CARP-CG: A robust parallel iterative solver for frequency-domain elastic wave modeling, application to the Marmousi2 model

Y. Li, Dept. Math., Harbin Institute of Technology, ISTerre, Université Grenoble Alpes, L. Métivier, LJK-CNRS, Université Grenoble Alpes, R. Brossier, ISTerre, Université Grenoble Alpes, B. Han, Dept. Math., Harbin Institute of Technology, J. Virieux, ISTerre, Université Grenoble Alpes*

SUMMARY

Solving the frequency-domain elastic wave equations relies on an efficient linear solver for the large, sparse, indefinite and ill-conditioned linear system derived from the discretization of the elastic wave equation. Direct solvers, which are mostly based on LU decomposition, are efficient for multiple right-hand sides problems, but the memory requirement is huge due to the fill-in effects. On the contrary, iterative solvers fully benefit from the sparsity of the system, but they require problem-specific preconditioners to ensure the convergence because of the ill-conditioning of the system. In this study, we investigate the performance of a robust iterative method named CARP-CG for frequency-domain elastic wave modeling. CARP-CG method turns the original system into a symmetric positive semi-definite system by Kaczmarz row-projections. Such a system can be efficiently solved by the conjugate gradient (CG) method. The row-projections can be seen as a purely algebraic preconditioning technique which is general and is easy to implement. The parallelization is straightforward through a row-block decomposition combined with a component-averaging method. We discretize the 2D frequency-domain elastic wave equation through a 4th order finite difference scheme. Numerical experiments on the Marmousi2 model exhibit a good scalability of CARP-CG. Comparisons between CARP-CG and standard Krylov iterative solvers (GMRES and CGNR) further emphasize the robustness and the fast convergence of CARP-CG method.

INTRODUCTION

An efficient linear solver of the time-harmonic elastic wave equations plays a key role in the frequency-domain implementation of reverse-time migration (RTM) and full waveform inversion (FWI) (Pratt, 1999; Virieux and Operto, 2009). The discretization of the time-harmonic wave equations yields a linear system related to a large sparse matrix, the so-called impedance matrix. Accounting for the subsurface heterogeneities described by space dependent parameters such as *P*-wave and *S*-wave velocities, density, or anisotropy parameters, yields an impedance matrix which is both indefinite and ill-conditioned. Thus, the resolution of the resulting system is a challenging task.

Solvers for tackling this challenge are divided into two main categories. Direct solvers have received an extensive popularity in the seismic imaging community because of their high efficiency in dealing with multiple-source problems. Direct solvers are based on a factorization of the impedance matrix which is done only once, the solution of each right-hand side being given by fast and simple computation tasks (forward and backward substitution for the LU decomposition for instance). However, the impedance matrix factorization is computationally expensive and requires a large amount of in-core memory. Even if the original impedance matrix is sparse, the fill-in introduces additional non-zero elements into the matrix, leading to a huge memory consumption. Although some improvements have been achieved recently in reducing the memory imprint (Wang et al., 2012; Weisbecker et al., 2013), the applications of direct factorization techniques is still limited for large problems, such as the ones related to realistic 3D elastic FWI applications.

Alternatively, iterative solvers fully benefit from the sparsity of the impedance matrix. The major computational cost of iterative solvers is related to matrix-vector products. It is both memory efficient and computationally cheap to perform these products since certain sparse format, such as compressed sparse row (CSR) format, can be applied to store the matrix and vectors. Nevertheless, the application of iterative solvers for solving frequency-domain wave equation is hampered by two important drawbacks: the solution of the linear system may require a large number of iteration to converge as the systems are ill-conditioned. In addition, iterative solvers lose the properties of handling multiple right-hand sides compared to direct solvers. The latter difficulty can be mitigated through the use of block Krylov subspace methods (O'Leary, 1980; Simoncini and Gallopoulos, 1996), or Galerkin projection methods (Chan and Ng, 1999). In this study, we focus on the first of these two difficulties and investigate the performances of a particular iterative solver for the solution of the frequency-domain elastic wave equations.

