

## 2D and 3D frequency-domain elastic wave modeling in complex media with a parallel iterative solver

Yang Li<sup>1</sup>, Ludovic Métivier<sup>2</sup>, Romain Brossier<sup>3</sup>, Bo Han<sup>4</sup>, and Jean Virieux<sup>3</sup>

### ABSTRACT

Full-waveform inversion and reverse time migration rely on an efficient forward-modeling approach. Current 3D large-scale frequency-domain implementations of these techniques mostly extract the desired frequency component from the time-domain wavefields through discrete Fourier transform. However, instead of conducting the time-marching steps for each seismic source, in which the time step is limited by the stability condition, performing the wave modeling directly in the frequency domain using an iterative linear solver may reduce the entire computational complexity. For 2D and 3D frequency-domain elastic wave modeling, a parallel iterative solver based on a conjugate gradient acceleration of the symmetric Kaczmarz row-projection method, named the *conjugate-gradient-accelerated component-averaged row projections* (CARP-CG) method, shows interesting convergence

properties. The parallelization is realized through row-block division and component averaging operations. Convergence is achieved systematically even when different physical factors such as the space-dependent Poisson's ratio, free-surface condition, and seismic attenuation are incorporated in the wave modeling. We determined that the scalability of CARP-CG was satisfactory, especially for large-scale applications, using up to several hundred computational cores. We found a potential improvement in computational complexity compared to time-domain modeling through numerical experiments. Finally, we achieved a convergence at 5 Hz in a 3D heterogeneous model, involving fast-slow-fast layers resembling waveguide geometries, with up to several hundred million unknowns, in fewer than 10 h on fewer than 200 cores. All of these results make CARP-CG a potential candidate of the forward modeling engine for seismic imaging on challenging models.

### INTRODUCTION

An efficient solver for the frequency-domain elastic wave equations is the cornerstone of any frequency-domain implementation of reverse time migration (RTM) and full-waveform inversion (FWI) (Pratt, 1999; Virieux and Operto, 2009). The discretization of the frequency-domain elastic wave equations with a volumetric method such as finite difference (FD) or finite element yields a large and sparse impedance matrix. Accounting for the subsurface heterogeneities, described by space-dependent parameters such as P-wave and S-wave velocities, density or anisotropy coefficients, yields an impedance matrix that is indefinite and ill conditioned. Solving the resulting linear system is a challenging task.

Strategies for obtaining frequency-domain modeling results may be divided into three main categories: First, the modeling can be performed in the time domain and the specific frequency components can be extracted through a discrete Fourier transform (DFT). This is the strategy implemented for some 3D FWI large-scale applications (Sirgue et al., 2007, 2008). Although it currently seems to be an efficient and flexible strategy to handle large-scale problems, one significant drawback is that the time-marching steps must be performed independently for each source. In addition, recent studies show that the DFT is a nonnegligible computational task (Brossier et al., 2014).

The second and third strategies rely in performing the wave modeling in the frequency domain. Direct solvers have received an extensive popularity in the seismic imaging community because of

Manuscript received by the Editor 8 October 2014; revised manuscript received 4 January 2015; published online 8 May 2015.

<sup>1</sup>Harbin Institute of Technology, Department of Mathematics, Harbin, China and Institut des Sciences de la Terre (ISTerre), Université Grenoble Alpes, CNRS, Grenoble, France. E-mail: liyang19880223@163.com.

<sup>2</sup>Laboratoire Jean Kuntzmann (LJK), Université Grenoble Alpes, CNRS, Grenoble, France. E-mail: ludovic.metivier@ujf-grenoble.fr.

<sup>3</sup>Université Grenoble Alpes, CNRS, Institut des Sciences de la Terre (ISTerre), Grenoble, France. E-mail: romain.brossier@ujf-grenoble.fr; jean.virieux@ujf-grenoble.fr. Institut des Sciences de la Terre (ISTerre), Université Grenoble Alpes, CNRS, Grenoble, France. E-mail: .

<sup>4</sup>Harbin Institute of Technology, Department of Mathematics, Harbin, China. E-mail: bohan@hit.edu.cn.

© 2015 Society of Exploration Geophysicists. All rights reserved.

their high efficiency in dealing with multiple-source problems. These solvers are based on a factorization of the impedance matrix that is performed only once. The solution of each seismic source (right-hand side) can be given by fast and simple computation tasks (forward and backward substitution for the lower and upper (LU) decomposition, for instance). However, impedance matrix factorization is computationally expensive and it requires a large amount of in-core memory. Even if the original impedance matrix is sparse, the fill-in effect introduces additional nonzero elements into the matrix, leading to a huge memory requirement. Although some improvements have been achieved recently in reducing the memory imprint by using some low-rank compressions (Wang et al., 2012; Weisbecker et al., 2013), the application of direct factorization techniques is still out of reach for large-scale 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. Performing these products is memory efficient and computationally cheap because sparse formats can be used to store the matrices and vectors. Moreover, for FWI applications, one could easily control the accuracy of the forward modeling with an iterative solver because great accuracy might not be needed in the initial stage of the inversion (van Leeuwen and Herrmann, 2014). Nevertheless, the application of iterative solvers for solving the frequency-domain wave equation is hampered by the uncertainty of the convergence and/or the large number of iterations required to achieve the convergence.

The advantages and disadvantages of these three strategies can be illustrated by analyzing the computational and memory complexities. We consider a 3D model of size  $N \times N \times N$ , where  $N$  denotes the number of grid points in one dimension. In realistic applications, the number of seismic sources is  $N_s = \mathcal{O}(N^2)$ , and the number of time-marching steps for the time-domain modeling is  $N_t = \mathcal{O}(N)$  according to the Courant-Friedrichs-Lewy stability condition. The number of iterations for frequency-domain modeling with an iterative solver is denoted by  $N_{\text{iter}}$ . We summarize the computational complexity and memory complexity for each strategy in Table 1. The computational complexity is  $N^3$  for each time-marching step of the time-domain modeling and for each iteration of iterative solvers. The computational complexity of  $\mathcal{O}(N^6)$  for the direct solver strategy is related to the factorization of the impedance matrix (Operto et al., 2007; Nihei and Li, 2007).

As can be seen in Table 1, using iterative solvers may present three advantages: First, it requires the same memory complexity as the combination of the time-domain modeling and the DFT, which

is one order smaller than using the direct solvers. Moreover, compared to the other two strategies where the computational complexities reach  $\mathcal{O}(N^6)$ , frequency-domain modeling through an iterative solver may have a smaller computational complexity if  $N_{\text{iter}}$  can be optimized so that  $N_{\text{iter}} = \mathcal{O}(N^\alpha)$  with  $\alpha < 1$ . Finally, one could combine the iterative solver with specific techniques to handle the multiple right-hand sides (MRHS) problems, such as the block Krylov subspace methods (O'Leary, 1980; Simoncini and Gallopoulos, 1996) or Galerkin projection methods (Chan and Ng, 1999). No similar methods exist for time-domain modeling.

Iterative solvers are generally based on Krylov subspace methods, such as the generalized minimum residual method (GMRES) (Saad, 1986), the biconjugate gradient stabilized (BiCGStab) method (Van der Vorst, 1992; Sleijpen and Fokkema, 1993) and the conjugate gradient (CG) applied to the normal equations (CGNR) (Saad, 2003). Solving the linear system derived from the wave equation may require a large number of iterations to converge because the linear system is indefinite and ill conditioned (Ernst and Gander, 2012). An efficient preconditioner based on a complex shifted Laplacian operator has been successfully implemented (Plessix, 2007; Riyanti et al., 2007; Erlangga and Nabben, 2008) for the acoustic wave equation. However, extensions to the elastic case have not been performed, to the best of our knowledge. In this study, we focus on the application of another iterative solver, called the *conjugate-gradient-accelerated component-averaged row projections* (CARP-CG) method, to the 2D and 3D elastic wave modeling. Kaczmarz row projections (Kaczmarz, 1937) are used to transform the indefinite ill-conditioned linear system into a symmetric positive semidefinite system. Using the CG method to solve the derived linear system leads to the so-called conjugate-gradient minimum norm (CGMN) method (Björck and Elfving, 1979; Gordon and Gordon, 2008). Gordon and Gordon (2010a) propose to combine a component averaging (CA) procedure with this method to provide an efficient parallelization. The resulting method is called *CARP-CG*; it has been proved efficient for solving the 2D and 3D acoustic Helmholtz equation in heterogeneous media, even for high frequencies (van Leeuwen et al., 2012; Gordon and Gordon, 2013). Applications on 2D frequency-domain elastic wave modeling in heterogeneous media are presented in Li et al. (2014a, 2014b). In this study, we further investigate the performances of the CARP-CG method for 2D and 3D the frequency-domain elastic wave equation in strongly heterogeneous media.

In the following, we first present a brief introduction of the CARP-CG methodology. Then, we present the 3D frequency-domain elastic wave equations in orthorhombic media and the fourth-order staggered-grid finite-difference scheme for its discretization (Virieux, 1986; Levander, 1988). An antilumped mass strategy is applied to reduce the numerical dispersion. The coefficients of the FD scheme and the mass averaging operations are optimized by minimizing the misfit between the normalized phase velocities and the unity. Several 2D elastic numerical experiments are performed to investigate the influence of different physical factors on the convergence of the CARP-CG method: Poisson's ratio, the free-surface boundary condition (FSBC), and seismic attenuation. CARP-CG is shown to be very robust; that is, it is capable of converging, despite these highly heterogeneous physical factors. For the 2D elastic Marmousi2 model (Martin et al., 2006), results using a different number of cores show the scalability and the performance at high frequencies of the CARP-CG method. In addition, we compare the convergence properties of CARP-CG with GMRES, CGNR, and

**Table 1. Computational and memory complexities for different 3D wavefields modeling strategies, without considering parallelization methods.**

	Computational complexity	Memory complexity
Time domain + DFT	$\mathcal{O}(N_s \times N_t \times N^3) \approx \mathcal{O}(N^6)$	$\mathcal{O}(N^3)$
Frequency domain + direct	$\mathcal{O}(N^6)$	$\mathcal{O}(N^4)$
Frequency domain + iterative	$\mathcal{O}(N_{\text{iter}} \times N_s \times N^3) \approx \mathcal{O}(N_{\text{iter}} \times N^5)$	$\mathcal{O}(N^3)$

BiCGStab for different frequencies. The CARP-CG method converges at a faster rate in each case. Through the experiments on 3D homogeneous and heterogeneous elastic models, CARP-CG is shown to be capable of handling 3D large-scale frequency-domain wave modeling. Moreover, CARP-CG exhibits a very regular convergence, even for modeling in the heterogeneous model, where fast-slow-fast “trap” layers generate a severely ill-conditioned linear system.

