Preconditioning and multiple-right hand sides strategies for the solution of the frequency-domain wave propagation problem using the CGMN method

Okba Hamitou*,†,‡, Ludovic Métivier†,‡, Stéphane Labbé†, Romain Brossier† and Jean Virieux†

* Speaker; † ISTerre, University Grenoble Alpes; ‡ LJJK, University Grenoble Alpes

SUMMARY
Frequency-domain waveform modeling in the acoustic and elastic approximations requires the solution of large ill-conditioned linear systems. In the context of frequency-domain full waveform inversion, the solutions of these systems are required for a large number of sources (i.e. right-hand sides). Because of their tremendous memory requirements, direct solvers are not yet adapted to the solution of 3D elastodynamics equations. We are thus interested in the use of efficient iterative solvers adapted to the solution of these systems. The CGMN method has shown robust convergence properties for 2D and 3D elastic problems in highly heterogeneous media, compared to standard Krylov methods, but still requires a large number of iterations to reach sufficient accuracy. In this study, the design of an efficient preconditioning strategy adapted to this method is investigated. This preconditioner is computed as a sparse approximate inverse of a heavily damped wave propagation operator. In addition, the single seed method is used to increase the efficiency of the solver for multiple right-hand sides. The efficiency of these two combined strategies is evaluated on the 2D BP2004 model in the visco-acoustic approximation, up to 40 Hz. An overall time speed-up equal to 3 and a reduction of the number of iterations by a factor 10 are observed.

INTRODUCTION
Solving 3D frequency-domain acoustic and elastodynamics equations is a highly challenging problem. After discretization, these problems amount to the solution of a large, sparse, ill-conditioned and non-positive linear systems. These difficulties are magnified by the high contrasts and discontinuities in the physical properties of the subsurface. In the context of frequency-domain Full Waveform Inversion (FWI), the solution of the wave propagation problem is needed for a large number of sources (up to several tens of thousands for realistic applications). In the acoustic approximation, direct methods can be used. The factorization of the matrix is performed once and the solutions of the linear systems relative to each right-hand side can be obtained efficiently by forward and backward substitutions. Current FWI studies demonstrate the ability of using direct solvers in the 3D acoustic approximation for reasonable target size (Operto et al., 2014, 2015). However, the high memory requirement of these methods, due to the fill-in associated with the matrix factorization, makes them still inappropriate for solving 3D elastic problems.

In the perspective of performing 3D frequency-domain elastodynamics FWI, two strategies could thus be adopted. The first consists in solving the elastodynamics in the time-domain and using a Discrete Fourier Transform (DFT) on-the-fly to transform the solution into the frequency-domain. This strategy is appealing because of its straightforward implementation, however, three limitations can be identified. First, no Multiple Right Hand Sides (MRHS) acceleration can be implemented. Second, the time-integration is controlled by a CFL condition depending on the highest P-wave-velocity and the smallest discretization step. For elastodynamics simulations in media with slow S-wave velocity, this could lead to severe restrictions on the time step. Finally, the integration time should be long enough for the DFT to be sufficiently accurate.

These limitations lead us to rather investigate the possibility of solving the elastodynamics equations using iterative solvers. These algorithms fully benefit from the sparsity of the linear system as they only require the computation of scalar products and sparse matrix vector products, and therefore present significantly lower memory requirement. In this study, we are interested in the CGMN iterative solver, introduced by Björck and Elfving (1979). This method is based on the row projection strategy proposed by Kaczmarz (1937) to transform the initial non-symmetric system into a symmetric positive one which can be solved using a Conjugate Gradient (CG) method. This strategy can be efficiently parallelized using row-block decomposition and averaging techniques (Gordon and Gordon, 2010). It has been successfully applied to the frequency-domain wave modeling in the 3D acoustic approximation (van Leeuwen et al., 2012) as well as in the 2D and 3D elastic approximations (Li et al., 2014). In the latter case, CGMN reveals to be extremely robust: the convergence is obtained in media with fast variations, presenting high Poisson’s ratio, and with a free surface boundary condition responsible for the generation of surface waves. In the same configuration, standard Krylov solvers such as GMRES, BiCGSTAB, or Conjugate Gradient on Normal Equations (CGNR) (Saad, 2003) fail to converge.