Iterative solvers are generally based on Krylov subspace methods such as the generalized minimum residual method (GMRES), the biconjugate gradient stabilized (BiCGSTAB) and the conjugate gradient applied to the normal equations (CGNR) (Saad, 2003). In the acoustic approximation, an efficient preconditioner based on a complex shifted Laplacian operator has been successfully implemented (Plessix, 2007; Riyanti et al., 2007; Erlangga and Nabben, 2008). A similar, algebraic version of this preconditioner has been proposed by Osei-Kuffuor and Saad (2010). Recently, Gordon and Gordon (2008) have proposed to use another iterative solver, based on the Kaczmarz method (Kaczmarz, 1937). The Kaczmarz row-projection strategy is used to transform an indefinite ill-conditioned linear system into a symmetric positive semi-definite system, which can be solved through a CG method. Later on, they combined a component averaging procedure with this method to provide an efficient parallelization. The resulting method is named as CARP-CG (Gordon and Gordon, 2010). This method has been proved efficient for the resolution of the 2D and 3D Helmholtz equation in heterogeneous media, even for high frequencies (van Leeuwen et al., 2012; Gordon and Gordon, 2013). The applications on the 2D frequency-domain elastic wave modeling in heterogeneous media with fixed Poisson's ratio were given in Li et al. (2014). In this study, we further investigate

DOI http://dx.doi.org/10.1190/segam2014-1181.1

the performances of the CARP-CG method for 2D frequency-domain elastic wave equation in strongly heterogeneous media.

In the following, we first present a brief review of the CARP-CG method. Then we present the 2D frequency-domain elastic wave equations and the 4th order finite-difference scheme used for their discretization. An anti-lumped mass strategy is implemented to reduce the number of discretization points. Validation of the CARP-CG solution is performed through a comparison with results from a time-domain code. Numerical experiments are implemented for the highly heterogeneous elastic Marmousi2 model (the Poisson's ratio reaches 0.48 in the elastic part of the model). Results with different numbers of processors show the scalability and the performance at high frequencies of CARP-CG. Finally, we compare the convergence properties of CARP-CG with GMRES and CGNR for different frequencies. CARP-CG achieves a fast convergence rate in all cases while the other two methods present at best slow convergence rates.

THE CARP-CG METHOD

We consider a linear system of n equations in n variables, $A\mathbf{x} = \mathbf{b}$. The Kaczmarz method (Kaczmarz, 1937) solve the system by cyclically projecting the iterates onto the rows of the matrix as follows:

$$\mathbf{x}' = \mathbf{x} + \omega_i(b_i - \mathbf{a}_{i*}^T\mathbf{x})\mathbf{a}_{i*}/\|\mathbf{a}_{i*}\|_2^2, \quad i = 1,\ldots,n, \qquad (1)$$

where the column vector $\mathbf{a}_{i*}$ denotes the i-th row of A, $\omega_i \in (0,2)$ is a relaxation parameter (in the experiments presented in this study, we use a constant relaxation parameter $\omega = 1$). Defining the matrices

$$Q_i = (I - \omega_i\mathbf{a}_{i*}\mathbf{a}_{i*}^T/\|\mathbf{a}_{i*}\|_2^2), \quad M_i = \omega_i\,\mathbf{a}_{i*}\mathbf{e}_i^T/\|\mathbf{a}_{i*}\|_2^2, \qquad (2)$$

where $\mathbf{e}_i, i = 1\ldots n$ is the canonical basis of $\mathbb{R}^n$, we can rewrite equation (1) as

$$\mathbf{x}' = Q_i\mathbf{x} + M_i\mathbf{b}. \qquad (3)$$

A loop of projections from the first to the last equations is referred to as a forward sweep. A backward sweep goes in the reverse order from n to 1. Then a double Kaczmarz sweep (a forward sweep followed by a backward sweep) is equivalent to:

$$\mathbf{x}' = Q\mathbf{x} + R\mathbf{b},\ Q = Q_1\cdots Q_nQ_n\cdots Q_1, \qquad (4)$$

$$R = Q_1\cdots Q_n\left(\sum_{i=1}^{n} Q_n\cdots Q_{i+1}M_i\right) + \sum_{i=1}^{n} Q_1\cdots Q_{i-1}M_i. \qquad (5)$$

