Application of matrix square root and its inverse to downward wavefield extrapolation

Polina Zheglova, Felix J. Herrmann

Seismic Laboratory for Imaging and Modeling
University of British Columbia

June 18, 2014

Motivation

- Downward wavefield extrapolation
- Matrix functions
- Low rank matrix compression (HSS)
- Combine to explore and develop efficient algorithms for modeling/imaging

Introduction

- In downward extrapolation, the goal is to solve Helmholtz equation

$$\frac{\partial^2 p(\mathbf{x}, z, \omega)}{\partial z^2} = -\left(\frac{\omega^2}{c^2(\mathbf{x}, z)} + \nabla_{\mathbf{x}}^2\right) p(\mathbf{x}, z, \omega) \tag{1}$$

  by stepping in depth from the boundary data $p(\mathbf{x}, z = z_0, \omega)$.
- Main advantage: reduction in dimensionality of extrapolation problem
- Main difficulty: evanescent modes

Introduction

Full wave equation depth extrapolation (Sandberg & Beylkin, 2009; Sandberg, Beylkin & Vassiliou, 2010)

- Operator $\mathcal{H}_2 = \frac{\omega^2}{c^2(\mathbf{x},z)} + \nabla^2_{\mathbf{x}}$ is projected to its non-negative invariant subspace:

$$\mathcal{H}_2 \to \mathcal{P}\mathcal{H}_2\mathcal{P}$$

- Downward extrapolation equation:

$$\frac{\partial^2 p(\mathbf{x}, z, \omega)}{\partial z^2} = -\mathcal{P}\mathcal{H}_2\mathcal{P}p(\mathbf{x}, z, \omega) \quad (2)$$

- Spectral projector is is computed by:

$$\mathcal{P} = \frac{1}{2}(I + \mathsf{sign}(\mathcal{H}_2))$$

where $\mathsf{sign}(\mathcal{H}_2)$ is found by recursion (e.g. Kenney & Laub, 1995)

$$S_0 = \frac{\mathcal{H}_2}{\|\mathcal{H}_2\|_2}, \qquad S_{k+1} = \frac{3}{2}S_k - \frac{1}{2}S_k^3$$

- Efficiency is achieved by low rank matrix compression (PLR, HSS), estimated cost $\sim O(N)$

Introduction

One way wave equation:

- Similar approach can be used in one way wave equation extrapolation
- Square root operator $\mathcal{H}_1 = \mathcal{H}_2^{1/2}$ can be computed by polynomial recursion
- Filtering of evanescent waves is still necessary
- Modeling of all propagating modes is possible
- Efficiency for large problems with matrix compression

Other uses:

- Correct modeling of a volume injection (e.g. air gun) source and scattering operators **(e.g. Wapenaar, 1990)**
- These require computation of inverse square root $\mathcal{H}_2^{-1/2}$

One way wave equation

- The one way wave equation is obtained by factoring the operator $\mathcal{H}_2 = \frac{\omega^2}{c^2(\mathbf{x},z)} + \nabla^2_{\mathbf{x}}$, and then neglecting the terms that account for the scattering (e.g. Grimbergen et al., 1998; Wapenaar, 1990):

$$\frac{\partial p^{\pm}}{\partial z} = \mp i \mathcal{H}_1 p^{\pm}$$

where

p^+, p^- - down and up going fields: $p = p^+ + p^-$,
$\mathcal{H}_1$ - propagator, $\mathcal{H}_1 \mathcal{H}_1 p = \mathcal{H}_2 p$.