Despite these relatively good properties, the CGMN method still requires a large number of iterations to reach sufficient accuracy. We are thus interested in the design of a suitable preconditioner for the CGMN method, allowing to increase the performance of this algorithm. In addition, in the perspective of FWI applications, a MRHS strategy based on the seed method proposed by Chan and Wan (1997) is applied to investigate the potential acceleration which can be expected.

A numerical example on the BP2004 case study is provided in the 2D visco-acoustic approximation. Using a 4th order finite-difference scheme the combination of the two strategies (preconditioning and MRHS) allow for an overall speed-up equal to 3 and a reduction of the total number of iterations by a factor 10 for a case study involving 1000 sources.
Preconditioning and MRHS strategies for frequency-domain wave propagation modeling

**THEORY**

**Overview of the CGMN method**

We consider the linear system derived from the discretization of the frequency-domain wave equation

\[ Ax = b, \tag{1} \]

where \( A \) is a complex and invertible matrix of size \( n \) and \( b \) is a complex vector. The rows of \( A \) are denoted by \( a_i \). Assuming these rows have been previously normalized, the Kaczmarz’s method cyclically projects the iterate \( x^k \) following

\[ x^{k+1} = x^k + \lambda \left( b_i - a_i^T x^k \right) a_i, \quad i = k + 1 \lfloor n \rfloor, \tag{2} \]

where \( \lambda \in [0; 2] \) is a relaxation parameter. Performing the projections \( (2) \) from the first to the last row of \( A \) is referred to as a forward sweep. Performing these projections in the reverse order is referred to as a backward sweep. A succession of a forward sweep and a backward sweep is referred to as a double sweep operation. Within this framework, the well known Symmetric Successive Over Relaxation (SSOR) iterative method (Saad, 2003) is equivalent to a sequence of double sweep operations. The solution of \( (1) \) through the SSOR method is computed through the iteration \( x^{k+1} = Q x^k + R b \), where \( Q \) and \( R \) are deduced from \( A \) and represent the double sweep operation. An acceleration of this fixed-point iteration can be obtained by considering the system

\[ (I - Q)x = Rb, \tag{3} \]

where the identity operator is denoted by \( I \). The matrix \( I - Q \) is symmetric and positive semi-definite (Björck and Elfving, 1979). Hence, the CG method can be applied to solve \( (3) \). The resulting method is called CGMN.

Using a finite-difference method for the discretization of the wave equation yields a sparse matrix \( A \) with \( s \) non-zero diagonals, \( s \) depending on the order of the discretization scheme. A double sweep requires \( 2ns \) operations (\( 2n \) projections on sparse rows), with \( s \) being in \( O(1) \). Using the double sweep operation, the complexity of the matrix vector product \( (I - Q)x \) is thus in \( O(n) \), despite the matrix \( I - Q \) is dense.

**Preconditioning strategy for the CGMN method**

The preconditioning method presented in this study relies on applying the CGMN method to the right preconditioned system

\[ APy = b, \quad x = Py, \tag{4} \]

where the matrix \( P \) is an approximate inverse of \( A \). Following earlier results on the use of complex-shifted Laplacian preconditioners (Erlangga and Nabben, 2008), a good preconditioner for \( A \) can be deduced from a heavily damped wave propagation operator \( B \). In this study, we build a preconditioner \( P \) as a sparse approximate inverse of \( B \). The \( k \)-th column of \( B^{-1} \), denoted by \( b_{k}^{-1} \), is the solution of the system

\[ b_{k}^{-1} e_k = b_{k}, \tag{5} \]

where \( e_k \) is a vector of size \( n \) with a single non-zero component equal to 1 at index \( k \). Assuming here a 2D approximation, let \((i_k, j_k)\) be the couple of indexes representing the spatial position associated with the index \( k \), the column \( b_{k}^{-1} \) can be identified with the pressure wavefield generated by a Dirac source located at the position \((i_k, j_k)\) in a strongly attenuating medium. Each column of \( P \) can thus be sparsely approximated by restraining \( b_{k}^{-1} \) to few discretization points around \((i_k, j_k)\) (see Figure 1). An efficient way to compute the preconditioner \( P \) is to solve the \( n \) local frequency-domain wave propagation problems in a strongly damped small medium (9 or 25 grid points according to the sparsity pattern which has been chosen). Perfectly Matched Layers (PML) (Bérenger, 1994) surround this local domain to absorb the outgoing waves. Using this method, the computation of the preconditioner presents low memory requirements as it only requires the solution of small frequency-domain wave propagation problems. In addition it is highly scalable as each column of the preconditioner can be computed independently.