Note that the Kaczmarz method used in double sweep with a constant relaxation parameter is equivalent to the Symmetric Successive Over Relaxation (SSOR) method applied to the normal equations. Considering the equation (4) as a fixed point iteration, we obtain the following linear system.

$$(I - Q)\mathbf{x} = R\mathbf{b}. \qquad (6)$$

The matrix $(I - Q)$ is symmetric and positive semi-definite (Gordon and Gordon, 2008). Then the CG method can be applied to (6) and the theoretical convergence is guaranteed. The resulting method is called CGMNC.

The CARP-CG method is a parallelization of CGMNC. The system of equations is first divided into t row-blocks. Denote by A^q, b^q the matrix and the right-hand side in block q, $1 \leqslant q \leqslant t$. For $1 \leqslant j \leqslant n$, we denote $I_j = \{1 \leqslant q \leqslant t \mid A_q$ contains an equation with a nonzero coefficient of $x_j\}$. We define the number of blocks which contain at least one equation with a nonzero coefficient of x_j as $s_j = |I_j|$. The component-averaging (CA) operator is a mapping $\mathrm{CA} : (\mathbb{R}^n)^t \to (\mathbb{R}^n)$, defined as follows:

$$\mathrm{CA}(y^1,\ldots,y^t)_j = \frac{1}{s_j}\sum_{q\in I_j} y_j^q,$$

$$\text{for} \quad y = \left[y^1,\ldots,y^t\right]^T \in (\mathbb{R}^n)^t, \quad \forall j = 1,\ldots,n, \qquad (7)$$

where y_j^q is the j-th component of y^q for $1 \leqslant q \leqslant t$. The definition of the CA operator allows to perform the double sweep operation in parallel: forward and backward sweeps are performed in parallel on each block assigned to a corresponding processor. The CA operator is applied after each forward and backward sweep to account for the coupling between the different blocks of the matrix.

If we denote by KSWP the forward sweep, BKSWP the backward sweep and DKSWP the double sweep, then DKSWP is formulated in Algorithm 1. In fact, the averaging operations are equivalent to add "averaging equations" to the original system (Gordon and Gordon, 2010). In addition, the projections on these averaging equations are independent of the sweeping order, which means that the averaging operations in DKSWP do not influence the symmetry and positive semi-definiteness of $(I - Q)$. Thus, the CG method remains applicable.

Algorithm 1 DKSWP.

for each $1 \leqslant q \leqslant t$ in parallel **do**
  $y^q = \mathrm{KSWP}(A^q, b^q, x)$
enddo
set $x' = \mathrm{CA}(y^1,\ldots,y^t)$.
for each $1 \leqslant q \leqslant t$ in parallel **do**
  $y^q = \mathrm{BKSWP}(A^q, b^q, x')$
enddo
set $x'' = \mathrm{CA}(y^1,\ldots,y^t)$.

2D FREQUENCY-DOMAIN ELASTIC WAVE EQUATION DISCRETIZATION

We consider the first-order velocity-stress formulation of the 2D frequency-domain elastic wave equations:

$$\begin{aligned} i\omega\rho v_x &= \partial_x\tau_{xx} + \partial_z\tau_{xz}, \quad i\omega\tau_{xx} = (\lambda+2\mu)\partial_x v_x + \lambda\partial_z v_z, \\ i\omega\rho v_z &= \partial_x\tau_{xz} + \partial_z\tau_{zz}, \quad i\omega\tau_{zz} = (\lambda+2\mu)\partial_z v_z + \lambda\partial_x v_x, \\ & \qquad\qquad i\omega\tau_{xz} = \mu(\partial_x v_x + \partial_z v_z), \end{aligned} \qquad (8)$$

where ρ is the density, λ and μ are the Lamé parameters, $v = (v_x, v_z)$ is the velocity displacement and $\tau = (\tau_{xx}, \tau_{zz}, \tau_{xz})$ are the stresses components. The source terms are omitted. Perfectly matched layers (PML) are applied to absorb outgoing waves (Bérenger, 1994; Collino and Tsogka, 2001).

