Th P10 01

A Robust Parallel Iterative Solver for Frequency-domain Elastic Wave Modeling

Y. LI* (Harbin Inst. Tech.; Univ. Grenoble Alpes), L. Métivier (LJK, CNRS, Université Grenoble Alpes), R. Brossier (ISTerre, Université Grenoble Alpes), B. Han (Harbin Inst. Tech., Dept. Math.) & J. Virieux (ISTerre, Université Grenoble Alpes)

SUMMARY

Frequency-domain elastic wave modeling relies on an efficient linear solver for the large, sparse, ill-conditioned linear system derived from the discretization of the elastic wave equation. Direct solvers are mostly based on LU decomposition. These methods are efficient for multiple right-hand sides problems, but they require significant memory resources. Conversely, iterative solvers benefit from the sparsity of the system, but they require sophisticated preconditioners to converge due to the system ill-conditioning. In this study, we investigate the performance of an iterative method named CARP-CG for frequency-domain elastic wave modeling. The CARP-CG method transforms the original system into a symmetric positive semi-definite system by cyclic row-projections. This system is efficiently solved with the conjugate gradient (CG) method. The cyclic row-projection transformation can be seen as a purely algebraic preconditioning technique which is easy to implement. The algorithm can be parallelized through a row-block decomposition combined with component-averaging operations. Numerical experiments on the 2D frequency-domain elastic problem with high Poisson's ratio exhibit a good scalability of CARP-CG. Comparisons for different frequencies between CARP-CG and standard Krylov iterative solvers (GMRES and CG on the normal equations) emphasize the robustness and the fast convergence of the method.

Introduction

Frequency-domain elastic wave modeling is a challenging problem. After discretization, the problem amounts to the resolution of a large-scale, sparse, severely ill-conditioned linear system. Standard solvers for these systems are divided in two categories: direct and iterative solvers. Direct solvers are mostly based on LU factorization of the matrix (Amestoy et al., 2000). These methods are particularly efficient in the context of multiple right-hand sides problems: once the LU decomposition is performed, the linear systems are solved through backward and forward substitution. The main drawback of these methods is the significant memory requirement, related to the storage of the LU factors which are dense, despite the sparsity of the matrix.

Iterative solvers are generally based on Krylov subspace methods such as GMRES, BiCG(STAB) or Conjugate Gradient applied to the normal equations (CGNR)(Saad, 2003). Contrary to direct solvers, iterative solvers fully benefit from the sparsity of the system, in terms of memory requirements and scalability, as only matrix-vector products are required. However, these methods might present poor convergence properties or even fail to converge without preconditioning. The design of appropriate preconditioning matrix is therefore a crucial issue for these solvers. For instance, combination of complex Laplacian shifting and multigrid technique have been proven efficient for the Helmholtz equation (Erlangga and Nabben, 2008). This preconditioner may however not be adapted for the more general frequency-domain elastic wave equation.

Recently, Gordon and Gordon (2008) have proposed to use the Kaczmarz method (Kaczmarz, 1937) to transform an ill-conditioned linear system into a symmetric positive semi-definite system, which can be solved through a Conjugate Gradient (CG) process. Later on, they have shown how a Component Averaging Row Projection method (CARP) can be used in combination of this method to provide an efficient parallelization. The resulting method is named as CARP-CG (Gordon and Gordon, 2010). This method has been proven 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). In this study, we investigate the performance of the CARP-CG method for 2D frequency-domain elastic wave equation.

We first briefly introduce the CARP-CG method. Then we make a comparison between the frequency and time-domain results using the Marmousi 2 model to ensure CARP-CG converges to the right solution. Numerical experiments with different numbers of processors show the scaling properties of CARP-CG. Finally, we compare the convergence properties of CARP-CG with GMRES and CGNR.

The CARP-CG method

Consider a linear system of n equations in n variables, $A\mathbf{x} = \mathbf{b}$. The Kaczmarz method (Kaczmarz, 1937) cyclically projects the iterate as follows:

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

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 $\omega = 1$). We refer the projections from $i = 1$ to n as a forward sweep. A backward sweep goes in the reverse order from n to 1. Defining the matrices

