Scalable Iterative Helmholtz Solvers

From Theory to Practice

Delft University of Technology

Vandana Dwarka, Kees Vuik*

January 5, 2022

Aim and Impact

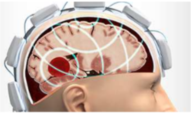
- Joint-work with PhD candidate Vandana Dwarka
- Contribute to broad research on Helmholtz solvers
- Understand inscalability (convergence)
- This presentation: improve convergence properties
  - Two-level methods
  - Multilevel methods (multigrid and deflation)

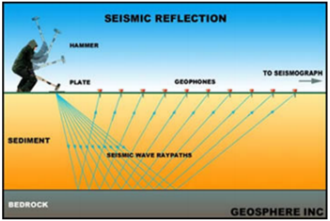
Introduction - The Helmholtz Equation

- Inhomogeneous Helmholtz equation + BC's

$$(-\nabla^2 - k^2)\, u(\mathbf{x}) = f(\mathbf{x}), \mathbf{x} \in \Omega \subseteq \mathbb{R}^n$$

- k is the wave number: $k = \frac{2\pi}{\lambda}$
- Practical applications in seismic/medical imaging and plasma fusion



Introduction - Numerical Model

- Start with analytical 1D model problem

$$-\frac{d^2u}{dx^2} - k^2\,u = \delta(x - \frac{1}{2}),$$
$$u(0) = 0, u(1) = 0,$$
$$x \in \Omega = [0,1] \subseteq \mathbb{R},$$

- Discretization using second-order FD with at least 10 gpw
- We obtain a linear system $A\hat{u} = f$

$$A = \frac{1}{h^2}\text{tridiag}[-1 \;\; 2 - (kh)^2 \;\; -1],$$

- A is real, symmetric, normal, indefinite and sparse
- Using Sommerfeld BC's A becomes non-Hermitian $\Rightarrow$ non-selfadjoint

Introduction - Challenges

- Negative & positive eigenvalues ⇒ limits Krylov based solvers
- Fast near-origin moving eigenvalues ⇒ slows convergence
  - CSLP (Helmholtz operator with complex shift)
  - Deflation + CSLP
  - Despite improvements problem remains
- Problems exacerbate in 2D & 3D and as k gets larger

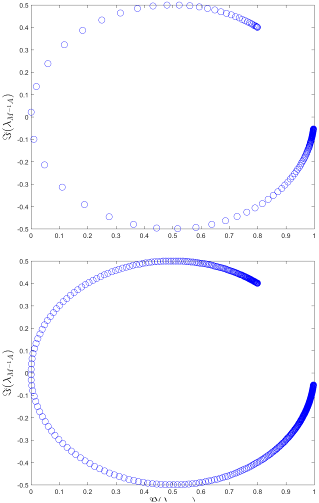
Preconditioning - CSLP

- Preconditioning to speed up convergence of Krylov subspace methods
- Solve $M^{-1}Au = M^{-1}f$, M is CSLP-preconditioner.

$$M = L - (\beta_1 - \beta_2 i)k^2 I,$$
$$(\beta_1, \beta_2) \in [0, 1]$$

- Increasing $k \Rightarrow$ eigenvalues move fast towards origin $\Rightarrow$ inscalable CSLP-solver
- Project unwanted eigenvalues onto zero = Deflation

Figure: $\sigma(M^{-1}A)$ for $k = 50$ (top) and $k = 150$ bottom.



Preconditioning - Deflation

- Projection principle: solve $PAu = Pf$

$$\tilde{P} = AQ \text{ where } Q = ZE^{-1}Z^T \text{ and } E = Z^TAZ,$$
$$P = I - \tilde{P},\ Z \in \mathbb{R}^{m \times n},\ m < n$$

- Columns of Z span deflation subspace
- Ideally Z contains eigenvectors
- In practice approximations: inter-grid vectors from multigrid
- Use DEF + CSLP combined $\Rightarrow$ spectral improvement