These experiments emphasize the robustness of CARP-CG for large-scale elastic 2D and 3D problems (especially compared to more standard Krylov solvers) as well as sublinear complexity and correct scaling properties. This suggests that CARP-CG should be investigated further for application in seismic imaging methods such as FWI. However, mandatory improvement in terms of the number of iterations should be provided before that. This could be achieved by proper preconditioning strategies and adaptation of the algorithm to the solution of MRHS.

### PRESENTATION OF THE CONJUGATE-GRADIENT-ACCELERATED COMPONENT-AVERAGED ROW PROJECTIONS ALGORITHM

In this paper, we consider the linear system of  $n$  linear equations in  $n$  variables:

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad 1 \leq i \leq n, \quad (1)$$

which can be written in the matrix form as

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{C}^{n \times n}, \quad \mathbf{x}, \mathbf{b} \in \mathbb{C}^n. \quad (2)$$

Matrix  $\mathbf{A}$  is the so-called impedance matrix, and the right-hand side corresponds to a seismic source.

#### The Kaczmarz method

The basis of the CARP-CG method is the Kaczmarz method (Kaczmarz, 1937) with a relaxation parameter. Starting from an arbitrary initial point, the Kaczmarz method cyclically projects the iterate  $\mathbf{x}^k$  onto one equation of system 2 to get the new iterate  $\mathbf{x}^{k+1}$ :

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \gamma \frac{b_{i(k)} - \langle \mathbf{a}_{i(k)*}, \mathbf{x}^k \rangle}{\|\mathbf{a}_{i(k)*}\|_2^2} \mathbf{a}_{i(k)*}, \quad (3)$$

$$i(k) = k \bmod n + 1,$$

where the column vector  $\mathbf{a}_{i*}$  denotes the  $i$ th row of  $\mathbf{A}$ ,  $\langle \cdot, \cdot \rangle$  denotes the scalar product of two vectors in  $\mathbb{C}^n$ , and  $0 < \gamma < 2$  is a relaxation parameter.

For the sake of implementation efficiency, we assume that the equations are normalized; that is, for  $1 \leq i \leq n$ , the  $i$ th equation has been divided by the norm  $\|\mathbf{a}_{i*}\|_2^2$ . Such an operation is also referred as *geometric row scaling*, which is a simple and useful diagonal preconditioner for nonsymmetric problems with discontinuous coefficients (Gordon and Gordon, 2010b). Throughout the rest of the paper, we assume that this normalization has been performed.

The Kaczmarz method cyclically performs the projections described by equation 3 until a certain convergence criterion is satisfied. A sequence of such projections performed from the first to the

last equations is referred to as a *forward Kaczmarz sweep*. A *backward sweep* is a sequence of projections in the reverse order. A *double sweep* is defined as a forward sweep followed by a backward sweep.

Note that the Kaczmarz method is actually equivalent to the successive overrelaxation (SOR) method (Saad, 2003) applied to the “normal equations” derived from system 2:

$$\mathbf{A}\mathbf{A}^H\mathbf{y} = \mathbf{b}, \quad \mathbf{x} = \mathbf{A}^H\mathbf{y}, \quad (4)$$

where  $H$  denotes the conjugate transpose operator. Performing the Kaczmarz method in double sweeps amounts to applying the symmetric SOR (SSOR) (Saad, 2003) method to the normal equations 4.

#### The conjugate gradient minimum norm method

Being based on the normal equations, the Kaczmarz method is robust and its convergence is guaranteed for consistent systems (Gordon and Gordon, 2005). However, its convergence rate can be extremely slow. Björck and Elfving (1979) propose the conjugate-gradient minimum norm (CGMN) method as a CG acceleration of the Kaczmarz method performed in double sweeps.

If we denote by  $\mathbf{S}_i$  the symmetric matrix  $\mathbf{S}_i = \mathbf{a}_{i*}\mathbf{a}_{i*}^H$  and set

$$\mathbf{Q}_i = \mathbf{I} - \gamma\mathbf{S}_i, \quad \mathbf{M}_i = \gamma\mathbf{a}_{i*}\mathbf{e}_i^T / \|\mathbf{a}_{i*}\|_2^2, \quad (5)$$

where  $\mathbf{e}_i, i = 1 \dots n$  is the canonical basis of  $\mathbb{R}^n$ , we can represent a single projection 3 by the following operation:

$$\mathbf{x}' = \mathbf{Q}_i\mathbf{x} + \mathbf{M}_i\mathbf{b}_i. \quad (6)$$

Hence, a double-sweep operation can be formulated as

$$\mathbf{x}^{k+1} = \mathbf{Q}\mathbf{x}^k + \mathbf{R}\mathbf{b}, \quad (7)$$

with

$$\mathbf{Q} = \mathbf{Q}_1 \cdots \mathbf{Q}_n \mathbf{Q}_n \cdots \mathbf{Q}_1,$$

$$\mathbf{R} = \mathbf{Q}_1 \cdots \mathbf{Q}_n \left( \sum_{i=1}^n \mathbf{Q}_n \cdots \mathbf{Q}_{i+1} \mathbf{M}_i \right) + \sum_{i=1}^n \mathbf{Q}_1 \cdots \mathbf{Q}_{i-1} \mathbf{M}_i. \quad (8)$$

If we consider equation 7 to be a fixed-point iteration method, we obtain the linear system

$$(\mathbf{I} - \mathbf{Q})\mathbf{x} = \mathbf{R}\mathbf{b}, \quad (9)$$

where the matrix  $\mathbf{I} - \mathbf{Q}$  is symmetric and positive semidefinite (Gordon and Gordon, 2010a). Linear system 9 can be solved using the CG method. The matrix-vector products for  $\mathbf{I} - \mathbf{Q}$  are performed through the double-sweep operations. The resulting method is the so-called CGMN method.

#### The conjugate-gradient-accelerated component-averaged row projections method

The CGMN method depends on double-sweep operations, which are inherently sequential. However, for high-performance comput-

ing applications such as the one involved in seismic imaging problems, it is mandatory for the solver to be parallel and scalable. The CARP-CG method is a parallelization of CGMN.

The impedance matrix is divided into row blocks,  $\mathbf{A}^1, \dots, \mathbf{A}^t$ . We denote by  $\mathbf{A}^q, \mathbf{b}^q$  the matrix block and the right-hand side corresponding to  $q$ ,  $1 \leq q \leq t$ . For  $1 \leq j \leq n$ , we denote

$$I_j = \{q | q \in \mathbb{N}, \exists i \in \mathbb{N}: (\mathbf{A}^q)_{i,j} \neq 0\}. \quad (10)$$

Therefore, the number of blocks that contain at least one equation with a nonzero coefficient of  $x_j$  is  $s_j = |I_j|$ , where  $|\cdot|$  denotes the cardinality of  $I_j$ .

Given the current iterate  $\mathbf{x}^k$ , the sweep operation in each block can be performed independently by different cores. This generates  $t$  different vectors  $\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^t$  in parallel. The CA operator acts as a map, merging these  $t$  vectors into a new vector. The  $j$ th component of the new vector is defined as follows:

$$\text{CA}(\mathbf{y}^1, \dots, \mathbf{y}^t)_j = \frac{1}{s_j} \sum_{q \in I_j} y_j^q, \quad 1 \leq j \leq n. \quad (11)$$

We follow the procedure used in [Gordon and Gordon \(2010a\)](#) in which the forward and backward sweeps are followed by an averaging operation.

As is mentioned by [Gordon and Gordon \(2005\)](#), the CA operations are equivalent to performing orthogonal projections onto certain hyperplanes, so by representing these hyperplanes as equations and adding these so-called averaging equations to the system, the projections can be done as part of the Kaczmarz iterations. However, these new projections need to be done with a relaxation parameter equal to 1, whereas the other projections usually have different relaxation parameters. The original CGMN assumed a fixed relaxation parameter, and it was extended to allow cyclic relaxation parameters, called *CGMNC* ([Gordon and Gordon, 2010a](#)). CARP-CG in the regular space is mathematically equivalent to CGMNC in the extended space. Thus, CARP-CG shares the same robustness as the CGMNC method.

In our particular application, we use a fourth-order FD discretization of the wave equation. The grid points are ordered in the  $z, y, x$  manner ( $z, x$  for 2D); that is,

$$\underbrace{v_{1,1,1}, v_{1,1,2}, \dots, v_{1,1,n_z}}_{}, \underbrace{v_{1,2,1}, \dots, v_{1,2,n_z}}_{}, \dots, \underbrace{v_{1,n_y,1}, \dots, v_{1,n_y,n_z}}_{}, \dots, \underbrace{v_{n_x,n_y,1}, \dots, v_{n_x,n_y,n_z}}_{}, \quad (12)$$

where  $v$  is, for instance, the horizontal particle velocity field. Hence, the partitioning of the matrix is equivalent to a division in the  $x$ -axis in terms of the physical domain. Because most of the 3D model has smaller  $y$ - and  $z$ -dimensions, such a division yields an impedance matrix with a smaller bandwidth. The number of rows in the matrix block is a multiple of the number of grid points in the corresponding subdomain. Each subdomain is assigned to one core. Because the derived impedance matrix is sparse and banded and the bandwidth is small enough, the components should be only shared by adjacent matrix blocks, which means that the averaging operations are limited to the subdomains that share a boundary. Therefore, the communication is limited to two cores, which will lead to good scalability of CARP-CG. Because the bandwidth of the matrix is related

to the order of the FD scheme, a high-order staggered-grid FD scheme may not be adapted for using CARP-CG. Although one could resort to a compact high-order FD scheme to reduce the bandwidth ([Britt et al., 2011](#); [Turkel et al., 2013](#)), it is difficult to extend such a scheme to the staggered-grid case, especially for the elastic wave equations. In our case, we show that a fourth-order scheme yields a sufficiently narrow banded matrix for CARP-CG to be scalable up to hundreds of cores.

## FINITE-DIFFERENCE DISCRETIZATION

### 3D elastic wave equation

We consider the 3D frequency-domain velocity-stress elastic system for the heterogeneous orthorhombic media with an explosive source  $\mathbf{s}$  applied on the normal stress components

$$\begin{cases} i\omega\rho v_x = \partial_x \tau_{xx} + \partial_y \tau_{xy} + \partial_z \tau_{xz} + f_x, \\ i\omega\rho v_y = \partial_x \tau_{xy} + \partial_y \tau_{yy} + \partial_z \tau_{yz} + f_y, \\ i\omega\rho v_z = \partial_x \tau_{xz} + \partial_y \tau_{yz} + \partial_z \tau_{zz} + f_z, \\ i\omega\tau_{xx} = C_{11}\partial_x v_x + C_{12}\partial_y v_y + C_{13}\partial_z v_z + i\omega s_1, \\ i\omega\tau_{yy} = C_{12}\partial_x v_x + C_{22}\partial_y v_y + C_{23}\partial_z v_z + i\omega s_2, \\ i\omega\tau_{zz} = C_{13}\partial_x v_x + C_{23}\partial_y v_y + C_{33}\partial_z v_z + i\omega s_3, \\ i\omega\tau_{yz} = C_{44}(\partial_y v_z + \partial_z v_y), \\ i\omega\tau_{xz} = C_{55}(\partial_x v_z + \partial_z v_x), \\ i\omega\tau_{xy} = C_{66}(\partial_x v_y + \partial_y v_x). \end{cases} \quad (13)$$

In equation 13,  $i$  denotes the imaginary unit,  $\mathbf{v} = (v_x, v_y, v_z)$  is the particle velocity vector,  $\boldsymbol{\tau} = (\tau_{xx}, \tau_{yy}, \tau_{zz}, \tau_{yz}, \tau_{xz}, \tau_{xy})$  is the vector that gathers the stress tensor components,  $\rho$  is the density,  $(f_x, f_y, f_z)$  are the external forces, and  $(s_1, s_2, s_3)$  are increments of the normal stresses that are related to an explosive source. For orthorhombic media, the stiffness matrix is described by nine independent parameters, such that

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}. \quad (14)$$

The reason we consider orthorhombic media is that the corresponding stiffness matrix has the same structure as for vertical transverse isotropic media and isotropic media. Therefore, we can easily switch between different types of media without modifying the FD scheme.