We apply a 4th-order staggered-grid scheme (Levander, 1988) and the parsimonious approach (Luo and Schuster, 1990) to discretize the wave equations. The latter reduces the size of the system by considering only the velocity displacement. An anti-lumped mass strategy is applied to minimize the numerical dispersion. The optimal weighting coefficients corresponding to the nodes of the stencil are obtained by minimizing the misfit between unity and the normalized phase-velocity of P- and S-waves (Min et al., 2000). The phase velocity dispersion curves with Poisson's ratio $\sigma = 0.25$ and 0.49 are presented in Figure 1. We notice that the grid dispersion are only weakly dependent on Poisson's ratio and the numerical anisotropy of

DOI http://dx.doi.org/10.1190/segam2014-1181.1

S-waves is largely reduced compared to the conventional 4th-order scheme. To keep the error of phase velocities within 1%, the optimized 4th-order scheme only requires 3.3 points per minimum S-wavelength.

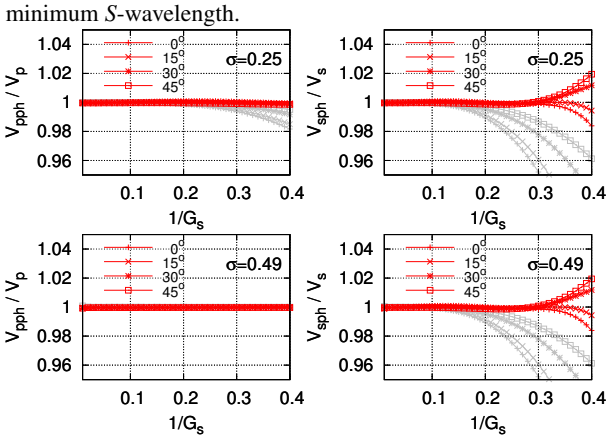

Figure 1: Normalized phase velocities for P-wave (left column) and S-wave (right column) with a Poisson's ratio of 0.25 (top row) and 0.49 (bottom row). G_s is the number of grid points per minimum S-wavelength. Dispersion curves are plotted for propagation angles of 0°, 15°, 30°, 45°. The curves in gray denote the conventional 4th-order scheme and the red curves are obtained by using the optimized coefficients.

NUMERICAL EXPERIMENTS

Comparison with a reference time-domain seismogram

We consider the Marmousi2 model (Figure 2) of size 3.5 km×17 km. Apart from the water layer on top in which the Poisson's ration is equal to 0.5 (the S-wave velocity is 0), the maximal Poisson's ratio of the water wet sand reaches 0.48. This configuration is challenging both for the finite-difference discretization and the convergence of the iterative solver. We compute the seismograms corresponding to a surface acquisition system with an explosive source located in the water layer at $x = 8500$ m, $z = 100$ m. The receivers are located along the axis $z = 100$ m. We use a Ricker source centered at 5 Hz. Since the minimum shear velocity apart from the water layer is approximately 270 m $\cdot$ s^{-1} and the grid step is $h = 10$ m, we have 5.4 points per minimum S-wavelength. A reference seismogram is computed using a $\mathcal{O}(\Delta t^2, \Delta x^4)$ time-domain elastic modeling code. The CARP-CG seismograms are obtained through the resolution of several linear systems for different frequencies and are transformed to the time domain through Fast Fourier Transform (FFT). The initial estimation of CARP-CG is taken as $\mathbf{x}^0 = 0$ and the iterations continue until $\|\mathbf{b} - A\mathbf{x}\|/\|\mathbf{b} - A\mathbf{x}^0\| < 10^{-6}$. The comparison of these seismograms is presented in Figure 3, which shows that CARP-CG converges to a solution consistent with the time-domain result.