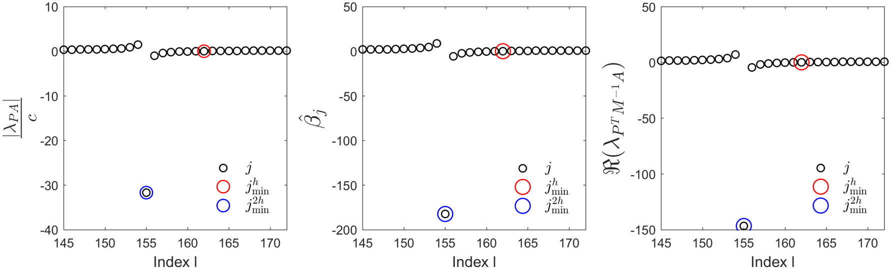
$$M^{-1}PAu = M^{-1}Pf$$

- Monitor eigenvalues using RFA (Dirichlet)

Preconditioning - Deflation

- Investigate near-null eigenvalue of all operators involved

Figure: $\lambda_j(PA), \beta^j, \lambda_j(P^TM^{-1}A)$ for $k = 500$



- Eigenvalues of PA and $P^TM^{-1}A$ behave like $\hat{\beta} = \frac{\lambda^l(A)}{\lambda^l(A_{2h})}$
- If near-kernel of A and A_{2h} misaligned $\Rightarrow$ near-null eigenvalues reappear!
- Equivalent to $j^h_{\min} \neq j^{2h}_{\min}$

Preconditioning - Deflation

- Recall: deflation space spanned by linear approximation basis vectors
- Transfer coarse-fine grid $\Rightarrow$ interpolation error
- Measure effect by projection error E
  $$E(kh) = \|(I - P)\phi_{j_{\min},h}\|^2,$$
  $$P = Z(Z^T Z)^{-1} Z^T$$

Figure: Restricted & interpolated eigenvectors (left $kh = 0.625$, right $k^3 h^2 = 0.625$

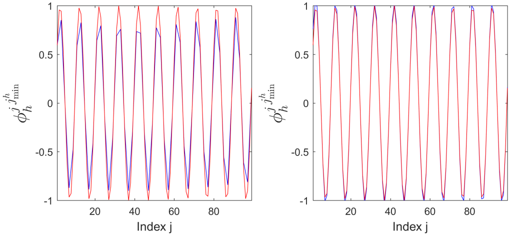


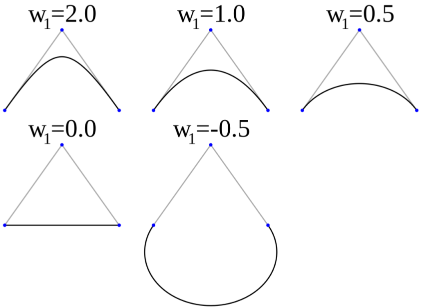
Table: Projection error DEF-scheme

k	$E(0.625)$	$E(0.3125)$
10^2	0.8818	0.1006
10^3	9.2941	1.0062
10^4	92.5772	10.0113
10^5	926.135	100.1382
10^6	9261.7129	1001.3818

Our Approach - Introduction

- Higher-order deflation vectors
- Rational quadratic Bezier curve $\Rightarrow$ one control-point
- Weight-parameter w to adjust control-point

Figure: Effect of changing weight



- w determined such that projection error minimized

Our Approach - Projection Error (1D)

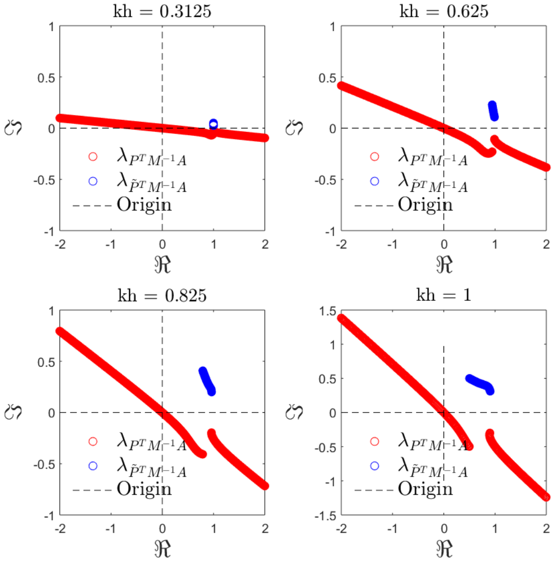
Table: Projection error $E(kh)$ for various w