This approach ensures \( P \) is composed of only few non-zero diagonals (9 or 25 in Figure 1). The number of non-zero diagonals of \( AP \), denoted by \( s \), increases with respect to the number \( s \) of non-zero diagonals of \( A \). A gain in number of arithmetic operations is obtained if the increase from \( s \) to \( s \) is compensated by the decrease of the number of iterations. This is quantified by the ratio \( \eta = (Ns)/(Ns) \), where \( N \) and \( N \) denote respectively the number of iterations performed by CGMN to solve the system \( (1) \) and \( (4) \).

**Seed method for linear systems with MRHS**

Consider the original linear system \( (1) \) with \( N_{rhs} \) MRHS

\[ AX = B, \tag{6} \]

where \( X = \left[ x^{(1)}, \ldots, x^{(N_{rhs})} \right] \) are the solutions of the linear system related to the right-hand sides \( B = \left[ b^{(1)}, \ldots, b^{(N_{rhs})} \right] \). The straightforward treatment for an iterative approach to solve the linear systems \( (6) \) is to solve each linear system independently. One way to accelerate this process is to compute initial guesses for the unsolved systems while solving the linear systems. This is efficiently done using the single seed method introduced by Smith et al. (1989) and Chan and Wan (1997).

The single seed method takes advantage of the Krylov subspace generated along the solution of one linear system in \( (6) \), referred to as the seed system. The residuals of the others systems are projected orthogonally on this Krylov subspace to compute an approximate initial guess. Once the seed system is solved, this process is repeated by selecting the next “non seed” systems as the seed system. This strategy is presented in Algorithm 1. For seismic experiments, the MRHS are associated with distinct source positions. As a consequence, the MRHS are orthogonal one to each other. In practice the seed method performs poorly in such configurations. This is understandable as in this case the residuals associated with one system are nearly orthogonal to the Krylov space generated through the solution of the seed system. Therefore the ini-

---

Figure 1: Sparsity pattern for computing the preconditioner using 9-points (left) and 25-points (right) per column.

© 2015 SEG
DOI http://dx.doi.org/10.1190/segam2015-5889724.1
SEG New Orleans Annual Meeting
Page 3613
Preconditioning and MRHS strategies for frequency-domain wave propagation modeling

Algorithm 1 Single seed method algorithm

1. Compute Conjugate Gradient initial residuals
2. for \( k = 1, \ldots, N_{\text{iter}} \) do
3. \hspace{1em} Set the \( k \)-th system as the seed system
4. \hspace{1em} for \( i = 1, \ldots, n \), until the seed system is solved do
5. \hspace{2em} Perform Conjugate Gradient iteration:
6. \hspace{3em} Compute \( p_1^{(k)} \), \( p_2^{(k)} \) and \( x_1^{(k)} \)
7. \hspace{1em} for \( j = k + 1, \ldots, N_{\text{iter}} \) do
8. \hspace{2em} Compute the projected solutions and residuals
9. \hspace{3em} \( \eta \leftarrow (p_1^{(k)})^T T_{i,j}^k ((p_1^{(k)})^T (I - Q)p_i^k) \)
10. \hspace{3em} \( x_{i+1} = x_i + \eta p_i^k \)
11. \hspace{3em} \( r_{i+1}^{(j)} = r_i^{(j)} - \eta(I - Q)p_i^k \)
12. \hspace{2em} end for
13. \hspace{1em} end for
14. for \( k = 1, \ldots, N_{\text{iter}} \) do
15. \hspace{1em} \( \gamma_0^{(k)} \leftarrow Rb^{(k)} - (I - Q)\gamma_0^{(k)} \)
16. end for