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

where $\mathbf{e}_i$ is a standard basis of $\mathbb{R}^n$, e.g., $\mathbf{e}_1 = (1,0,\ldots,0)^T$, we can rewrite equation (1) as

$$\mathbf{x}' = Q_i\mathbf{x} + M_i\mathbf{b} \tag{3}$$

A double Kaczmarz sweep (a forward sweep followed by a backward sweep) amounts to:

$$\mathbf{x}' = Q\mathbf{x} + R\mathbf{b},\ Q = Q_1 \cdots Q_n Q_n \cdots Q_1,\ \ R = Q_1 \cdots Q_n \left(\sum_{i=1}^{n} Q_n \cdots Q_{i+1} M_i \right) + \sum_{i=1}^{n} Q_1 \cdots Q_{i-1} M_i, \tag{4}$$

If we consider equation (4) as a fixed point iteration we obtain the following linear system.

$$(I - Q)\mathbf{x} = R\mathbf{b}. \tag{5}$$

One can show that $(I-Q)$ is symmetric and positive semi-definite. Then CG can be applied to (5) 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 CA : $(\mathbb{R}^n)^t \to (\mathbb{R}^n)$, defined as follows:

$$\mathrm{CA}(y^1,\ldots,y^t)_j = \frac{1}{s_j}\sum_{q\in I_j} y_j^q, \quad \text{for} \quad y = \left[y^1,\ \ldots, y^t\right]^T \in (\mathbb{R}^n)^t, \quad \forall j = 1,\ldots,n, \tag{6}$$

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 parallelize the double sweep operation: forward and backward sweeps are performed in parallel on each block. The CA operator is applied after each forward and backward sweep to account for the coupling between the different blocks of the matrix.

Numerical results

We consider the first-order velocity-stress elastic wave equations:

$$\begin{aligned} \rho\partial_t v_x &= \partial_x\tau_{xx} + \partial_z\tau_{xz}, & \partial_t\tau_{xx} &= (\lambda+2\mu)\partial_x v_x + \lambda\partial_z v_z,\\ \rho\partial_t v_z &= \partial_x\tau_{xz} + \partial_z\tau_{zz}, & \partial_t\tau_{zz} &= (\lambda+2\mu)\partial_z v_z + \lambda\partial_x v_x,\\ & & \partial_t\tau_{xz} &= \mu(\partial_x v_x + \partial_z v_z), \end{aligned} \tag{7}$$

where ρ is the density, λ and μ are the Lamé parameters, $v = (v_x, v_z)$ is the velocity displacement and $\tau = (\tau_{xx}, \tau_{zz}, \tau_{xz})$ is the stress. A Fourier transform is used to transfer the system into the frequency-domain. Perfectly matched layers (PML) are introduced to absorb outgoing waves. The discretization is performed through a second-order staggered-grid finite-difference scheme (Virieux, 1986). We perform our numerical experiment using the Marmousi 2 P-wave model (Figure 1) of size 3.5 km$\times$17 km. The S-wave model is obtained by fixing the Poisson's ratio σ to 0.4. The density is constant equal to 1000 kg/m^3.

Comparison with the time-domain result

We first compute time seismograms corresponding to a surface acquisition system with an explosive source located at $x = 8500$ m, $z = 100$ m and the receivers equally spaced along $z = 100$ m. We use a Ricker source centered at 5 Hz. The discretization step is $h = 7.2$ m to ensure 15 gridpoints per minimum wavelength. A reference seismogram is computed using a time-domain elastic modeling code. The CARP-CG time seismogram is obtained through the resolution of several linear systems for different frequencies and is transformed to the time domain through Fast Fourier Transform (FFT). The comparison of these seismograms is presented in Figure 2, which shows that CARP-CG converges to a solution consistent with the time-domain result.

Scaling properties

The Marmousi 2 model presented in Figure 1 is used to investigate the scaling properties of CARP-CG. The initial estimation is taken as $\mathbf{x}^0 = 0$ and the iterations continue until $\|\mathbf{b} - A\mathbf{x}\|/\|\mathbf{b} - A\mathbf{x}^0\| < 10^{-6}$. We perform several experiments by increasing frequencies from 1.25 Hz to 20 Hz. We keep 15 points per wavelength for each experiment. A Dirac source located at $x = 8500$ m,$z = 100$ m is used. The size of the linear system increases from 174,460 to 44,006,720 (Table 1). Figure 3 presents the scaling curves and the convergence histories using the number of cores given in Table 1. For each experiment, the scaling curve is close to the ideal scaling for low number of cores. For small size problems (low frequency), the use of more cores yields more communication and degrade the scaling efficiency. However, for larger scale problems (10 Hz, 20 Hz), the scaling remains close to the ideal scaling until 128 cores. The scaling properties of the CARP-CG method thus seems encouraging. Due to the CA operator, using more cores only slightly increases the total number of iterations, especially for high frequency case. This is however expected since the matrix obtained through second-order finite-difference discretization is sparse and banded and the averaging operations involve only adjacent blocks (i.e., $s_j = 2$, $1 \leqslant j \leqslant n$). The number of iterations and computation time are summarized in Table 1.