k	$w = 0.1250$	$w = 0.0575$	$w = 0.01875$	$w = 0.00125$
	$kh = 1$	$kh = 0.825$	$kh = 0.625$	$kh = 0.3125$
10^2	0.0127	0.0075	0.0031	0.0006
10^3	0.0233	0.0095	0.0036	0.0007
10^4	0.0246	0.0095	0.0038	0.0007
10^5	0.0246	0.0095	0.0038	0.0007
10^6	0.0246	0.0095	0.0038	0.0007

- Weight-parameter w chosen to minimize projection error
- In all cases projection error *strictly* < 1
- RFA confirms favourable spectrum

Our Approach - Spectral Analysis (1D)

Figure: Spectrum of old (red) and new (blue) method for $k = 10^6$



Two-Level Deflation - 2D

Table: GMRES-iterations with tol $= 10^{-6}$ using Sommerfeld BC's and MG-approximation of CSLP(1,1). AD contains no CSLP.

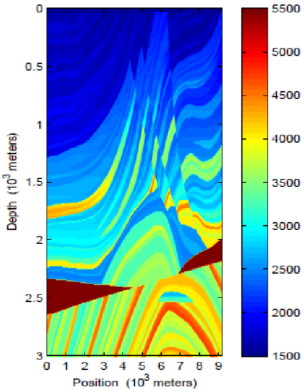
k	APD(0.1250)	APD(0.0575)	AD(0)
	$kh = 0.625$	$kh = 0.3125$	$kh = 0.3125$
100	4	4	3
250	5	4	4
500	5	5	5
750	7	5	5
1000	8	8	7

- DEF + CSLP needs 471 iterations for $k = 250$
- Close to wavenumber independence
- Weight-parameter w and CSLP less important as kh decreases

Two-Level Deflation - 2D Marmousi

Table: Solve time (s) and GMRES-iterations for 2D Marmousi

	DEF-TL	APD-TL	DEF-TL	APD-TL
	10 gpw			
f	Solve time (s)		Iterations	
1	1.72	4.08	3	4
10	7.20	3.94	16	6
20	77.34	19.85	31	6
40	1175.99	111.78	77	6
	20 gpw			
1	9.56	3.83	3	5
10	19.64	15.45	7	5
20	155.70	122.85	10	5
40	1500.09	1201.45	15	5



Two-Level Deflation - 3D

Table: GMRES-iterations with tol $= 10^{-6}$ using Sommerfeld BC's and MG-approximation of CSLP(1,1). AD contains no CSLP.

k	APD(0.125)	AD(0)
	Iterations	Iterations
10	4	4
25	4	5
50	4	5
75	4	5

- DEF + CSLP takes 66 iterations for $k = 40$
- Wavenumber independent convergence
- Two-level method memory $\Rightarrow$ multilevel methods

Multilevel methods

Multilevel Deflation

- Pros

  Close to linear complexity

  Memory efficient

  Recursive structure

  Use as preconditioner with FGMREs
- Cons

  Needs more inner cycles

  Convergence depends weakly on k

Multigrid

- Pros

  Linear complexity

  Memory efficient

  Recursive structure

  Use as stand-alone or preconditioner
- Cons

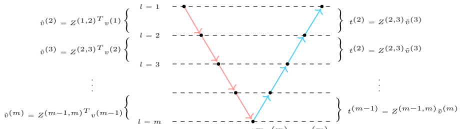
  Diverges for Helmholtz

  Slow convergence for small k

New research on convergent multigrid solver!