We use the fourth-order staggered-grid FD scheme ([Virieux, 1986](#); [Levander, 1988](#)) to discretize system 13. The corresponding stencil is illustrated in Figure 1. For instance, the first  $x$ -derivative of  $v_x$  on the staggered grids is approximated as

$$\begin{aligned} [\partial_x v_x]_{i,j,k} &\approx \frac{1}{\Delta} [c_1 (v_{xi+1/2,j,k} - v_{xi-1/2,j,k}) \\ &\quad + c_2 (v_{xi+3/2,j,k} - v_{xi-3/2,j,k})], \end{aligned} \quad (15)$$

where  $c_1 = 9/8$ ,  $c_2 = -(1/24)$ , and  $\Delta$  is the discretization spatial step. Following the parsimonious approach proposed by [Luo and Schuster \(1990\)](#), we can eliminate the stress terms to derive a FD scheme associated to the second-order elastic wave equation based

on the particle velocities only. This efficiently decreases the size of the linear system to be solved. Nevertheless, the underlying FD grid is still staggered, preserving the capability of handling high-contrast velocity and high-Poisson's-ratio media.

### Antilumped mass strategy

To decrease the numerical dispersion and the required number of grid points per wavelength, we apply an antilumped mass strategy following Marfurt (1984) and Jo et al. (1996). As a result, the mass acceleration term is approximated by a linear combination of the

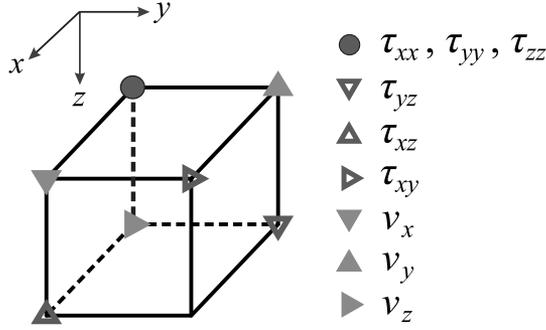


Figure 1. 3D staggered-grid finite-difference stencil.

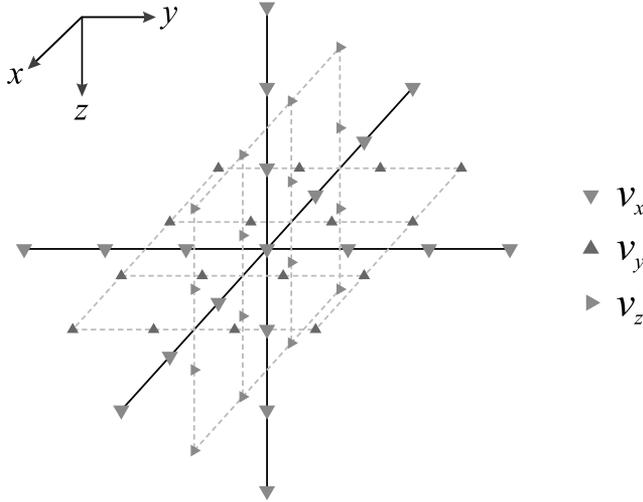


Figure 2. Grid points involved with  $v_{xi+1/2,j,k}$  for the mass acceleration term. The  $v_{xi+1/2,j,k}$  locates at the center point. The other  $v_x$  points involved are all along the axes through  $v_{xi+1/2,j,k}$ . The  $v_y$  points are on the  $x$ - $y$  plane, and the  $v_z$  points are on the  $x$ - $z$  plane. The grid points corresponding to the stress components are not shown here.

lumped mass and the consistent mass matrix operators. The mass acceleration term is distributed over the 19 grid points involved in the stencil (Figure 2). We take the horizontal particle velocity  $v_x$  as an example. The mass acceleration term is approximated as

$$\begin{aligned}
 (\rho v_x)_{i+1/2,j,k} = & a_1(\rho v_x)_{i+1/2,j,k} \\
 & + a_2[(\rho v_x)_{i+3/2,j,k} + (\rho v_x)_{i-1/2,j,k} + (\rho v_x)_{i+1/2,j+1,k} \\
 & + (\rho v_x)_{i+1/2,j-1,k} + (\rho v_x)_{i+1/2,j,k+1} + (\rho v_x)_{i+1/2,j,k-1}] \\
 & + a_3[(\rho v_x)_{i+5/2,j,k} + (\rho v_x)_{i-3/2,j,k} + (\rho v_x)_{i+1/2,j+2,k} \\
 & + (\rho v_x)_{i+1/2,j-2,k} + (\rho v_x)_{i+1/2,j,k+2} + (\rho v_x)_{i+1/2,j,k-2}] \\
 & + a_4[(\rho v_x)_{i+7/2,j,k} + (\rho v_x)_{i-5/2,j,k} + (\rho v_x)_{i+1/2,j+3,k} \\
 & + (\rho v_x)_{i+1/2,j-3,k} + (\rho v_x)_{i+1/2,j,k+3} + (\rho v_x)_{i+1/2,j,k-3}],
 \end{aligned} \tag{16}$$

where  $a_1, a_2, a_3$  and  $a_4$  are the weighting coefficients.

The weighting coefficients  $a_i, i = 1, \dots, 4$  and the fourth-order coefficients  $c_i, i = 1, 2$  are determined by solving an optimization problem that minimizes the misfit between the normalized phase velocities and the unity (Min et al., 2000). The Levenberg-Marquardt algorithm is applied to solve this optimization problem (Gosselin-Cliche and Giroux, 2014). The parameters involved in the optimization are Poisson's ratio  $\sigma$ ; the propagation angles  $\theta$  and  $\phi$  with respect to the  $x$ -axis and the  $z$ -axis, respectively; and the number of grid points per shear wavelength  $G_S$ . The optimization is performed with  $\sigma$  and  $1/G_S$  ranging from 0.01 to 0.33 in steps of 0.01, and with  $\theta$  and  $\phi$  ranging from 0 to  $\pi/4$  in steps of  $\pi/12$ . The optimal coefficients obtained are given in Table 2. Dispersion curves of the normalized phase and group velocities for compressional and shear waves are shown in Figures 3 and 4 for  $\sigma$  equal to 0.25 and 0.49, and for  $\theta$  and  $\phi$  ranging from 0 to  $\pi/4$ . The normalized velocities are plotted as a function of  $1/G_S$ .

In Figures 3 and 4, we observe that the fourth-order scheme with optimal coefficients yields better results than the conventional scheme for the phase and group velocities. The normalized phase velocities using the optimal coefficients are less dispersive and less anisotropic. In addition, the dispersion curves behave similarly for different Poisson's ratios, which is consistent with an expected insensitivity of the staggered-grid scheme to the Poisson's ratio (Levander, 1988). The error for the phase velocities is below 1% by using 3.3 grid points per minimum shear wavelength (Figure 3). Despite the fact that the group velocities are more dispersive and anisotropic than the phase velocities, the error still remains lower than 1% by using four grid points per minimum shear wavelength (Figure 4). Finally, for a given shear-wave velocity model, fewer grid points per wavelength are required, which leads to a reduction

Table 2. Optimal coefficient values obtained by minimizing the misfit between the normalized phase velocity and unity.

	$a_1$	$a_2$	$a_3$	$a_4$	$c_1$	$c_2$
Conventional	1	0	0	0	9/8	-1/24
Optimized	0.926585	6.99990E - 2	-3.28692E - 2	6.28252E - 3	1.25274	-5.45995E - 2

Figure 3. Dispersion curves of normalized P- and S-wave *phase* velocities using the conventional coefficients (in gray) and the optimal coefficients (in black). The curves are plotted for propagation angles with respect to the  $x$ - and the  $z$ -axes ranging from 0 to  $\pi/4$ , and for Poisson's ratios equal to  $\sigma = 0.25$  (left panels) and 0.49 (right panels).  $G_S$  is the number of grid points per minimum S-wavelength.