- Extrapolation is done by finite differences or matrix exponentiation by scaling and squaring algorithm

One way wave equation

- $\mathcal{H}_1 = \mathcal{H}_2^{1/2}$ is a non-local pseudo-differential operator
- Approximate square root by a polynomial or rational function $\Rightarrow$ paraxial wave equation
  - Efficient with finite differences and operator splitting
  - Propagating modes up to certain angle from the main propagation direction
- Modal decomposition of the discretized operator H_2 (e.g. Grimbergen et al.; Margrave et al., 2002; Lin & Herrmann, 2007)
  - Discretize $\mathcal{H}_2 \to H_2$ by finite differences
  - All propagating modes in the main propagation direction
  - Requires eigenvalue decomposition, not practical for large problems
- Our goal: use polynomial recursion with matrix compression instead of modal decomposition

Square root calculation

- Assume:
  - Absorbing boundary conditions in x are decoupled, H_2 is self-adjoint
  - Negative eigenvalues have been removed by spectral projector: $\tilde{H}_2 = \mathcal{P} H_2 \mathcal{P}$ - no evanescent modes
- Principal root of matrix H with no nonpositive eigenvalues can be computed by Shultz iteration (Higham, 2008):

$$Y_0 = \frac{H}{\|H\|_2}, Z_0 = I$$
$$Y_{k+1} = \frac{3}{2} Y_k - \frac{1}{2} Y_k Z_k Y_k$$
$$Z_{k+1} = \frac{3}{2} Z_k - \frac{1}{2} Z_k Y_k Z_k$$

- Derived by applying polynomial recursion for matrix sign function to $\begin{bmatrix} 0 & H \\ I & 0 \end{bmatrix}$ and Newton's method
- $Y_k \longrightarrow \left(\frac{H}{\|H\|_2}\right)^{1/2}$, $Z_k \longrightarrow \left(\frac{H}{\|H\|_2}\right)^{-1/2}$ quadratically

Square root calculation

- The square root polynomial recursion is poorly conditioned since $\tilde{H}_2$ has zeros eigenvalues (numerically they are very small complex numbers)
- Shultz iteration applied to $\tilde{H}_2$: $Y_k \longrightarrow \left(\frac{\tilde{H}_2}{\|\tilde{H}_2\|_2}\right)^{1/2}$ in $\sim O(10)$ with high accuracy.
- Z_k part causes the iteration eventually to diverge, since $\tilde{H}_2$ does not have an inverse $\Rightarrow$ careful stopping criterion is needed
- Stopping criterion we use: (a) difference between iterates $\|Y_{k+1} - Y_k\|$, (b) misfit $\|\tilde{H}_1^2 - \tilde{H}_2\|$, (c) update direction

Computation of pseudo inverse square root

Pseudo inverse $\tilde{H}_1^\dagger$ is needed to implement volume injection source as a boundary condition:

$$S^{\pm}(\mathbf{x}, z = 0, \omega) = \frac{i\omega^2}{2} \tilde{H}_1^\dagger I(\mathbf{x}, z = z_0, \omega)$$

(e.g. Wapenaar, 1990)

To compute pseudo inverse $\tilde{H}_1^\dagger$:

- Compute ${H_2}^{-1}$, e.g. by recursion (Ben Israel and Cohen, 1966): well conditioned, stable, quadratic convergence, known to be slow initially
- Apply spectral projector to ${H_2}^{-1}$: ${H_2}^{-1} \rightarrow \tilde{H}_2^\dagger = \mathcal{P} {H_2}^{-1} \mathcal{P}$
- Compute pseudo inverse of $\tilde{H}_1^\dagger$ by Shultz iteration from $\tilde{H}_2^\dagger$

Evanescent modes are discarded in all calculations.

Convergence

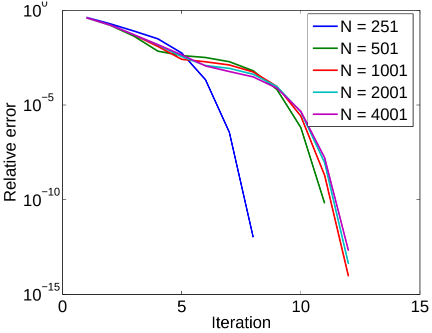


Figure 1: Convergence of iteration for the square root, relative error $= \frac{\|\tilde{H}_1^2 - \tilde{H}_2\|}{\|\tilde{H}_2\|}$