Multilevel Deflation

- Apply two-level method recursively
- Only 1 FGMRES it. per level



- Krylov 'smoother' vs Multigrid
- 10 iterations on indefinite levels
- 1 Jacobi iteration on all others
- Reduce time and memory

Algorithm 3.1 Two-level Deflation FGMRES

Initialization:
Choose u_0 and dimension k of the Krylov subspaces.
Define $(k+1) \times k$ $\bar{H}_k$ and initialize to zero.
Arnoldi process: $r_0 = f - Au_0$, $\beta = \|r_0\|_2$, $v_1 = r_0/\beta$.
for $j = 1, 2, \ldots k$ **do**
  $\hat{v} = Z^T v_j$
  $\tilde{v} = E^{-1}\hat{v}$
  $t = Z\tilde{v}$
  $s = At$
  $\tilde{r} = v_j - s$
  $r = M^{-1}\tilde{r}$
  $x_j = r + t$
  $w = Ax_j$
  for $i = 1, 2, \ldots j$ **do**
    $h_{i,j} = (w, v_j)$ $w = w - h_{i,j} v_i$
  end
  Compute $h_{j+1,j} = \|w\|_2$ and $v_{j+1} = w/h_{j+1,j}$.
  Define $X_k = [x_1, x_2, \ldots, x_k]$
  $\bar{H}_k = \{h_{i,j}\}_{1 \leqslant i \leqslant j+1, 1 \leqslant j \leqslant k}$
end
Form approximate solution:
Compute $u_k = u_0 + X_k y_k$ where $y_k = \arg\min_y \|\beta e_1 - \bar{H}_k y\|_2$.
Restart:
If satisfied stop, else set $u_0 \leftarrow u_k$ and repeat Arnoldi process.

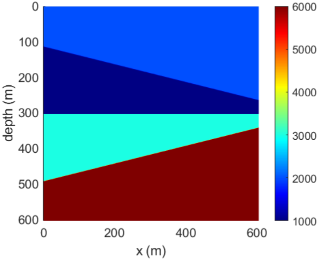
Multilevel Deflation - 3D

Table: Number of outer FGMRES-iterations for $kh = 0.625$. Column 1 quadratic, column 2 linear deflation vectors.

k	APD	DEF
	Iterations	Iterations
10	9	11
20	9	12
40	11	17
80	14	45

- Both methods benefit from multilevel implementation
- Reduced time and memory
- Convergence APD slightly depends on wavenumber
- What about heterogeneous models?

Multilevel Deflation - 2D Wedge



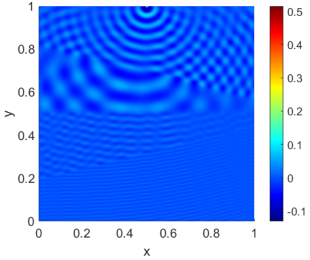
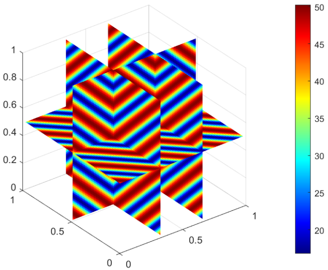


Table: Number of outer FGMRES-iterations for $kh = 0.625$

$\mathbf{k = 2\pi f}$	n	$c(x,y) \in [500, 3000]$ m/s		$c(x,y) \in [1000, 6000]$ m/s	
f (Hz)		Iterations	CPU(s)	Iterations	CPU(s)
10	10.201	9	0.428	9	0.598
20	41.209	11	2.112	14	11.148
40	162.409	17	47.080	19	86.171
60	366.025	21	157.143	22	325.960
80	648.025	23	459.561	25	774.926