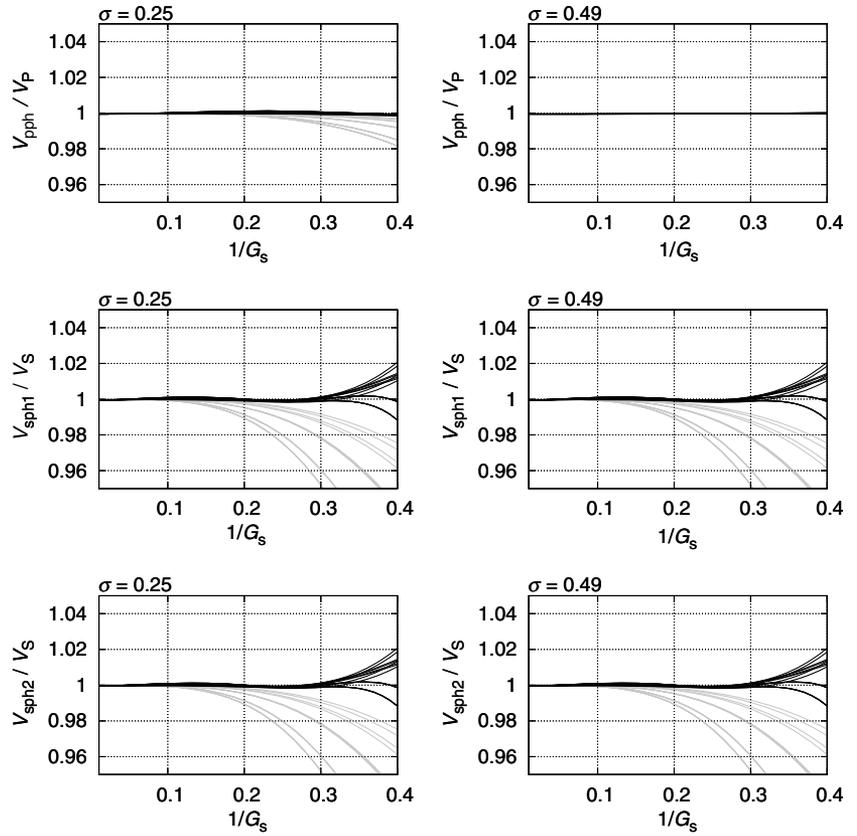
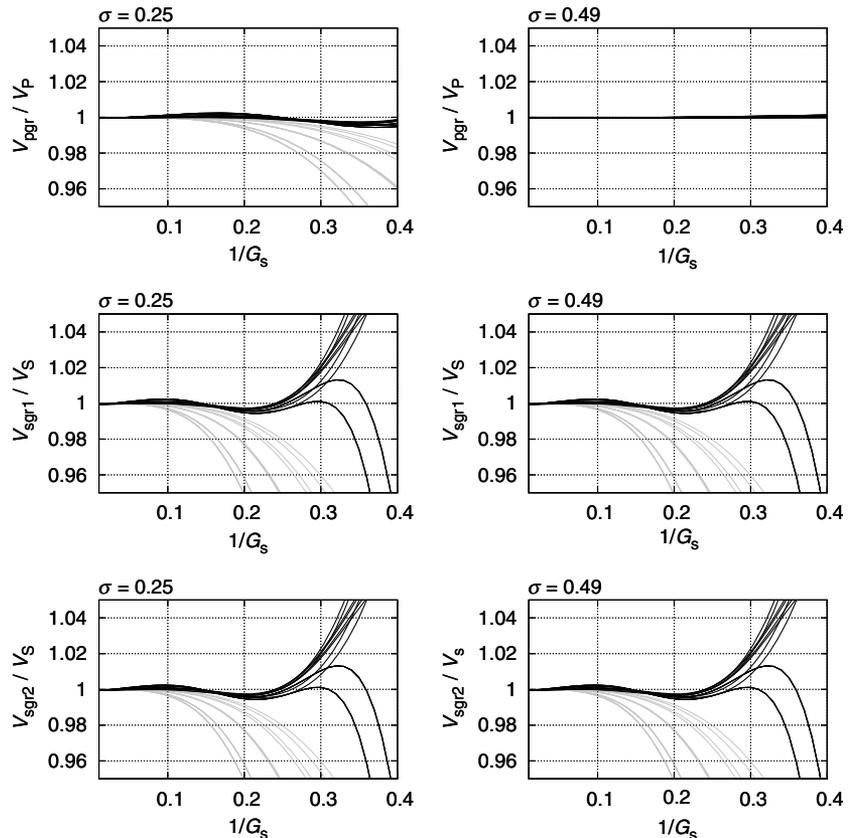


Figure 4. Dispersion curves of normalized P- and S-wave *group* velocities using the conventional coefficients (in gray) and the optimal coefficients (in black). The curves are plotted for propagation angles with respect to the  $x$ -s and  $z$ -axes ranging from 0 to  $\pi/4$ , and for the Poisson's ratios equal to  $\sigma = 0.25$  (left panels) and 0.49 (right panels);  $G_S$  is the number of grid points per minimum S-wavelength.



of the computational cost as the size of the corresponding linear system is decreased.

## NUMERICAL EXPERIMENTS

Several numerical experiments are presented to illustrate the performance of the CARP-CG method for the 2D and 3D frequency-domain elastic wave modeling in different types of media. All the computations are conducted on the Grenoble Alpes University High Performance Computing Center (CIMENT). Each node we use has two Intel Xeon E5-2670 2.6 GHz octacore CPUs, that is, 16 computing cores, and 64 GB of memory.

### Influence of physical factors

In this section, we investigate the influence of different physical factors on the convergence of CARP-CG and the condition number of the underlying system  $\mathbf{I} - \mathbf{Q}$ . We study the sensitivity to the following:

- Poisson's ratio
- FSBC
- seismic attenuation.

The wave modeling is carried out in 2D using the elastic Marmousi2 model. The P- and S-wave velocities, Poisson's ratio, and density of the original model are given in Figure 5. The size of the model is  $3.5 \times 17$  km, with a 450-m-deep water layer. The maximal Poisson's ratio in the water-wet sand is approximately 0.48, corresponding to a minimum S-wave velocity of approximately 270 m/s. As can be seen in Figure 5, the models have complex geometries, fast variations, and discontinuities. The resulting impedance matrix is ill conditioned, and the solution of this linear system with an iterative solver is challenging. The stopping criterion for the iterations is based on the relative residual:

$$\frac{\|\mathbf{b} - \mathbf{A}\mathbf{x}^k\|}{\|\mathbf{b} - \mathbf{A}\mathbf{x}^0\|} < 10^{-4}, \quad (17)$$

where  $\mathbf{x}^0$  is the starting point, which is set to  $\mathbf{0}$  in the following experiments. Perfectly matched layers (PMLs) are used as absorbing boundary conditions (Bérenger, 1994).

Note that in the numerical implementations, the number of matrix blocks is always equal to the number of cores. The relaxation parameter used for the Kaczmarz projections is set to 1, which is the optimal value selected from several experiments.

#### Poisson's ratio

Poisson's ratio, denoted as  $\sigma$ , is an important factor for describing an elastic medium. It is related to the P- and S-wave velocities through

$$\frac{V_S}{V_P} = \sqrt{\frac{0.5 - \sigma}{1 - \sigma}}. \quad (18)$$

If  $\sigma = 0.5$ , we see that  $V_S = 0$ , which corresponds to an acoustic medium. A Poisson's ratio close to 0.5 corresponds to soft materials where the ratio of  $V_S/V_P$  is small. This discrepancy between  $V_S$  and  $V_P$  enlarges the size of the resulting linear system and yields more complex problems to solve, as shear waves with small wavelength

are propagated. A lower Poisson's ratio, 0.25 for instance, corresponds to harder media for which the discrepancy between  $V_S$  and  $V_P$  is smaller, which yields linear systems that are easier to solve.

To investigate the influence of the Poisson's ratio, wave modeling is performed in three cases. In the first step, a constant Poisson's ratio of 0.25 is used. The ratio is increased to 0.4 in the second case. Finally, we use the space-dependent Poisson's ratio from the original Marmousi2 model. We keep the discretization setup with respect to the original Poisson's ratio for all the three cases. This is not the optimal option in practice because the grid could be coarser when considering an elastic model with a smaller Poisson's ratio. Here, we just want to rule out the influence of the grid size and focus on the effect of the Poisson's ratio.

Wave modeling is conducted on a Marmousi2 model discretized with a  $351 \times 1701$  grid and a spatial step equal to 10 m. The frequency is set to 5 Hz. An explosive Dirac source located at  $x = 8500$  m,  $z = 100$  m is used. The computation time and the number of iterations are summarized in Table 3. All results are obtained using 128 cores. The convergence curves for different Poisson's ratios are presented in Figure 6. The results indicates that CARP-CG requires more iterations to converge as the Poisson's ratio increases and even more when the spatial variations in Poisson's ratio are included. This can be related to the conditioning of matrix  $\mathbf{I} - \mathbf{Q}$ . Because calculating the eigenvalues is a computationally heavy problem itself, we extract a small part of the model (size  $36 \times 73$ ), marked by the gray frame in Figure 5, to form the impedance matrix  $\mathbf{A}$  and the corresponding matrix  $\mathbf{I} - \mathbf{Q}$ . The distribution of the eigenvalues and the condition numbers of  $\mathbf{I} - \mathbf{Q}$  for these different configurations are presented in Figure 6 and Table 3, respectively. The eigenvalues exhibit similar distributions for  $\sigma = 0.4$  and the variable  $\sigma$ . However, the condition numbers, which are determined by the ratio between the maximum and the minimum eigenvalues, are very different. The ratios  $\text{Cond}_{(\text{variable } \sigma)} / \text{Cond}_{(\sigma=0.4)} = 2.18$  and  $\text{Cond}_{(\sigma=0.4)} / \text{Cond}_{(\sigma=0.25)} = 1.47$ . This explains the larger deviation between the convergence trend for  $\sigma = 0.4$  and the variable  $\sigma$ , compared to the deviation between

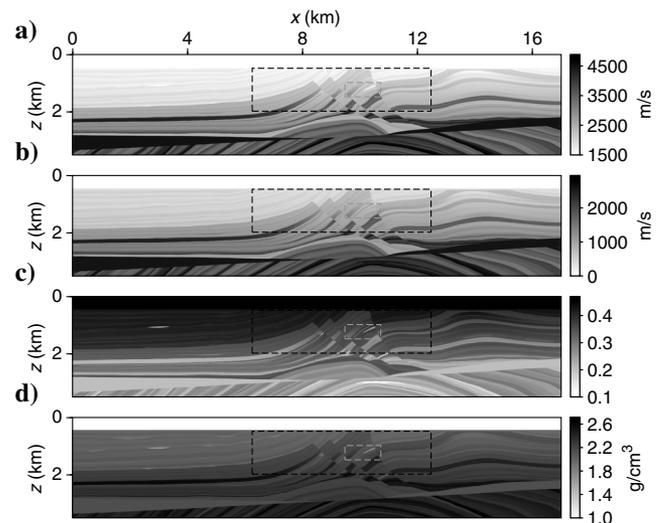


Figure 5. The Marmousi2 model (Martin et al., 2006): (a) P-wave velocity, (b) S-wave velocity, (c) Poisson's ratio, and (d) density. The big part in the black frame is used for comparisons with GMRES, CGNR, and BiCGStab. The small part in the gray frame is for the calculation of the eigenvalues.

$\sigma = 0.25$  and  $\sigma = 0.4$ . Hence, the complexity that arises from the variable Poisson's ratio seems to have a strong influence on the conditioning of the impedance matrix and the convergence of the CARP-CG method.

*Free-surface boundary condition*

A proper implementation of the FSBC is one of the crucial problem in elastic wave modeling. Because the staggered-grid scheme uses the first-order velocity-stress formulation of the elastodynamic equations, stress-free conditions at the free surface can be easily

**Table 3. Computation time and number of iterations  $N_{iter}$  for the wave modeling with different Poisson's ratios. We use the Marmousi2 model (Figure 5) of size  $3.5 \times 17$  km sampled on a  $351 \times 1701$  grid. The spatial step size is 10 m, the frequency is 5 Hz, and the number of cores is 128. The condition numbers of the matrix  $\mathbf{I} - \mathbf{Q}$  calculated from a small part of the Marmousi2 model are listed in the last column.**

$\sigma$	Time (s)	$N_{iter}$	Condition number
0.25	110.25	12,634	$1.08 \times 10^5$
0.4	126.56	14,491	$1.59 \times 10^6$
Variable	169.59	19,451	$3.45 \times 10^6$

described. We use an accurate and simple implementation of the elastic free surface for the staggered-grid scheme proposed by [Mittel \(2002\)](#). Unlike the imaging method ([Levander, 1988](#); [Graves, 1996](#)), this approach does not require additional grid points above the free surface; hence, it keeps the size of the linear system unchanged.