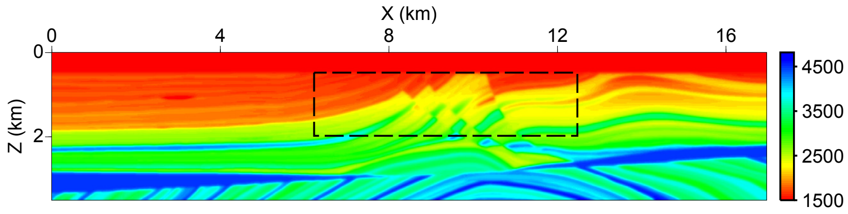

Figure 1 *The Marmousi 2 P-wave model. The small model in black frame is the one used for comparison with GMRES and CGNR.*

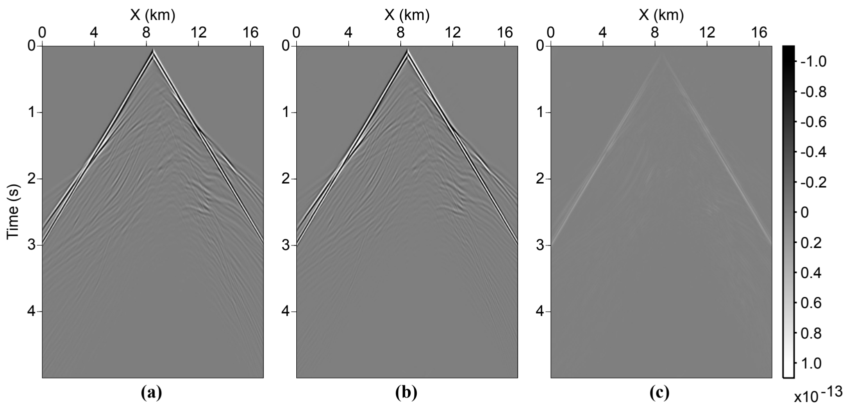

Figure 2 *Reference seismogram obtained from a time-domain code (a). CARP-CG seismogram obtained after FFT (b). Difference between the seismograms (c).*

f (Hz)	h (m)	$n_x \times n_z$	time (s)	iterations
1.25Hz	28.69	123×590	6.9 (64 cores)	7844
2.5Hz	14.40	244×1178	25.7 (64 cores)	10326
5Hz	7.21	486×2355	138.0(128 cores)	18790
10Hz	3.61	971×4709	598.6(128 cores)	25747
20Hz	1.81	1940×9416	2904.5 (128 cores)	35205

Table 1 *Geometry parameters, iteration counts and computation time for different frequencies using CARP-CG.*

Comparisons with GMRES and CGNR

Finally, we compare the performance of CARP-CG with standard methods: GMRES (Frayssé et al., 2003) and CGNR. This comparison is made for sequential method: CARP-CG is, therefore, degraded to CGMNC. We perform the simulations on a smaller domain extracted from the Marmousi 2 model (Figure 1). The Poisson's ratio is kept equal to 0.4. We use a Dirac source located at $x = 150$ m, $z = 3125$ m (local coordinates). The frequencies are 1.25 Hz, 2.5 Hz and 5 Hz. For each experiment, we ensure at least 15 discretization points per minimum wavelength. Figure 4 presents the convergence histories of the methods. The results indicate that CGMNC has a faster convergence rate and converges for all frequencies, while, without preconditioners, GMRES and CGNR converge slowly and fails to converge after 10^6 iterations when the frequency reaches 5 Hz.

Conclusions

CARP-CG seems to be a promising method for the frequency-domain elastic waveform modeling. A comparison with a time-domain code shows that the method converges to a consistent solution. The scalability of the method is satisfactory, more particularly, the row-block decomposition of the matrix and the Component Averaging operator do not degrade the convergence properties of the method. Comparison with standard Krylov iterative solvers (GMRES, CGNR) emphasizes the good performance and robustness of CARP-CG.