Convergence and scalability

We perform several numerical experiments for the Marmousi2 model to investigate the convergence and the scalability of CARP-CG. The initial estimation is $\mathbf{x}^0 = 0$ and the stopping criterion is $\|\mathbf{b} - A\mathbf{x}\|/\|\mathbf{b} - A\mathbf{x}^0\| < 10^{-5}$. We increase the frequency from 1.25 Hz to 20 Hz. We keep 5.4 points per minimum S-wavelength in the following experiments to ensure the accuracy of the finite-difference scheme. A Dirac source located at $x = 8500$ m, $z = 100$ m is used. The real parts of

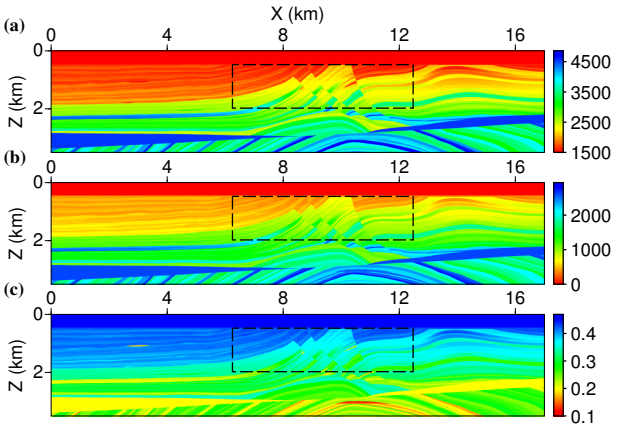

Figure 2: Marmousi2 model: (a) P-wave velocity, (b) S-wave velocity, and (c) Poisson's ratio. The part in black frame is used for comparison with GMRES and CGNR.

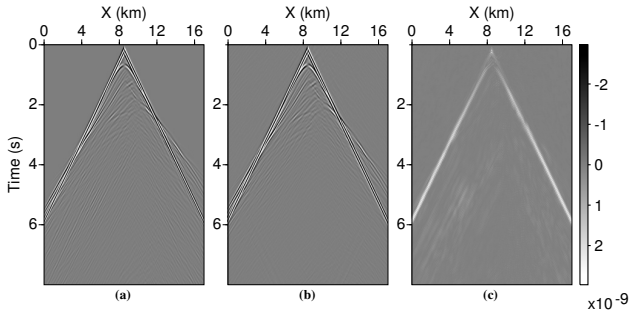

Figure 3: (a) Reference P-wave seismograms obtained from a time-domain code. (b) CARP-CG P-wave seismograms obtained after FFT. (c) Difference between the seismograms.

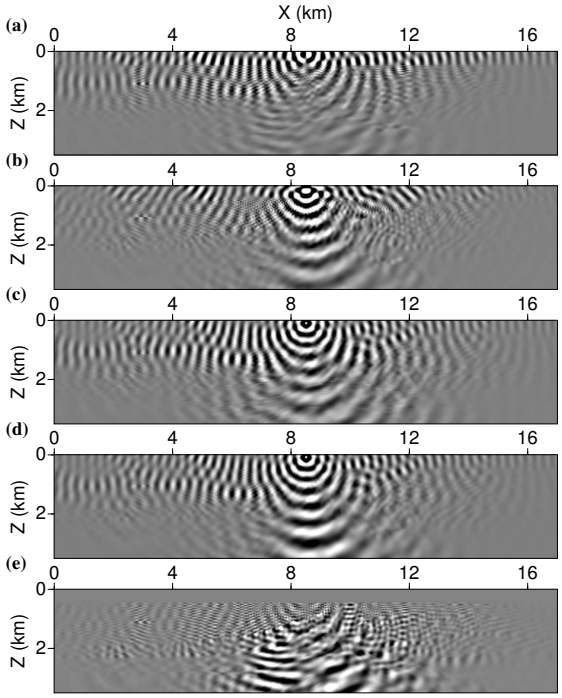

Figure 4: Real part of the 5 Hz wavefields corresponding to (a) v_x and (b) v_z, (c) σ_{xx}, (d) σ_{zz}, (e) σ_{xz}.