To illustrate the influence of the FSBC, a comparative experiment with the PML on the surface is performed. PMLs are attached to the other three boundaries of the model in both cases. We use the numerical settings described in the last section, retaining the variable Poisson's ratio from the original Marmousi2 model. Analysis of the condition numbers of the matrix  $\mathbf{I} - \mathbf{Q}$  with and without the FSBC is carried out as well. The convergence curves and the distribution of the eigenvalues are presented in Figure 7. The computation time, number of iterations, and condition numbers for each case are summarized in Table 4. Two representative wavefields of the horizontal particle velocity  $v_x$  are given in Figure 8.

Applying the FSBC generates the propagation of surface waves interacting with the P- and S-waves. The influence of the surface waves on the wavefields can be seen in Figure 8. Not surprisingly, this seems to strongly affect the conditioning of the impedance matrix. As shown in Figure 7, despite the maximum eigenvalues being very close to each other, integrating the FSBC generates a significant drop in the amplitude of the smallest eigenvalues. This leads to a condition number of  $\mathbf{I} - \mathbf{Q}$ , which is approximately 20 times larger when implementing the FSBC (Figure 7):  $7.24 \times 10^7$  for the FSBC and  $3.45 \times 10^6$  for the PML surface. The slower convergence rate observed for the experiments with the FSBC is thus consistent with the analysis of the condition number.

Figure 6. (a) Convergence curves for wave modeling at 5 Hz and (b) the eigenvalue distribution of the matrix  $\mathbf{I} - \mathbf{Q}$  calculated from a small part of the Marmousi2 model for different Poisson's ratios. The vertical axis of panel (a) and both axes of panel (b) are plotted on log scales. The horizontal axis of panel (b) contains the indexes of the eigenvalues, and the vertical axis corresponds to the values of the eigenvalues.

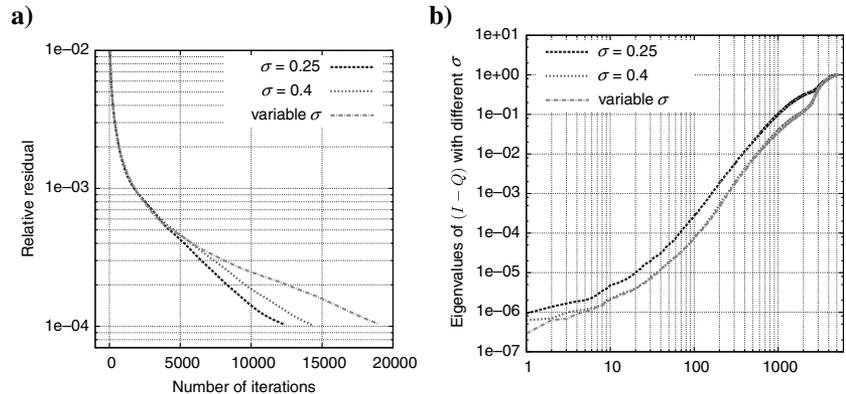
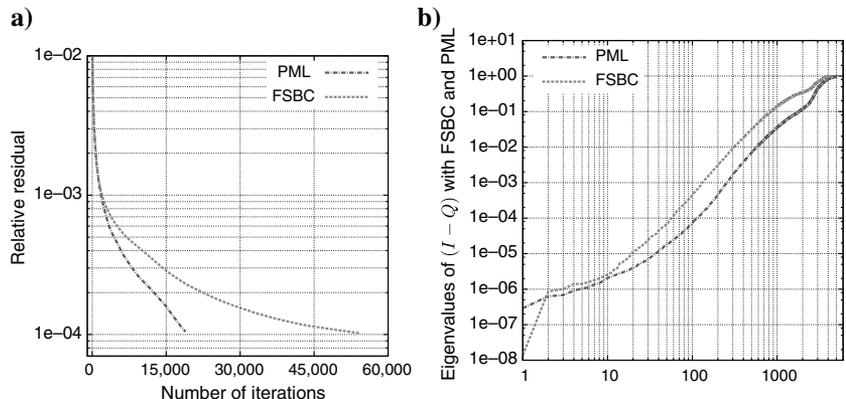


Figure 7. (a) Convergence curves for the wave modeling at 5 Hz and (b) eigenvalue distribution of the matrix  $\mathbf{I} - \mathbf{Q}$  calculated from a small part of the Marmousi2 model, with FSBC and PML as the surface boundary condition, respectively. The vertical axis of panel (a), and both axes of panel (b) are plotted on log scales. The horizontal axis of panel (b) contains the indexes of the eigenvalues, and the vertical axis corresponds to the values of the eigenvalues.



Seismic attenuation

Compared to time-domain wave modeling, incorporating the intrinsic attenuation of the medium is straightforward in the frequency domain by using complex-valued wave velocities. In this study, we use the Kolsky-Futterman model without the dispersion term such that

$$\bar{c}(\omega) = c \left( 1 - i \operatorname{sign}(\omega) \frac{1}{2Q_{\text{atten}}} \right), \quad (19)$$

where  $c$  is the velocity of the P- or S-wave,  $i$  denotes the imaginary unit,  $\omega$  is the angular frequency, and  $Q_{\text{atten}}$  is the quality factor (Kolsky, 1956; Futterman, 1962). Note that the smaller that  $Q_{\text{atten}}$  is, the

**Table 4. Computation time and number of iterations  $N_{\text{iter}}$  for wave modeling with the PML and FSBC as the surface boundary conditions. The numerical settings are the same as in Table 3. The condition numbers of the matrix  $\mathbf{I} - \mathbf{Q}$  calculated from a small part of the Marmousi2 model are listed in the last column.**

	Time (s)	$N_{\text{iter}}$	Condition number
FSBC	473.53	56,602	$7.24 \times 10^7$
PML	170.53	19,451	$3.45 \times 10^6$

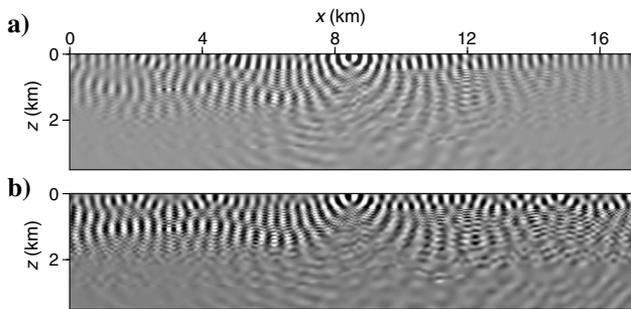


Figure 8. Frequency-domain horizontal particle velocity  $v_x$  wavefields at 5 Hz with (a) the PML surface, and (b) the FSBC.

more attenuating the medium is, whereas using larger values for  $Q_{\text{atten}}$  corresponds to almost nonattenuating medium.

Several experiments considering different levels of attenuation are performed to investigate the influence of seismic attenuation on the impedance matrix conditioning and on the convergence rate of CARP-CG. The same numerical settings described in last sections are used, with the spatial-dependent Poisson’s ratio and the FSBC. In addition, analysis of the condition numbers of matrix  $\mathbf{I} - \mathbf{Q}$  derived from different values of  $Q_{\text{atten}}$  is conducted. The convergence curves and the distribution of the eigenvalues are presented in Figure 9. The computation time, number of iterations, and condition numbers for different  $Q_{\text{atten}}$  values are summarized in Table 5. The representative wavefields with quality factor  $Q_{\text{atten}}$  equal to 1, 10, 50, and 1000 are shown in Figure 10. The wavefields for  $Q_{\text{atten}} = 100, 10,000$  are omitted because their difference from  $Q_{\text{atten}} = 1000$  is negligible.

Figure 9 indicates that CARP-CG converges faster when more attenuation is incorporated. This can be related to the more clustered eigenvalues (Figure 9b), and the smaller condition number (Table 5), for the smaller  $Q_{\text{atten}}$  value. Compared to the modeling without considering attenuation, we observe a convergence that is 2.60 times faster for  $Q_{\text{atten}} = 50$  and 1.76 times faster for  $Q_{\text{atten}} = 100$ . This phenomenon is interesting when considering that attenuation is necessary for practical wave modeling and it accelerates the convergence of the iterative solver. Introducing attenuation adds a pure imaginary term on the diagonal of the impedance matrix. This improves the property of the linear system because the ratio of the amplitude between diagonal and off-diagonal terms increases as quality factor  $Q_{\text{atten}}$  decreases. This is consistent with the results from the applications of the complex shifted-Laplace preconditioning technique (Plessix, 2007; Riyanti et al., 2007; Erlangga and Nabben, 2008), in which an approximate inverse of a heavily damped wave equation acts as a preconditioner. Figure 10 illustrates the attenuating effect on the propagation of the wavefields.

Numerical efficiency

Scalability on multiple cores

We perform a scalability test using the Marmousi2 model. The frequencies increase from 1.25 to 20 Hz. We keep 5.4 grid points per minimum S-wavelength with respect to each frequency to ensure the accuracy of the FD scheme. The size of the linear system thus increases for each frequency. For each system, we increase the number of cores and compare the required computational time to

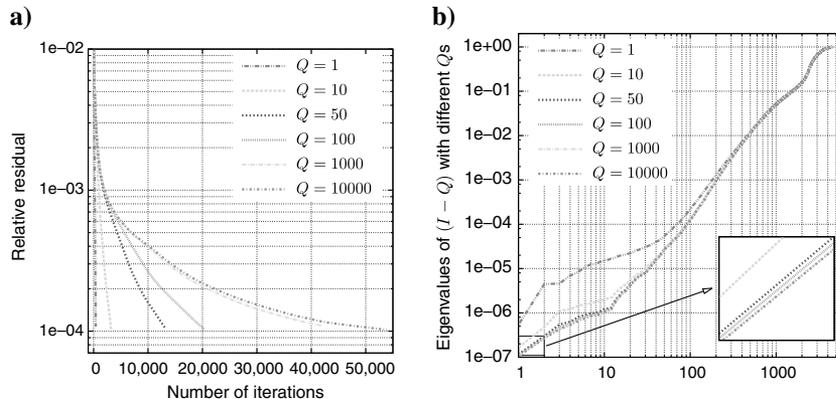


Figure 9. (a) Convergence curves for the wave modelings at 5 Hz and (b) eigenvalues distribution of the matrix  $\mathbf{I} - \mathbf{Q}$  calculated from a small part of the Marmousi2 model, with the quality factor  $Q_{\text{atten}}$  equal to 1, 10, 50, 100, 1000, and 10,000, respectively. The vertical axis of panel (a) and both axes of panel (b) are plotted on log scales. The horizontal axis of panel (b) contains the indices of the eigenvalues, and the vertical axis corresponds to the values of the eigenvalues.

reach the convergence. We use the space-dependent Poisson's ratio presented in Figure 5 and a quality factor of  $Q_{\text{atten}} = 1000$ . An explosive Dirac source is excited on the top middle part of the medium ( $x_S = 8500$  m,  $z_S = 100$  m).