Multilevel Deflation - 3D Sine



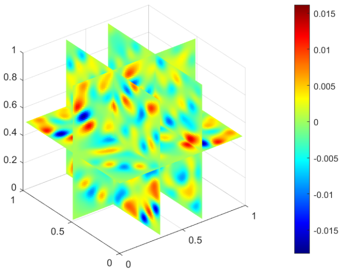


Table: Number of outer FGMRES-iterations for $kh = 0.625$

		$k(\mathbf{x})^2 = \alpha + \beta \sin(8\pi\mathbf{x}), \alpha = 0.5(k_1^2 + k_2^2), \beta = 0.5\lvert k_2^2 - k_1^2\rvert$			
		$\gamma = 1$		$\gamma = 2$	
$[k_1, k_2]$	n	Iterations	CPU(s)	Iterations	CPU(s)
$[8, 25]$	68.921	8	3.041	6	4.026
$[16, 50]$	531.441	26	133.688	15	123.218
$[24, 75]$	1.771.561	49	1095.185	28	959.926

Multilevel Deflation - 3D Elastic Wave

- Coupled vector equations for time-harmonic
- Wedge domain
- 20 gpw (grid points per wavelength)

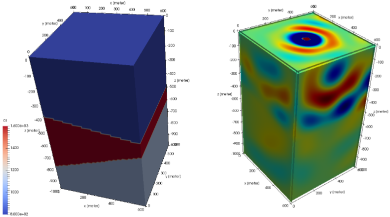


Table: Number of outer FGMRES-iterations.

$k = 2\pi f$	n	$\gamma = 1$		$\gamma = 2$	
$f(Hz)$		Iterations	CPU(s)	Iterations	CPU(s)
0.5	3.159	9	0.428	9	0.598
2	147.033	20	69.214	18	69.971
4	1.127.463	25	1174.589	22	937.294

Multigrid

- Standard multigrid diverges for small k
- Open problem in math for 30 years
- But, convergence if:

  Higher-order prolongation/restriction

  Coarsening on CSLP instead of original Helmholtz operator
- Small number of smoothing steps using $\omega-$Jacobi
- GMRES(3) smoothing gives a fast solver
- Works for both $V-$ and $W-$cycles

Multigrid - 2D

- Constant k using Sommerfeld BC

Table: Number of V- ($\gamma = 1$) and W-cycles ($\gamma = 2$) for constant k using tol. 10^{-5}. ν denotes the number of ω-Jacobi smoothing steps.

	$k = 50$		$k = 100$		$k = 150$		$k = 200$		$k = 250$	
	$N = 6724$		$N = 26244$		$N = 57600$		$N = 102400$		$N = 160000$	
	$N_D = 8$		$N_D = 8$		$N_D = 4$		$N_D = 8$		$N_D = 4$	
γ	1	2	1	2	1	2	1	2	1	2
$\nu = 4$	58	58	104	108	155	159	209	213	267	271
$\nu = 5$	58	58	104	104	150	166	194	229	238	287
$\nu = 6$	55	58	99	102	139	167	183	222	226	283
$\nu = 7$	53	60	97	101	136	163	179	219	221	280
$\nu = 8$	53	60	95	104	131	161	178	212	218	277

- Coarsening on CSL (shift $= 0.7$)
- No level-dependent parameters!
- Linear interpolation diverges ($k = 50, \gamma = 1$)
- What about GMRES(3) smoothing?

Multigrid - 2D

- Constant wave number using Sommerfeld BC

Table: Number of V- ($\gamma = 1$) and W-cycles ($\gamma = 2$) for constant k using tol. 10^{-5}. ν denotes the number of GMRES(3) smoothing steps.

	$k = 50$		$k = 100$		$k = 150$		$k = 200$		$k = 250$	
	$N = 6724$ $N_D = 8$		$N = 26244$ $N_D = 8$		$N = 57600$ $N_D = 4$		$N = 102400$ $N_D = 8$		$N = 160000$ $N_D = 4$	
γ	1	2	1	2	1	2	1	2	1	2
$\nu = 1$	14	7	24	10	39	19	51	24	64	29
$\nu = 2$	8	5	13	7	22	10	28	13	34	16
$\nu = 3$	6	5	10	6	16	9	20	10	24	12
$\nu = 4$	6	5	8	5	12	7	15	9	18	10
$\nu = 5$	5	5	7	5	11	7	13	8	15	9

- Coarsening + on CSL (shift = k^{-1})
- Iteration count with $\gamma = 2$ close to k-independent
- Linear interpolation 199 iterations ($k = 50, \gamma = 1$)
- What about heterogeneous problems?