Convergence

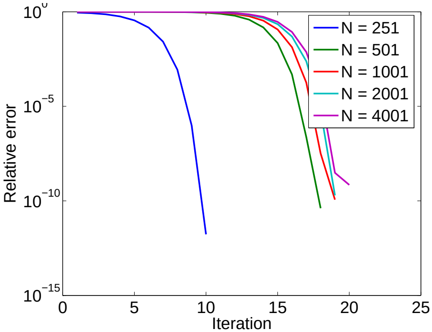


Figure 2: Convergence of iteration for the matrix inverse, relative error $= \frac{\|H_2^{-1} - \mathrm{pinv}(H_2)\|}{\|\mathrm{pinv}(H_2)\|}$

Convergence

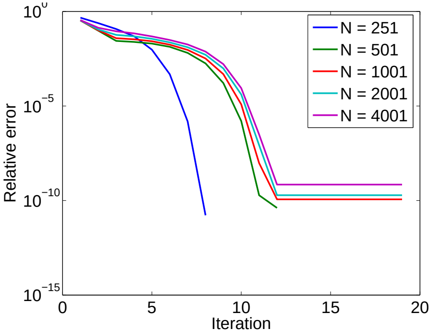
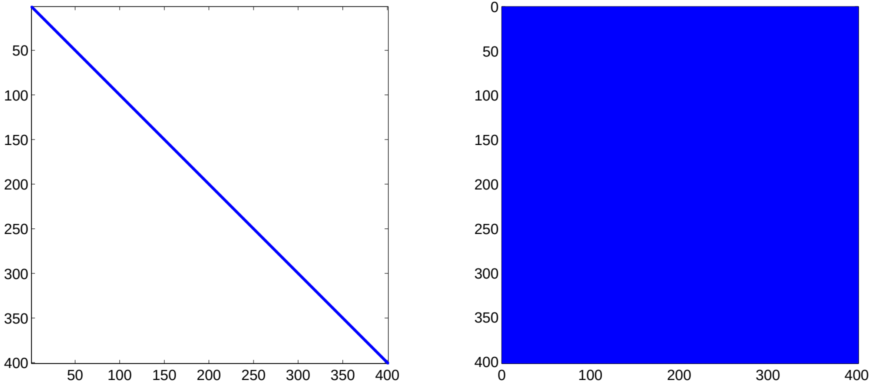


Figure 3: Convergence of iteration for the pseudo inverse of square root, relative error $= \frac{\|\tilde{H}_1^{\dagger 2} - \tilde{H}_2^{\dagger}\|}{\|\tilde{H}_2^{\dagger}\|}$

Compression



(a) 0.75% of non-zeros (b) 100% of non-zeros

Figure 4: Structure of example H_2 matrix and its inverse: 1-d case, simple finite differences

Compression

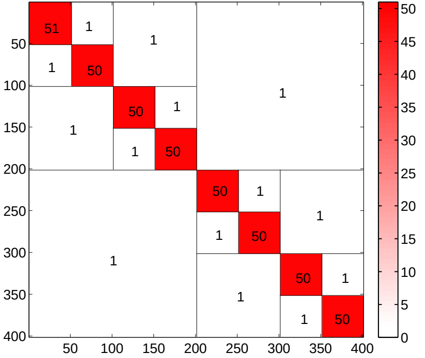


Figure 5: Rank representation of the inverse

- Diagonal blocks have full rank
- Off-diagonal blocks have rank 1

Compression

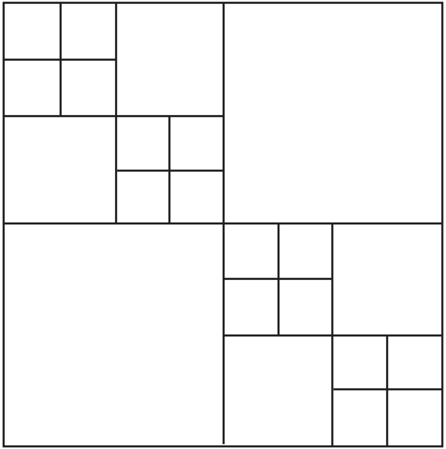
Matrices that have HSS structure

- Have large blocks with low numerical rank
- Often arise in solutions of PDEs, e.g. integral operators:

$$u(x) = \int K(x, y) f(y) dy,$$

  where $K(x, y)$ decays fast away from $x = y$ or is smooth
- Discretized Helmholtz operator (and functions of thereof) have HSS structure. This has been proven for some functions **(e.g. Beylkin et al., 1999 - sign function)**
- Seismic data being the Green's function can also be represented with HSS **(Kumar et al., 2013)**