We first conduct a set of experiments without the FSBC, using a single-precision code. The size of the FD grid, total number of unknowns, computing time, and number of iterations are summarized in Table 6. The size of the linear system increases from approximately  $1.2 \times 10^5$  to  $19.7 \times 10^6$  (Table 6). Figure 11 presents the scaling properties and the convergence curves for using different numbers of cores. The gray lines represent the ideal scalability for parallel computing, and the colored curves are for the CARP-CG method. For each frequency, the curve for CARP-CG is close to the ideal scalability for a small number of cores. For low-frequency problems (related to models of smaller size), using more cores generates more communications and degrades the scaling efficiency. However, for larger scale problems (e.g., 10, 20 Hz), the scaling

**Table 5. Computation time and number of iterations  $N_{\text{iter}}$  for the wave modeling with the FSBC and different levels of attenuation. The numerical settings are the same as in Table 3. The condition numbers of the matrix  $\mathbf{I} - \mathbf{Q}$  calculated from a small part of the Marmousi2 model are listed in the last column.**

$Q_{\text{atten}}$	Time (s)	$N_{\text{iter}}$	Condition number
1	3.67	412	$1.856 \times 10^6$
10	27.51	3269	$6.009 \times 10^6$
50	113.36	13,565	$8.522 \times 10^6$
100	173.10	20,729	$8.934 \times 10^6$
1000	406.55	48,622	$9.341 \times 10^6$
10,000	475.38	57,356	$9.384 \times 10^6$

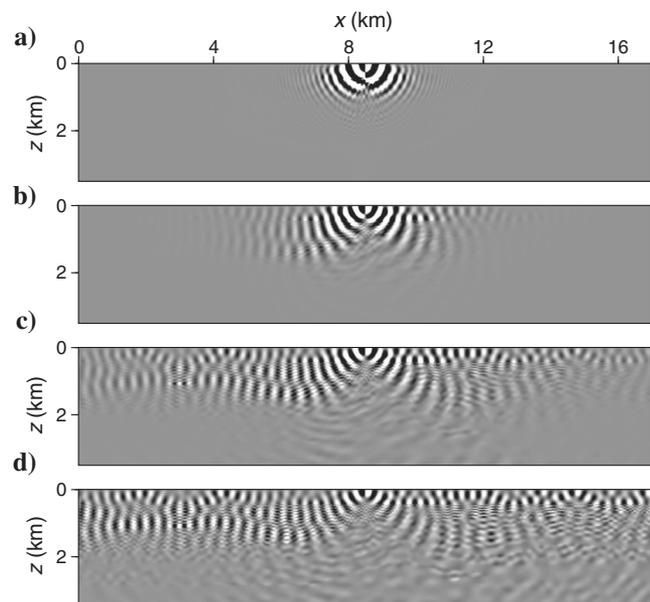


Figure 10. Frequency-domain horizontal particle velocity  $v_x$  wavefields at 5 Hz with quality factors of (a)  $Q_{\text{atten}} = 1$ , (b)  $Q_{\text{atten}} = 10$ , (c)  $Q_{\text{atten}} = 50$ , and (d)  $Q_{\text{atten}} = 1000$ .

remains close to the ideal scalability for using a large number of cores. Because we focus on the larger scale applications, the CARP-CG method seems encouraging. Note also that CARP-CG scales better for higher frequencies: When the number of cores increases, the percentage increase of the iteration numbers is smaller than in the case of the low frequencies (Gordon and Gordon, 2011). This could be verified by the more clustered convergence curves for high frequencies in Figure 11.

The results for the modeling at 20 Hz using a small number of cores are omitted in Figure 11. During the numerical simulation, a divergence occurs when using 4, 8, or 16 cores. This seems to be due to a lack of accuracy because this phenomenon is observed only in the single-precision implementation. This divergence was not observed when we conducted the same experiments in the double-precision implementation. However, it is interesting to note that CARP-CG converges again with the single-precision implementation when we increase the number of cores. When more cores are used, the sweeping operations are performed in smaller row blocks and more CA operations are performed. This suggests that these operations alleviate the accumulation of rounding error compared to sweeping in large row blocks.

However, when the FSBC is taken into account, the divergence for high frequencies (10 and 20 Hz) is observed systematically, even when the number of computation cores is increased. This could be due to the complexity arising from the FSBC, which may increase the accumulation of the round-off errors and counteract the effects of block division. We thus use a double-precision implementation in the following experiments to avoid the problem caused by the lack of accuracy. Although the memory cost doubles when moving from single to double precision, the computing time appears to increase much less than two times. For instance, for the 20-Hz modeling in the 2D Marmousi2 model with the FSBC and using 128 cores, the computing time is 4093.75 and 5503.54 s, respectively, for the single- and double-precision implementation. The corresponding experimental settings are presented in Table 7. Representative wavefields of the horizontal particle velocity  $v_x$  at each frequency are shown in Figure 12. The corresponding decrease of the average wavelength can be observed as the frequency increases. Figure 13 presents the scaling properties and the convergence histories for each frequency using different numbers of cores: A satisfactory scalability of CARP-CG is observed. The convergence curves are

**Table 6. Geometry parameters, total number of unknowns  $N_u$ , computing time, and number of iterations  $N_{\text{iter}}$  for the wave modeling at different frequencies without the FSBC. The time values are the minimum computing time we obtain, corresponding to the number of cores given in the  $N_{\text{proc}}$  column.**

Without FSBC						
$f$ (Hz)	$h$ (m)	$n_z \times n_x$	$N_u$	Time (s)	$N_{\text{proc}}$	$N_{\text{iter}}$
1.25	40	$88 \times 426$	119,296	12.19	64	9086
2.5	20	$176 \times 851$	384,912	34.79	64	9930
5	10	$351 \times 1701$	1,361,462	169.78	128	19,451
10	5	$701 \times 3401$	5,099,562	766.95	128	31,245
20	2.5	$1401 \times 6801$	19,715,762	4093.75	128	50,069

not as regular as in the last experiment due to the influence of the FSBC, but the more clustered curves for high frequencies further emphasizes the better scalability of CARP-CG for large-scale problems.

*Complexity analysis*

In Figure 14, we present a complexity analysis of the CARP-CG method in the previous two experiments. The number of iterations  $N_{iter}$  is plotted as a function of the geometric mean size of one dimension:

$$N = \sqrt[2]{N_x \times N_z}, \quad (20)$$

on a log-log scale. The solid line with squares indicates a linear dependence between  $N_{iter}$  and  $N$ ; that is,  $N_{iter} = \mathcal{O}(N)$ . The dotted line with triangles shows the results from the CARP-CG method without and with the FSBC. Performing a standard linear regression (indicated by the dotted-dashed line in Figure 14), we obtain two slopes equal to 0.59 and 0.65, respectively. This corresponds to a complexity such that

$$N_{iter} = \begin{cases} \mathcal{O}(N^{0.59}), & \text{without the FSBC,} \\ \mathcal{O}(N^{0.65}), & \text{with the FSBC.} \end{cases} \quad (21)$$

The irregularity of the complexity curve for the modeling with the FSBC may arise from the influence of the surface waves and eigenmodes appearing in same frequencies. Nevertheless, the slope of the two curves seems to be similar. Following the analysis in the introduction, a total computational complexity of CARP-CG for 2D modeling with one source is

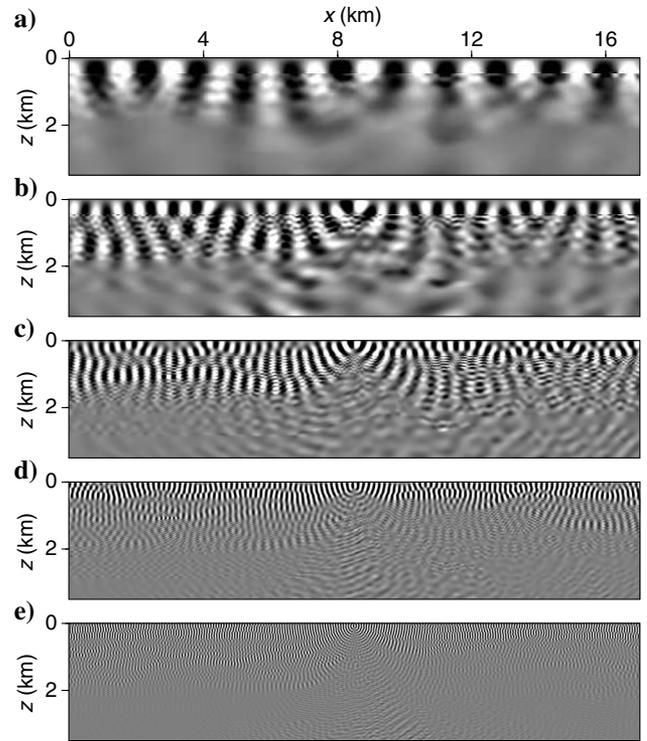


Figure 12. 2D frequency-domain wavefields in the Marmousi2 model at (a) 1.25, (b) 2.5, (c) 5, (d) 10, and (e) 20 Hz.

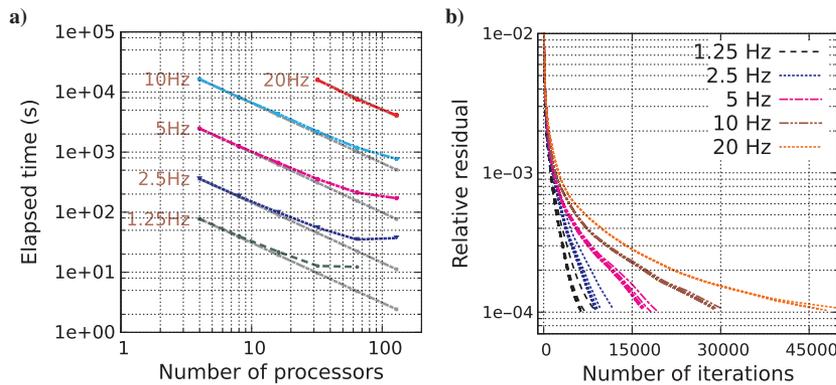


Figure 11. (a) Scaling properties and (b) convergence histories using different numbers of cores for the elastic wave modeling in single precision without the FSBC. The frequencies are 1.25, 5, 10, and 20 Hz, respectively. The gray lines represent the ideal scalability for parallel computing, and the colored curves are for the CARP-CG method.