The initial guess computed through the seed method is close to zero. However, one can circumvent this problem by transforming the system (6) into

\[
AY = MB, \quad X = M^{-1}Y, \tag{7}
\]

where \( M \in \mathbb{R}^{N_{\text{rhs}} \times N_{\text{rhs}}} \) is an invertible matrix which we will refer to as the blending matrix. For the numerical experiments, we will choose the matrix \( M \) of the form

\[
M = \begin{pmatrix}
1 - \gamma & 1 & \cdots & 1 \\
1 & \ddots & \ddots & \vdots \\
& \ddots & \ddots & 1 \\
1 & \cdots & 1 & 1 - \gamma
\end{pmatrix}, \tag{8}
\]

with \( \gamma \in [0, 1] \). The choice of \( \gamma \) should be done with caution. Indeed, as \( \gamma \) tends towards 1, the seed strategy tends to be particularly efficient, as the right hand side \( MB \) tends to be similar. However, the condition number of \( M \) also increases, and the multiplication of \( Y \) by \( M^{-1} \) degrades the accuracy of the solution. The accuracy required for the solution of \( Y \) should thus be done accordingly, which decreases the efficiency of the seed strategy. In the case study presented here, a choice of \( \gamma = 0.2 \) seems to represent an acceptable trade-off between these two requirements.

TEST ON 2D MODEL

We consider the frequency-domain visco-acoustic wave equation

\[
-\rho(x) v_p(x)^2 p(x) - \text{div} \left( \frac{1}{\rho(x)} \nabla p(x) \right) = s(x, \omega), \tag{9}
\]

where \( \omega = 2\pi f \) is the angular frequency, \( v_p(x) \) is the P-wave velocity, \( \rho(x) \) is the density, \( s(x, \omega) \) is the source term and \( p(x, \omega) \) is the pressure wavefield.

The attenuation is accounted for using the Kolsky model, resulting in a complex-valued P-wave velocity. The velocity is therefore defined by \( v_p(x) = v_p(x)(1 - 0.5i q_{\text{att}}) \), where \( q_{\text{att}} \) is the quality factor. FMI are attached to the four sides of the domain to mimic an infinite medium and absorb outgoing energy.

We perform the numerical experiments using the 67 km long and 12 km deep 2004 BP velocity and density model (see Figure 2). The quality factor \( q_{\text{att}} \) is taken equal to 1000, which corresponds to a non-attenuating medium. The discretization of the equation (9) is performed using the fourth-order staggered-grid finite-difference scheme of Leveander (1988). At least 5 discretization points per wavelength are used for all frequencies. The pressure wavefield solution of (9) for \( f = 5 \) Hz is presented in Figure 3. The number of iterations performed by the iterative solver to converge is denoted by \( N_{\text{iter}} \). The geometric mean size of the domain is denoted by \( N = \sqrt{n} \). The initial estimation is taken as \( s^0 = 0 \) and the stopping criterion is fixed to \( 10^{-4} \) on the relative residual \( \|Ax - b\|/\|Ax^0 - b\| \). The relaxation parameter \( \lambda \) is equal to 1.2.

Evaluating the efficiency of the preconditioner

The preconditioner \( P \) is computed upon a strongly damped medium with \( q_{\text{att}} = 1 \) and using the 9-points sparsity pattern (Fig. 1). The results obtained with this preconditioner are shown in Table 1. The computation time is divided by a factor up to 2.2. The gain in computational cost \( \eta \) reaches a factor 2. The preconditioner reduces the number of iterations up to a factor 6.9. However this gain is compensated by the increase of the

![Figure 2: 2004 BP model: velocity (top) and density (bottom).](image2)

![Figure 3: Pressure wavefield at frequency \( f = 5 \) Hz for a Dirac source located at \( x = 33 \) km and \( z = 180 \) m.](image3)

<table>
<thead>
<tr>
<th>( f ) (Hz)</th>
<th>( n )</th>
<th>CGMN</th>
<th>PCGMN</th>
<th>CGMN</th>
<th>PCGMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11551</td>
<td>845</td>
<td>164</td>
<td>2.3</td>
<td>1.4</td>
</tr>
<tr>
<td>5</td>
<td>232617</td>
<td>3313</td>
<td>569</td>
<td>206.9</td>
<td>109.7</td>
</tr>
<tr>
<td>10</td>
<td>904128</td>
<td>5959</td>
<td>951</td>
<td>1443.5</td>
<td>728.5</td>
</tr>
<tr>
<td>20</td>
<td>3564132</td>
<td>9525</td>
<td>1424</td>
<td>9195</td>
<td>4355.4</td>
</tr>
<tr>
<td>40</td>
<td>14152068</td>
<td>13469</td>
<td>1942</td>
<td>51480.9</td>
<td>23402.6</td>
</tr>
</tbody>
</table>

Table 1: Number of iterations and time of computations performed by the CGMN method with and without the preconditioner.
Preconditioning and MRHS strategies for frequency-domain wave propagation modeling

Figure 4: Convergence curves of the normalized residuals of the solution of the frequency-domain wave propagation problem using the CGMN method (black solid curve) and the preconditioned CGMN method (red dotted curve) for \( f = 5 \) Hz.