Compression



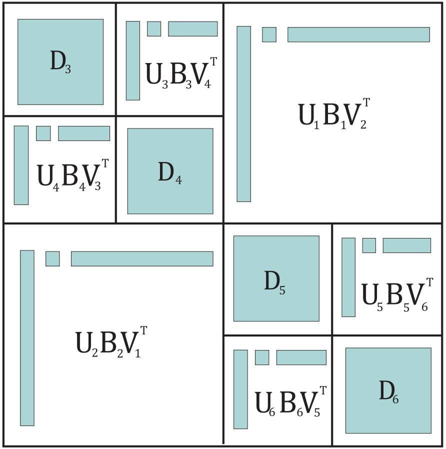
Xia, 2012, Lyons, 2005

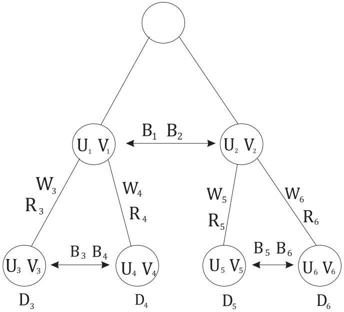
- Off-diagonal blocks have low numerical rank
- Each low rank approximation is a product of
  - a tall matrix
  - a small matrix and
  - a thin matrix
- The hierarchy is organized in a binary tree

Compression

Hierarchically semiseparable (HSS) representation of matrices allows us to

- Store dense matrices with less memory
- Do matrix operations - multiplication, addition, scaling, etc. - fast (e.g. $O(n)$ vs $O(n^3)$ flops)
- Results are also HSS matrices
- Approximate but can be made arbitrarily accurate by increasing the rank of the block approximants



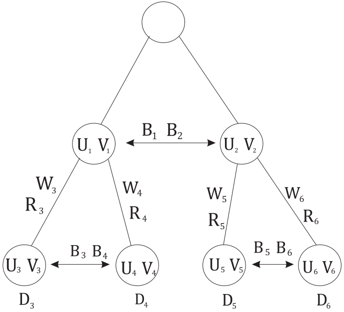


Xia, 2012, Lyons, 2005

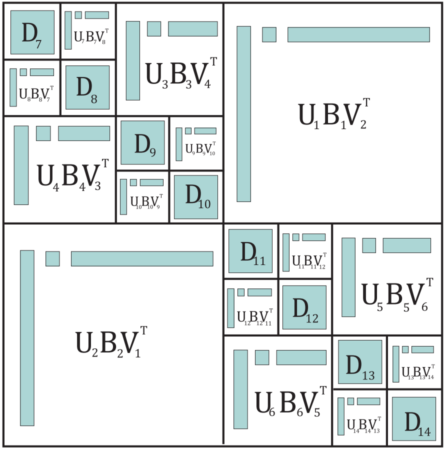
Compression

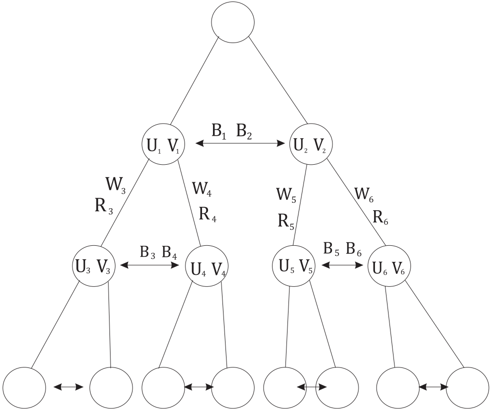
- Store only lowest U's and V's in the hierarchy
- Store B's, R's, W's for each level - small matrices, much smaller than U's and V's
- Higher U's and V's are determined from lower U's and V's via R's and W's
- Store the lowest D's in the hierarchy as dense matrices
- Optimized for matrix-vector multiplication