**Table 7. Geometry parameters, total number of unknowns  $N_u$ , computing time, and number of iterations  $N_{iter}$  for wave modeling at different frequencies with the FSBC. The time values are the minimum computing time that we obtain, corresponding to the number of cores given in the  $N_{proc}$  column.**

With FSBC						
$f$ (Hz)	$h$ (m)	$n_z \times n_x$	$N_u$	Time (s)	$N_{proc}$	$N_{iter}$
1.25	40	$88 \times 426$	100,656	24.85	32	12,053
2.5	20	$176 \times 851$	349,272	143.3	64	27,915
5	10	$351 \times 1701$	1,291,822	556.32	128	42,728
10	5	$701 \times 3401$	4,961,922	1647.92	128	45,278
20	2.5	$1401 \times 6801$	19,442,122	7161.45	128	58,457

$$\mathcal{O}(N_{\text{iter}} \times N^2) = \begin{cases} \mathcal{O}(N^{2.59}), & \text{without the FSBC,} \\ \mathcal{O}(N^{2.65}), & \text{with the FSBC,} \end{cases} \quad (22)$$

representing an improvement compared to  $\mathcal{O}(N^3)$ , which corresponds to the combination of the 2D time-domain modeling and the Fourier transform (Nihei and Li, 2007).

*Comparison with the generalized minimum residual, conjugate gradient applied to the normal equations, and biconjugate gradient stabilized methods*

We compare the CARP-CG method with the restarted GMRES method with a restarted parameter of 50 (Saad, 1986; Frayssé et al., 2003), the CGNR method (Saad, 2003), the CGNR method (Saad,

2003), and the BiCGStab method (Van der Vorst, 1992; Sleijpen and Fokkema, 1993). To make a fair comparison, all of the methods are applied to the normalized linear system; that is, each equation is divided by the norm of the corresponding row vector. The experiments are conducted using one core. The CARP-CG method is thus equivalent to the CGMN method. We perform the wave modeling in a part of the Marmousi2 model (Figure 5) with the spatial-dependent Poisson's ratio and the FSBC at different frequencies. Table 8 summarizes the parameter settings, the number of iterations, and the computing time for each method. Figure 15 presents the convergence curves. We observe that the convergence of CGMN is achieved in all cases, whereas CGNR converges only for low frequencies, and GMRES and BiCGStab fail to converge for each frequency. GMRES gets stagnant after a few iterations. This could be caused by the lim-

Figure 13. Scaling properties and convergence histories using different numbers of cores for the elastic wave modeling in double precision with the FSBC. The frequencies are 1.25, 5, 10, and 20 Hz, respectively. The gray lines represent the ideal scalability for parallel computing, and the colored curves are for the CARP-CG method.

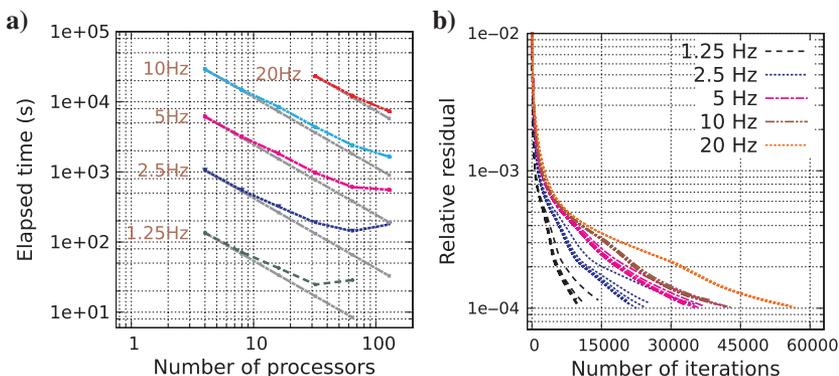
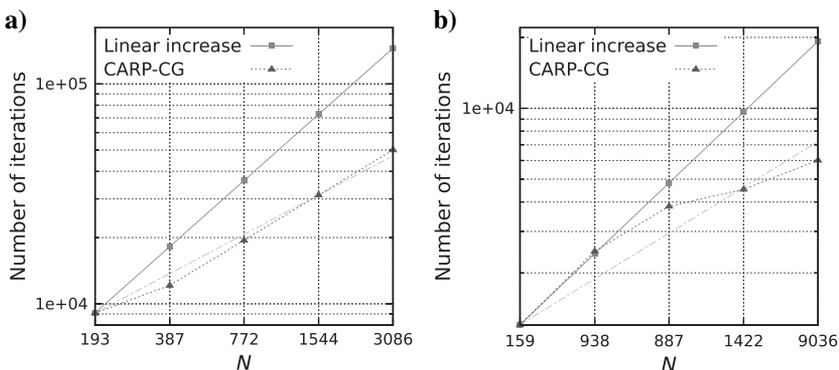


Figure 14. Complexity analysis of the CARP-CG method for the elastic wave modeling (a) without and (b) with the FSBC. The number of iterations  $N_{\text{iter}}$  is plotted as a function of the average model dimension  $N$  on a log-log scale. The straight dotted-dashed line indicates the results obtained from line regression.



**Table 8. Geometry parameters, total number of unknowns  $N_u$ , number of iterations  $N_{\text{iter}}$ , and computing time for the modeling in a part of the Marmousi2 model (marked by the black frame in Figure 5), using CGMN, GMRES, CGNR, and BiCGStab. The numbers in the bracket are the elapsed time (in s) for the convergence. The symbol “\*” indicates convergence achieved before  $10^5$  iterations, and the number with symbol “ $\diamond$ ” is the relative residual obtained after  $10^5$  iterations. The results for BiCGStab are omitted because the relative residual increases to greater than  $10^{10}$  after several iterations.**

$f$ (Hz)	$h$ (m)	$n_z \times n_x$	$N_u$	CGMN	GMRES	CGNR
1.25	40	$39 \times 157$	17,346	6826* (40.6)	$3.20E - 3 \diamond$	27,759* (94.6)
2.5	20	$76 \times 313$	57,276	10,961* (228.8)	$2.96E - 3 \diamond$	43,402* (493.0)
5	10	$151 \times 626$	208,012	17,083* (1257.6)	$1.91E - 3 \diamond$	66,524* (2781.7)
10	5	$301 \times 1251$	790,562	24,738* (6790.3)	$2.36E - 4 \diamond$	95,974* (15,410.2)

ited restart size. However, a too-large restart size will increase the requirement of memory cost and becomes unfeasible. BiCGStab exhibits severe instabilities (the relative residual reaches  $10^{10}$  after several iterations). CGNR seems to have a slow but stable convergence. This can be explained by the fact that CGNR is based on the normal equations, whose diagonal elements are relatively large, even when  $\mathbf{A}$  has large off-diagonal elements (Gordon and Gordon, 2008). This improves the robustness of the convergence, as is the

case for the CGMN method. However, the convergence rate of CGNR is extremely slow, whereas the CGMN method achieves faster convergence. The nonconvergence for GMRES, CGNR, and BiCGStab stems from the complexity of the Marmousi2 model. Numerical tests we do not present here show that the convergence of these methods can be achieved if a less complicated model, a homogeneous model for instance, is considered. In this case, an even-faster convergence rate can be obtained for the BiCGStab and GMRES.

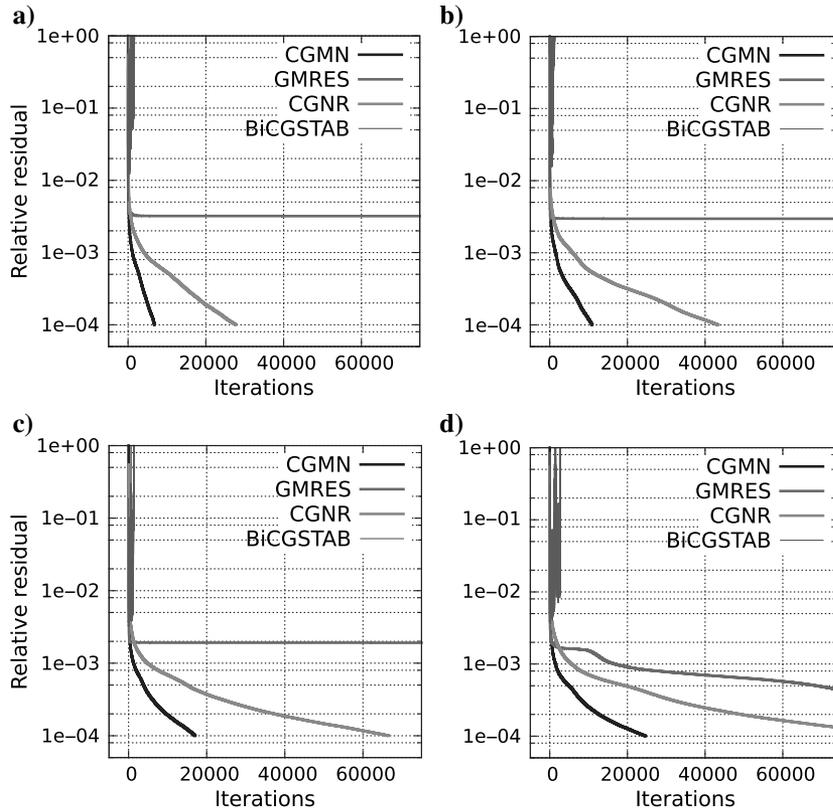


Figure 15. Convergence curves of CGMN, GMRES, CGNR, and BiCGStab for the wave modeling at (a) 1.25, (b) 2.5, (c) 5, and (d) 10 Hz.

### Test on 3D models

#### Homogeneous model

A set of 3D homogeneous models is used to investigate the performance of the CARP-CG method for the 3D applications. The P-wave velocity is set to 5000 m/s, and the S-wave velocity is 2886.75 m/s with respect to a constant Poisson’s ratio equal to 0.25. The grid varies from  $51 \times 51 \times 51$  to  $501 \times 501 \times 501$  as the frequency increases from 2.5 to 25 Hz. The number of grid points per minimum S-wave-length is kept to 4. The number of unknowns varies from  $1 \times 10^6$  to  $4 \times 10^8$ . PMLs are attached to the boundaries. An explosive Dirac source is located at the center of the model. The geometry parameters, total number of unknowns  $N_u$  (with  $N_{PML} = 10$ ), computing time, number of cores  $N_{proc}$ , estimated memory used per core, and number of iterations  $N_{iter}$  are summarized in Table 9. The convergence curves are presented in Figure 16a. Four representative wavefields of the horizontal particle velocity  $v_x$  at 10, 15, 20, and 25 Hz are presented in Figure 17. The increment of the wavenumber can be observed as the frequency goes from 10 to 25 Hz.

The same complexity analysis as in the previous section is carried out. The results are presented in Figure 16b. The number of iterations

Table 9. Geometry parameters, total number of unknowns  $N_u$ , computing time, number of cores  $N_{proc}$ , estimated memory used per core, and number of iterations  $N_{iter}$  for the 3D elastic wave modeling in a homogeneous media at different frequencies.

