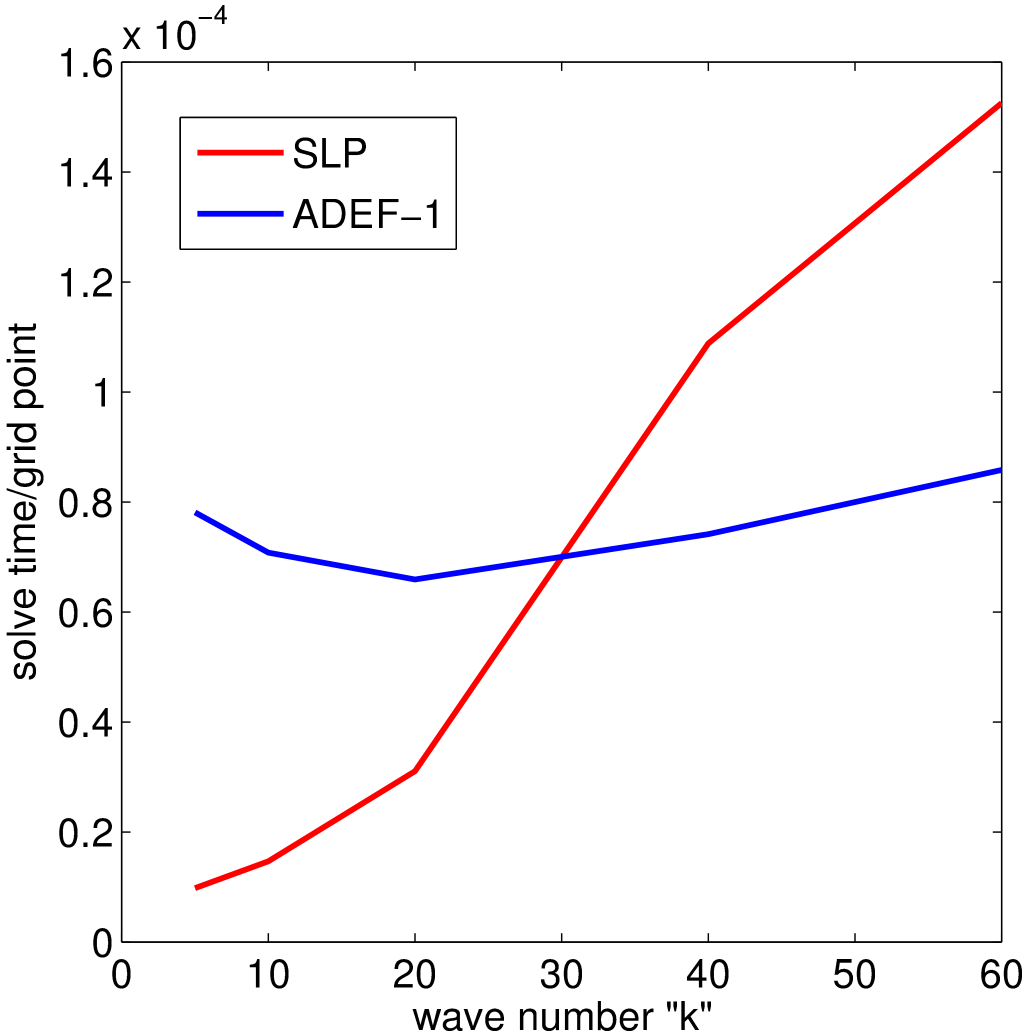
Application: geophysical survey

Cube with constant k

Wave number	Solve Time		Iterations	
k	SLP-F	ADEF1-F	SLP-F	ADEF1-F
5	0.04	0.32	7	8
10	0.48	2.32	9	9
20	8.14	17.28	20	9
40	228.29	155.52	70	10
60	1079.99	607.45	97	11

Application: geophysical survey

Cube with constant k



Application: geophysical survey

Cube with variable k

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

k	CLSP(time)	ADEF1(time)	CLSP	ADEF1
5	0.09	0.24	9	11
10	1.07	1.94	15	12
20	16.7	18.9	32	16
40	1304	214	331	24

Application: geophysical survey

Cube with variable k

Grid size h is such that $kh \approx 0.3125$

k	CLSP(time)	ADEF1(time)	CLSP	ADEF1
5	0.6	1.4	9	9
10	7.5	10.04	14	9
20	324.1	79.2	72	9
30	3810.9	361.7	285	11

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 is 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

- Y.A. Erlangga and C.W. Oosterlee and C. Vuik A Novel Multigrid Based Preconditioner For Heterogeneous Helmholtz Problems SIAM J. Sci. Comput.,27, pp. 1471-1492, 2006
- M.B. van Gijzen, Y.A. Erlangga and C. Vuik. Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian. SIAM J.of Sc. Comp. 2007.
- Y.A. Erlangga and R. Nabben. On a multilevel Krylov method for the Helmholtz equation preconditioned by shifted Laplacian. ETNA, 2008.
- H. Knibbe and C.W. Oosterlee and C. Vuik GPU implementation of a Helmholtz Krylov solver preconditioned by a shifted Laplace multi-grid method. Journal of Computational and Applied Mathematics, 236, pp. 281-293, 2011
- A.H. Sheikh, D. Lahaye and C. Vuik. On the convergence of shifted Laplace preconditioner combined with multi-grid deflation. Numerical Linear Algebra with Applications, 20, pp. 645-662, 2013
- http://ta.twi.tudelft.nl/nw/users/vuik/pub_it_helmholtz.html