Xia, 2012, Lyons, 2005



Compression





Xia, 2012, Lyons, 2005

Compression

Complexity of algorithms for HSS matrix operations (Sheng et al., 2007):

Operation	Cost with HSS	Cost without HSS
Matrix-vector mutiplication	$O(nr^2)$	$O(n^2)$
Matrix-matrix multiplication	$O(nr^3)$	$O(n^3)$
Matrix addition	$O(nr^2)$	$O(n^2)$
LU decomposition	$O(nr^3)$	$O(n^3)$
Matrix inverse	$O(nr^3)$	$O(n^3)$
Transpose	$O(nr)$	$O(n^2)$
HSS construction	$O(nr)$	Not applicable

- r is maximum rank of off-diagonal blocks
- Efficient implementation is non-trivial
- Current implementation in Matlab (MSN toolbox and Lina Miao)

Examples

- True model: 2D SEG model, background model: smoothed 2D SEG model
- Absorbing boundary conditions: taper the wavefield at each depth step (Serjan et al., 1985)
- Model parameters:
  - 85 grid points $\times$ 338 grid points, model size 840 $\times$ 3370 m
  - Spacing: $\Delta x = \Delta z = 10$ m
  - Source: Ricker wavelet with central frequency 15 Hz
  - Sources: 100 m spacing from $x = 100$ m to $x = 3300$ m at depth $z = 0$ m
  - Receivers: at every grid point at depth $z = 0$ m
- Data is generated by the linearized constant density acoustic frequency domain forward modeling operator