Figure 5: Complexity analysis of the CGMN method (lower red dotted curve) with and (upper red dotted curve) without the preconditioner. The black solid line and the blue “x”-marked line represent respectively the linear increase and trend line.

The cost of each iteration. Indeed, the matrix \( AP \) has 45 non-zero values per row whereas the matrix \( A \) has only 13 non-zero entries. The improvement of the convergence profile of the residuals for \( f = 5 \) Hz is shown in Figure 4. A complexity analysis of the CGMN method is presented in Figure 5. The number of iterations \( N_{iter} \) performed by the preconditioned and non-preconditioned CGMN method is plotted as a function of \( N \) on a log-log scale. The computational complexity is sub-linear in \( N \), which is consistent with previous studies (Li et al., 2014). Based on this analysis, the CGMN complexity is evaluated as \( N_{iter} \approx 20.9 N^{0.79} \). Using the preconditioner, the complexity is improved to \( N_{iter} \approx 5.9 N^{0.71} \).

A multiple right-hand side strategy

The gain in computational time and number of iterations of the single seed method is evaluated for \( f = 5 \) Hz yielding 209 grid points along the vertical direction, 1113 grid points along the horizontal direction and a space grid step \( d = 60 \) m. The number of sources \( N_{rhs} \) is set to 1000. The sources are located at \( z_0 = 180 \) m and regularly spaced by a space grid step \( d \) along the horizontal direction. The blending matrix \( M \) is defined with \( \gamma = 0.2 \).

Applying the single seed method as it has been introduced in the previous section over all the \( N_{rhs} \) right-hand sides leads to poor results in terms of computational time. The large number of sources makes the computational cost related to the additional projection operations (lines 7 to 12 in algorithm 1) too expensive, and not compensated by the reduction of the overall number of iterations. Therefore, we apply the seed method only on smaller subsets of right-hand sides. We divide the 1000 RHS into 50 subsets of 20 RHS and apply our method sequentially on each subset. The improvement of the convergence profile when solving the linear systems with the subset of RHS is shown in Figure 6. Linear systems using the projected initial guesses converge in an increasingly smaller number of iterations. Compared to the CGMN method using the preconditioner, the computational time is then divided by a factor 1.5 whereas the number of iterations is divided by a factor 1.6. Compared to the CGMN method without preconditioner, the computational time is divided by a factor 2.9 whereas the number of iterations is divided by a factor 10.

CONCLUSION

A sparse approximate inverse preconditioner for frequency-domain wave equation and a MRHS strategy are presented in this study. The preconditioner is built as an approximate inverse of a strongly damped frequency-domain wave propagation operator. Its evaluation relies on the solution of local systems of small size which can be performed independently making its construction highly scalable. The preconditioner accelerates substantially the convergence of the CGMN solver. However this gain is compensated by the increase of the computational cost of each iteration and the final speed-up is around 2.2 for a system of 14 million unknowns. The single seed method is used to accelerate the solution of the system for 1000 right-hand-sides. When applied to small subsets of the right-hand sides, an additional speed up of 1.5 is observed.

Further work will be dedicated to the application of the preconditioner to 2D and 3D realistic elastic systems, and adapted to the parallel version of the CGMN algorithm, named as the CARP-CG method (Gordon and Gordon, 2010). Block conjugate gradient method (O’Leary, 1980) will be investigated together with the single seed method for improving the MRHS strategy. Ultimately, this method should serve as an efficient forward modeling engine for performing 3D frequency-domain FWI in the elastodynamics approximation.

ACKNOWLEDGMENTS

This study was funded by the SEISCOPE consortium (http://seiscope2.osug.fr), sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SCHLUMBERGER, SHELL, SINOPEC, STATOIIL, TOTAL and WOODSIDE. This study was granted access to the HPC resources of CIMENT infrastructure (https://ciment.ujf-grenoble.fr) and CINES/IDRIS under the allocation 046091 made by GENCI.
EDITED REFERENCES
Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2015 SEG Technical Program Expanded Abstracts have been copyedited 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