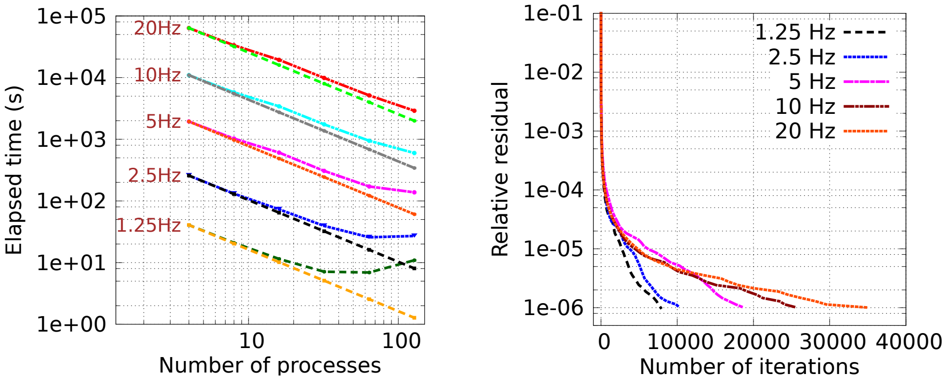

Figure 3 *Scaling properties (left) and convergence histories (right) of CARP-CG. The straight lines on the left panel correspond to the ideal scaling and the curves are the results from CARP-CG.*

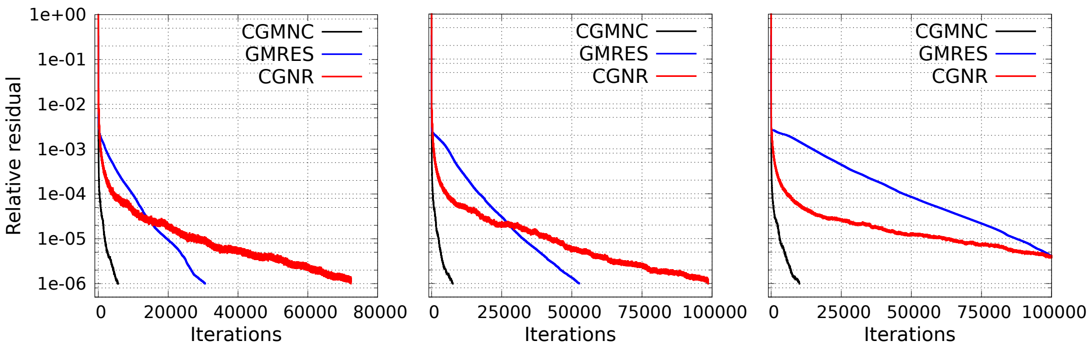

Figure 4 *The convergence curves of CGMNC, GMRES and CGNR at 1.25 Hz (left), 2.5 Hz (middle) and 5 Hz (right).*

Further research will focus on the application of CARP-CG on 3D elastic wave propagation and the influence of attenuation on the convergence. The combination of CARP-CG with higher order finite-difference schemes could be studied. The design of general preconditioners compatible with CARP-CG will also be investigated. For seismic imaging purpose, combination of CARP-CG method with multiple right-hand sides acceleration, such as block-CG methods, will be investigated.

Acknowledgements

This study was funded by the SEISCOPE II consortium (*http://seiscope2.osug.fr*), sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SHELL, SINOPEC, STATOIL, TOTAL and WESTERNGECO, the 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)

References

Amestoy, P., Duff, I.S. and L'Excellent, J.Y. [2000] Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering*, **184**(2-4), 501–520.

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

Frayssé, V., Giraud, L., Gratton, S. and Langou, J. [2003] A set of GMRES routines for real and complex arithmetics on high performance computers. Tech. Rep. 3, CERFACS, Technical Report TR/PA/03/3.

Gordon, D. and Gordon, R. [2008] CGMN revisited: Robust and efficient solution of stiff linear systems derived from elliptic partial differential equations. *ACM Trans. on Mathematical Software*, **35**(3), 18:1–18:27.

Gordon, D. and Gordon, R. [2010] CARP-CG: A robust and efficient parallel solver for linear systems, applied to strongly convection dominated PDEs. *Parallel Computing*, **36**, 495–515.

Gordon, D. and Gordon, R. [2013] Robust and highly scalable parallel solution of the Helmholtz equation with large wave numbers. *Journal of Computational and Applied Mathematics*, **237**(1), 182–196.

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.

Saad, Y. [2003] *Iterative methods for sparse linear systems*. SIAM, Philadelphia.

van Leeuwen, T., Gordon, D., Gordon, R. and Herrmann, F. [2012] Preconditioning the Helmholtz equations via row projections. *Expanded Abstracts*, EAGE, A002.

Virieux, J. [1986] P-SV wave propagation in heterogeneous media: velocity stress finite difference method. *Geophysics*, **51**, 889–901.