DOI http://dx.doi.org/10.1190/segam2014-1181.1

5 Hz wavefields corresponding to v_x, v_z σ_{xx}, σ_{zz} and σ_{xz} are presented in Figure 4 respectively. The interface between water and sediments can be seen in these fields, especially in the shear stress σ_{xz}.

Depending on the frequencies, the size of the linear system increases from 96,336 to 19,715,762 (Table 1). Figure 5 presents the scaling curves and the convergence histories using the number of processors given in Table 1. For each frequency, the scaling curve is close to the ideal scaling for low number of processors. For low frequency problems (related to models of smaller size), using more processors generates more communications and degrades the scaling efficiency. However, for larger scale problems (e.g., 10 Hz, 20 Hz), the scaling remains close to the ideal scaling until 128 processors. The scaling properties of the CARP-CG method thus seem encouraging.

The number of iterations and elapsed time are summarized in Table 1. The number of iterations does not increase as fast as the size of the linear system. The scaling curves seem to indicate that the elapsed time for large models could be further decreased by using more processors.

f (Hz)	h (m)	$n_x \times n_z$	time (s)	iterations
1.25Hz	40	88×426	7.6 (64)	3793
2.5Hz	20	176×851	23.6 (64)	4677
5Hz	10	351×1701	74.2 (128)	5958
10Hz	5	701×3401	242.2 (128)	7325
20Hz	2.5	1401×6801	1031.7 (128)	8656

Table 1: Geometry parameters, iteration counts and computation time for different frequencies using CARP-CG. The numbers in bracket correspond to the number of processors used when minimum elapsed time is achieved for each frequency.

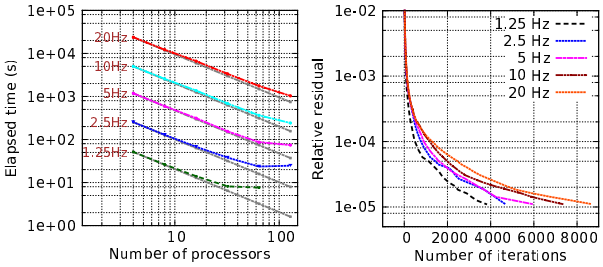

Figure 5: Scaling properties (left) and convergence histories (right) of CARP-CG. The gray lines on the left panel correspond to the ideal scaling of the computational time and the colored curves are the elapsed time of CARP-CG for each frequency.

Comparisons with GMRES and CGNR

We compare the performance of CARP-CG with two standard Krylov methods: GMRES (Frayssé et al., 2003) and CGNR. This comparison is carried out for sequential method: CARP-CG, therefore, reduces to CGMNC. We perform the simulations on a small domain extracted from the Marmousi2 model (Figure 2). We use a Dirac source located at $z = 150$ m, $x = 3125$ m (local coordinates). The frequencies are 1.25 Hz, 2.5 Hz, 5 Hz and 10 Hz. For each experiment, we ensure 5.4 grid points per minimum shear wavelength. In Figure 6, we present the convergence histories of these methods. The numbers of iterations and the elapsed time for each frequency are summarized in Table 2. The results indicate that CGMNC has a fast convergence rate in all cases, while, GMRES converges slowly or fails to converge after 10^5 iterations. CGNR is relatively robust but less efficient than CGMNC. This confirms the robustness and the interest of the CARP-CG method for solving the frequency-domain elastic wave equations.