Examples

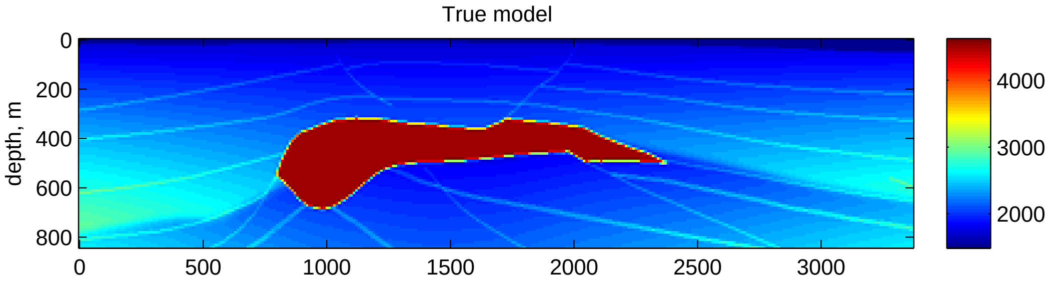


Figure 6: True velocity model

Examples

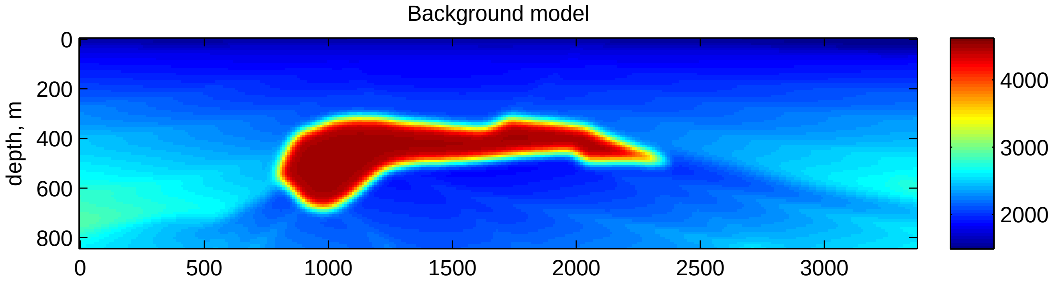


Figure 7: True velocity model

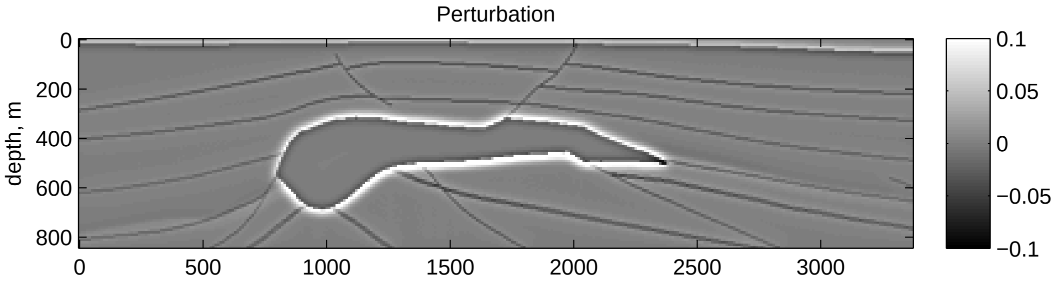


Figure 8: Model perturbation

Examples

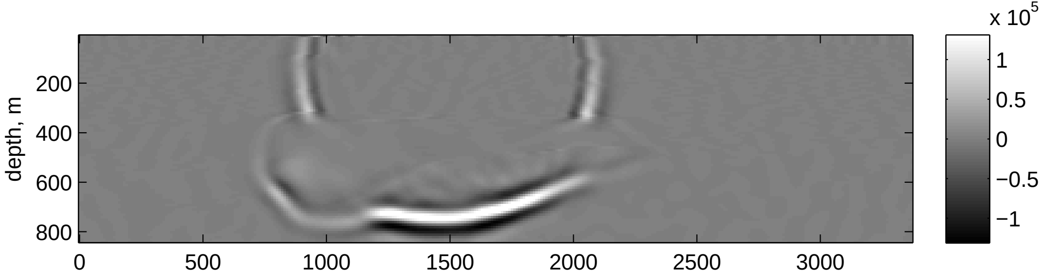


Figure 9: Wavefield time slice at $t = 0.35$ sec, source $x = 1500$ m

Examples

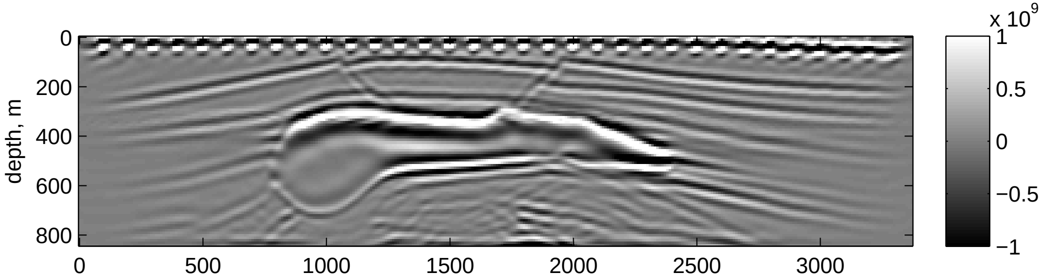


Figure 10: One way wave equation migration result

Examples

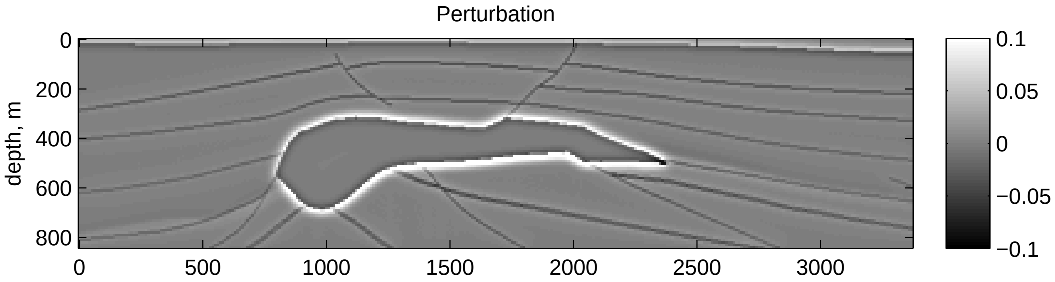


Figure 11: Model perturbation

Examples

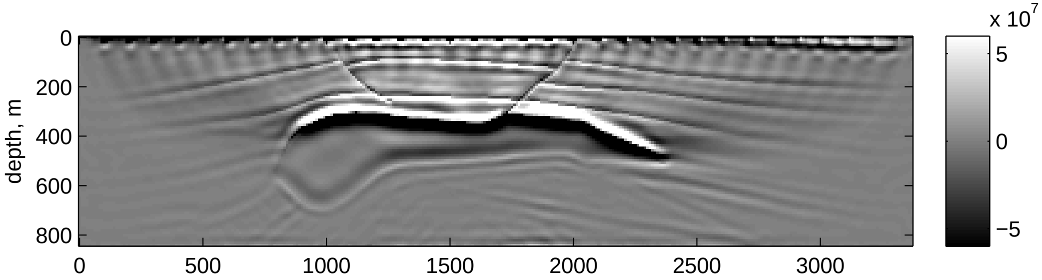


Figure 12: Reverse time migration result

Future work

- Implementation of matrix compression - currently use Matlab implementation that is not optimal
- Do more tests with HSS compression: precision seems to depend on HSS approximation accuracy, and does not get worse with increase of matrix size
- 3D implementation

Acknowledgement

This work was in part financially supported by the National Sciences and Engineering Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, BP, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Total SA, WesternGeco, and Woodside.

References

- G. Beylkin, N. Coult, M. Mohlenkamp, 1999. Fast spectral projection algorithms for density-matrix computations, Journal of Computational Physics, 152, 32-54.
- A. Ben-Israel & D. Cohen, 1966. On iterative computation of generalized inverses and associated projections, J. SIAM Numer. Anal., **3**, pp. 410-419.
- J. L. T. Grimbergen, F. J. Dessing, and K. Wapenaar, 1998. Modal expansion of one-way operators in laterally varying media, Geophysics, 63(3), 995âĂŞ1005.