Multigrid - 2D random k (high-contrast)

Figure: $k(x, y)$

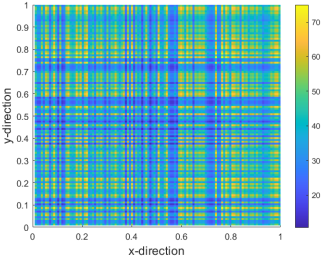


Figure: $u(x, y)$

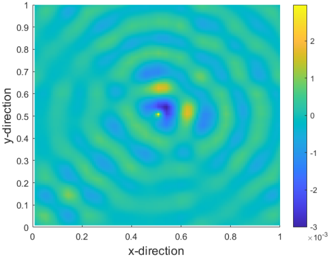


Table: Number of V- ($\gamma = 1$) and W-cycles ($\gamma = 2$) with tol 10^{-5}.
ν denotes the number of ω-Jacobi smoothing steps.

	$(k_1, k_2) = (10, 50)$		$(k_1, k_2) = (10, 75)$	
γ	1	2	1	2
$\nu = 4$	102	96	111	107
$\nu = 5$	97	95	103	105
$\nu = 6$	95	95	101	104
$\nu = 7$	94	94	102	104
$\nu = 8$	94	94	102	104

Multigrid - 2D random k (high-contrast)

Figure: $k(x, y)$

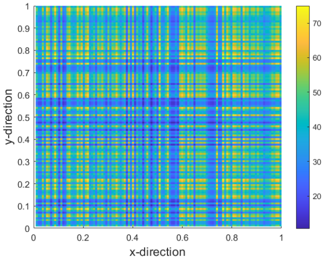


Figure: $u(x, y)$

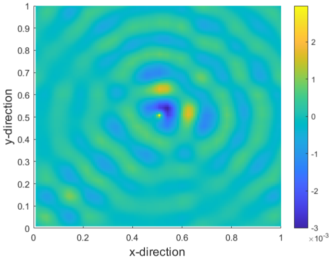


Table: Number of V- ($\gamma = 1$) and W-cycles ($\gamma = 2$) with tol 10^{-5}.
ν denotes the number of GMRES(3) smoothing steps.

	$(k_1, k_2) = (10, 50)$		$(k_1, k_2) = (10, 75)$	
γ	1	2	1	2
$\nu = 1$	28	12	31	12
$\nu = 2$	16	8	17	7
$\nu = 3$	12	7	12	6
$\nu = 4$	10	6	10	6
$\nu = 5$	9	6	9	6

Conclusion

- Deflation projects unwanted eigenmodes to zero
- Misalignment of near-zero eigenvalues affects convergence
- New deflation scheme: higher-order approximation
- Two-level method wavenumber independent convergence but memory constrained
- Use higher-order scheme in multilevel methods
  1. Multilevel deflation (with FGMRES)
  2. Multigrid (preconditioner or stand-alone solver)
- Upcoming work: research on interpolation schemes and large-scale applications using parallel computing

References

- Upcoming articles: multilevel deflation and multigrid methods. Reports available at: `http://ta.twi.tudelft.nl/users/vuik//pub_it_helmholtz.html`
- Further reading

V. Dwarka, C. Vuik.
Scalable Convergence Using Two-Level Deflation Preconditioning for the Helmholtz Equation
SIAM Journal on Scientific Computing, 42(3):A901–A928, 2020.

V. Dwarka, R. Tielen, M. Moller and C. Vuik
Towards Accuracy and Scalability: Combining Isogeometric Analysis with Deflation to Obtain Scalable Convergence for the Helmholtz Equation
Computer Methods in Applied Mechanics and Engineering, 377:113694, 2021.

V. Dwarka and C. Vuik
Pollution and Accuracy of solutions of the Helmholtz Equation: A Novel Perspective from the Eigenvalues
Journal of Computational and Applied Mathematics, 395:113549, 2021.