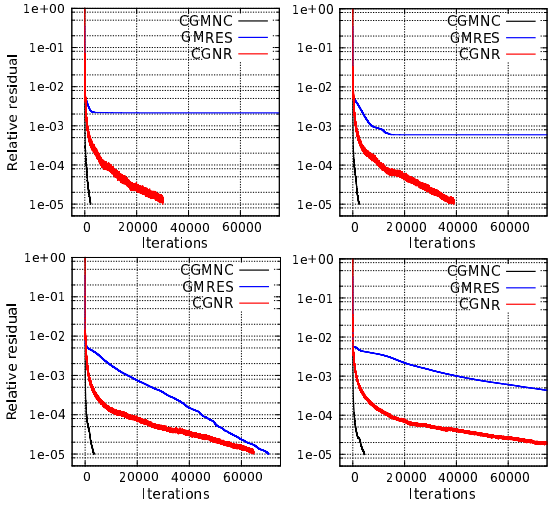

Figure 6: The convergence curves of CGMNC, GMRES and CGNR at 1.25 Hz (top left), 2.5 Hz (top right), 5 Hz (bottom left) and 10 Hz (bottom right).

f (Hz)	CGMNC	GMRES	CGNR
1.25Hz	2232 (20.2)	-	30079 (192.3)
2.5Hz	2425 (59.7)	-	38966 (715.8)
5Hz	3374 (258.4)	70535 (3531.3)	64645 (3400.6)
10Hz	4494 (1209.3)	-	-

Table 2: Number of iterations and elapsed time in seconds (in brackets) for CGMNC, GMRES and CGNR. The symbol ' - ' indicates nonconvergence after 10^5 iterations.

CONCLUSION

CARP-CG is a promising method for solving the system of equations describing frequency-domain elastic waves. The scalability of the method through row-block decomposition and the component averaging operator is satisfactory. Comparison with standard Krylov iterative solvers (GMRES, CGNR) emphasizes the good performance and robustness of CARP-CG. These results, obtained without any dedicated preconditioning technique, suggest that CARP-CG combined with an efficient preconditioner should be an efficient solver for this class of equations. This will be the topic of further studies, as well as the application of CARP-CG to 3D frequency-domain elastic wave modeling and its combination with block Krylov methods or Galerkin projection method to handle efficiently multiple right hand sides.

ACKNOWLEDGEMENTS

This study was partially funded by the SEISCOPE II consortium (*http://seiscope2.osug.fr*), National Nature Science Foundation of China under Grant No.91230119 and Chinese Scholarship Council. This study was granted access to the HPC facilities of CIMENT (Université Joseph Fourier Grenoble)

DOI http://dx.doi.org/10.1190/segam2014-1181.1

http://dx.doi.org/10.1190/segam2014-1181.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Bérenger, J.-P., 1994, A perfectly matched layer for absorpetion of electromagnetic waves: Journal of Computational Physics, **114**, no. 2, 185–200, http://dx.doi.org/10.1006/jcph.1994.1159.

Chan, T. F., and M. K. Ng, 1999, Galerkin projection methods for solving multiple linear systems: SIAM Journal on Scientific Computing, **21**, no. 3, 836–850, http://dx.doi.org/10.1137/S1064827598310227.

Collino, F., and C. Tsogka, 2001, Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media: Geophysics, **66**, 294–307, http://dx.doi.org/10.1190/1.1444908.

Erlangga, Y. A., and R. Nabben, 2008, On a multilevel Krylov method for the Helmholtz equation preconditioned by shifted Laplacian: Electronic Transactions on Numerical Analysis, **31**, 403–424.

Frayssé, V., L. Giraud, S. Gratton, and J. Langou, 2003, A set of GMRES routines for real and complex arithmetics on high-performance computers: Technical Report 3, CERFACS.

Gordon, D., and R. Gordon, 2008, CGMN revisited: Robust and efficient solution of stiff linear systems derived from elliptic partial differential equations: ACM Transactions on Mathematical Software, **35**, 18:1–18:27.

Gordon, D., and R. Gordon, 2010, CARP-CG: A robust and efficient parallel solver for linear systems, applied to strongly convection dominated PDEs: Parallel Computing, **36**, no. 9, 495–515, http://dx.doi.org/10.1016/j.parco.2010.05.004.