- N.J. Higham, 2008. Functions of matrices: theory and computations, SIAM.
- C.S. Kenny, A.J. Laub, 1995. The matrix sign function, Automatic Control, IEEE Transactions, 40(3), 1330-1348.
- R. Kumar, H. Mansour, A. Y. Aravkin, F. J. Herrmann, 2013. Reconstruction of seismic waveïňĄelds via low-rank matrix factorization in the hierarchical-separable matrix representation, SEG Expanded Abstracts.
- T. Y. Lin & F.J. Herrmann, 2007. Compressed wavefield extrapolation, Geophysics, **72**, pp. SM77-SM93.
- W. Lyons, 2005. Fast algorithm with applications to PDEs, Ph.D. thesis, University of California, Santa Barbara.
- G.F. Margrave, M.P. Lamoureux, P. Gibson, R.A. Bale & J. Grossman, 2002. Exact wavefield extrapolation in 2D for v(x), CREWES Research Reports, http://www.crewes.com/ForOurSponsors/ResearchReports/2002/2002-13.pdf.

References

- K. Sandberg, G. Beylkin, 2009. Full-wave-equation depth extrapolation migration, Geophysics, 74(6), WCA121-WCA128.
- k. Sandberg, G. Beylkin, A. Vassiliou, 2010. Fill-wave-equation depth migration using multiple reflections. SEG Expanded Abstracts.
- Z. Sheng, P. Dewilde, S. Chandrasekaran, 2007. Algorithms to solve hierarchically semi-separable systems, Operator Theory: Advances and Applications Volume 176, 255-294
- C.P.A. Wapenaar, 1990. Representation of seismic sources in the one-way wave equations, Geophysics, **55**, pp.786-790.
- J. Xia, 2012. On the complexity of some hierarchical structured algorithms. SIAM Journal of Matrix Analysis and Applications, 33(2), 388-410.