Gordon, D., and R. Gordon, 2013, Robust and highly scalable parallel solution of the Helmholtz equation with large wave numbers: Journal of Computational and Applied Mathematics, **237**, no. 1, 182–196, http://dx.doi.org/10.1016/j.cam.2012.07.024.

Kaczmarz, S., 1937, Angenäherte Auflösung von Systemen linearer Gleichungen (English translation by Jason Stockmann): Bulletin International de l'Académie Polonaise des Sciences et des Lettres, **35**, 355–357.

Levander, A. R., 1988, Fourth-order finite-difference P-SV-seismograms: Geophysics, **53**, 1425–1436, http://dx.doi.org/10.1190/1.1442422.

Li, Y., L. Métivier, R. Brossier, B. Han, and J. Virieux, 2014, A robust parallel iterative solver for frequency-domain elastic wave modeling: Presented at the 76th Annual International Conference and Exhibition, EAGE.

Luo, Y., and G. T. Schuster, 1990, Parsimonious staggered grid finite-differencing of the wave equation: Geophysical Research Letters, **17**, no. 2, 155–158, http://dx.doi.org/10.1029/GL017i002p00155.

Min, D.-J., C. Sin, B.-D. Kwon, and S. Chung, 2000, Improved frequency-domain elastic wave modeling using weighted-averaging difference operators: Geophysics, **65**, 884–895, http://dx.doi.org/10.1190/1.1444785.

DOI http://dx.doi.org/10.1190/segam2014-1181.1

O'Leary, D. P., 1980, The block conjugate gradient algorithm and related methods: Linear Algebra and Its Applications, **29**, 293–322, http://dx.doi.org/10.1016/0024-3795(80)90247-5.

Osei-Kuffuor, D., and Y. Saad, 2010, Preconditioning Helmholtz linear systems: Applied Numerical Mathematics, **60**, 420–431. (Special Issue: {NUMAN} 2008).

Plessix, R. E., 2007, A Helmholtz iterative solver for 3D seismic-imaging problems: Geophysics, **72**, no. 5, SM185–SM194, http://dx.doi.org/10.1190/1.2738849.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, part I: theory and verification in a physic scale model: Geophysics, **64**, 888–901, http://dx.doi.org/10.1190/1.1444597.

Riyanti, C. D., A. Kononov, Y. A. Erlangga, C. Vuik, C. Oosterlee, R. E. Plessix, and W. A. Mulder, 2007, A parallel multigrid-based preconditioner for the 3D heterogeneous high-frequency Helmholtz equation: Journal of Computational Physics, **224**, no. 1, 431–448, http://dx.doi.org/10.1016/j.jcp.2007.03.033.

Saad, Y., 2003, Iterative methods for sparse linear systems: SIAM.

Simoncini, V., and S. Gallopoulos, 1996, A hybrid block GMRES method for nonsymmetric systems with multiple right-hand sides: Journal of Computational and Applied Mathematics, **66**, no. 1–2, 457–469, http://dx.doi.org/10.1016/0377-0427(95)00198-0.

van Leeuwen, T., D. Gordon, R. Gordon, and F. Herrmann, 2012, Preconditioning the Helmholtz equations via row projections: 74th Annual International Conference and Exhibition, EAGE Extended Abstracts, A002.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, http://dx.doi.org/10.1190/1.3238367.

Wang, S., J. Xia, M. V. de Hoop, and X. S. Li, 2012, Massively parallel structured direct solver for equations describing time-harmonic qP-polarized waves in TTI media: Geophysics, **77**, no. 3, T69–T82, http://dx.doi.org/10.1190/geo2011-0163.1.

Weisbecker, C., P. Amestoy, O. Boiteau, R. Brossier, A. Buttari, J.-Y. L'Excellent, S. Operto, and J. Virieux, 2013, 3D frequency-domain seismic modeling with a block low-rank algebraic multifrontal direct solver: Presented at the 83rd Annual International Meeting, SEG.

DOI http://dx.doi.org/10.1190/segam2014-1181.1