$f$ (Hz)	$h$ (m)	$n_x \times n_y \times n_z$	$N_u$	Time (s)	$N_{proc}$	Mem/core (MB)	$N_{iter}$
2.5	288	$51 \times 51 \times 51$	1,073,733	81.14	8	83.9	546
5	144	$101 \times 101 \times 101$	5,314,683	438.05	16	207.8	989
7.5	96	$151 \times 151 \times 151$	15,000,633	1244.46	24	391.0	1402
10	72	$201 \times 201 \times 201$	32,381,583	2777.40	32	633.1	1785
12.5	57	$251 \times 251 \times 251$	59 707 533	5202.29	40	933.8	2162
15	48	$301 \times 301 \times 301$	99,228,483	8697.03	48	1293.3	2488
17.5	41	$351 \times 351 \times 351$	153,194,433	13,932.63	56	1711.4	2817
20	36	$401 \times 401 \times 401$	223,855,383	20,611.18	64	2188.2	3101
22.5	32	$451 \times 451 \times 451$	313,461,333	29,323.91	72	2723.6	3389
25	28	$501 \times 501 \times 501$	424,262,283	41,838.52	80	3317.8	3755

Figure 16. (a) Convergence histories for the 3D elastic wave modeling in a homogeneous media at different frequencies. The frequency increases from 1.25 to 25 Hz. A complexity analysis for the 3D homogeneous elastic wave modelings is given in panel (b), in which the number of iterations  $N_{iter}$  is plotted as a function of the average model dimension  $N$  on a log-log scale. The straight dotted-dashed line in (b) indicates the results obtained from line regression.

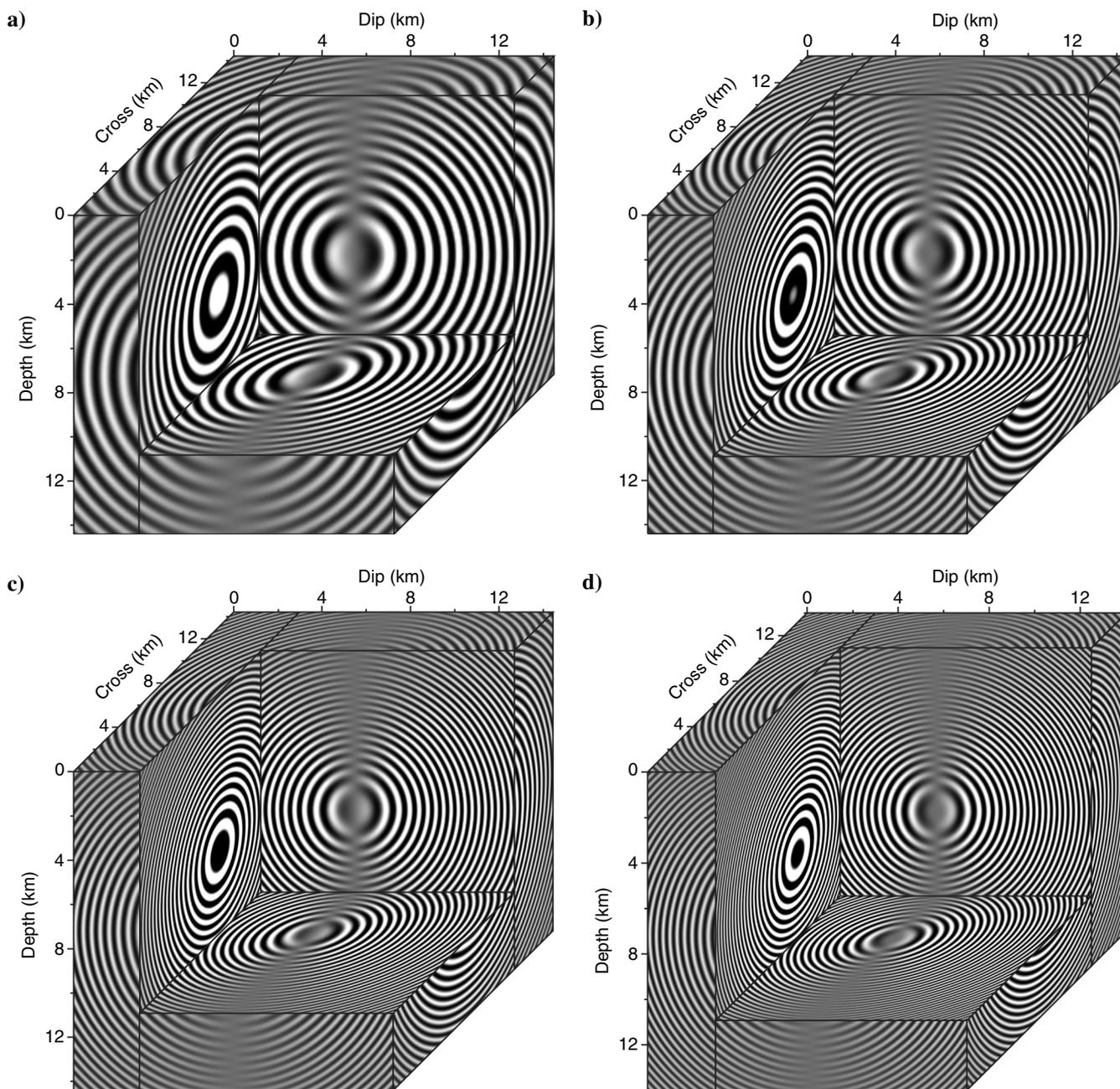
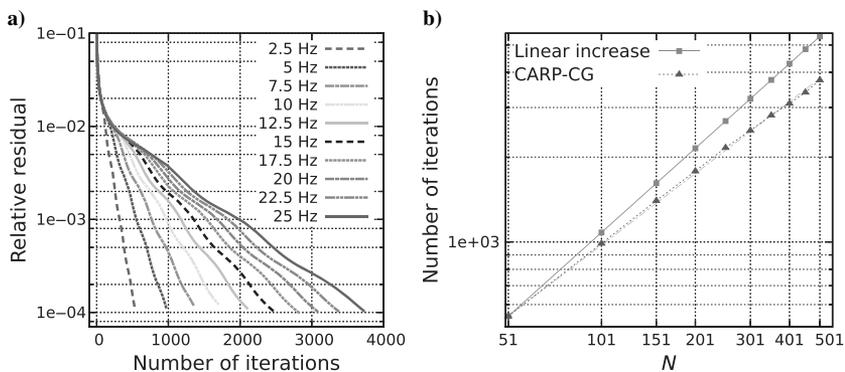


Figure 17. 3D frequency-domain elastic wavefields of the horizontal particle velocity  $v_x$  at (a) 10, (b) 15, (c) 20, and (d) 25 Hz.

$N_{\text{iter}}$  is plotted as a function of the geometric mean size of one dimension

$$N = \sqrt[3]{N_x \times N_y \times N_z}, \quad (23)$$

on a log-log scale. We obtain a slope of 0.84, leading to  $N_{\text{iter}} = \mathcal{O}(N^{0.84})$ . The total computational complexity for one source using CARP-CG is

$$\mathcal{O}(N_{\text{iter}} \times N^3) = \mathcal{O}(N^{3.84}), \quad (24)$$

which is more efficient than  $\mathcal{O}(N^4)$  for the combination of the 3D time-domain modeling and the Fourier transform (Nihei and Li, 2007).

### Heterogeneous model

CARP-CG is applied to perform the 3D frequency-domain elastic wave modelings at different frequencies using a 3D heterogeneous elastic model (courtesy of Shell). The size of the model is

$48 \times 20 \times 4$  km. The P- and S-wave velocities and the density models are presented in Figure 18. The minimum S-wave velocity is approximately 837 m/s. The 3D optimized staggered-grid FD scheme is applied to discretize the elastic system 13. The PMLs are applied at the boundaries. An explosive Dirac source located at (20, 10, 0.55 km) is used. To ensure the accuracy of the FD modeling, the grid size is adapted to keep 4.2 grid points per minimum S-wavelength. The geometry parameters, total number of unknowns  $N_u$  (with  $N_{\text{PML}} = 10$ ), computing time, number of cores  $N_{\text{proc}}$ , estimated memory used per core, and number of iterations  $N_{\text{iter}}$  at different frequencies are summarized in Table 10. The convergence curves are presented in Figure 19a. Representative wavefields of the horizontal particle velocity  $v_x$  at different frequencies are given in Figure 20. The shallow part generates a larger number of waves corresponding to the smaller wave speed.

Note that the source is located in a layer with a slow wave speed surrounded by two layers with a higher wave speed, similarly to a waveguide geometry. The propagation of the guided waves leads to a particularly challenging problem for the iterative solvers. Such

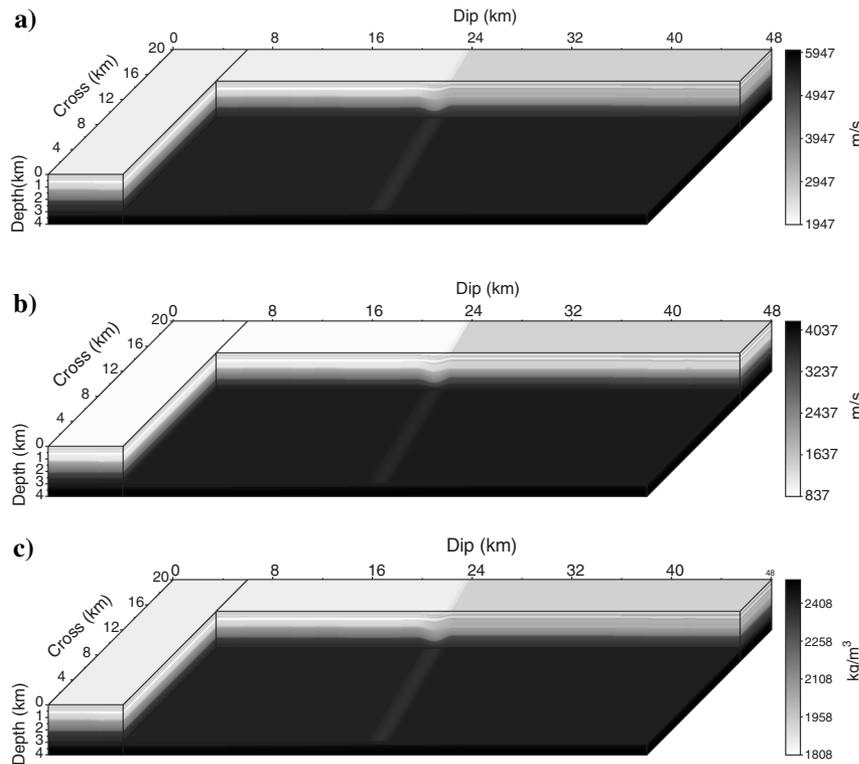


Figure 18. 3D elastic model: (a) P-wave velocity, (b) S-wave velocity, and (c) density.

Table 10. Geometry parameters, total number of unknowns  $N_u$ , computing time, number of cores  $N_{\text{proc}}$ , estimated memory used per core, and number of iterations  $N_{\text{iter}}$  for the 3D elastic wave modeling in a heterogeneous model at different frequencies.

$f$ (Hz)	$h$ (m)	$n_x \times n_y \times n_z$	$N_u$	Time (s)	$N_{\text{proc}}$	Mem/core (MB)	$N_{\text{iter}}$
0.625	320	$151 \times 63 \times 13$	1,405,107	182.3	16	54.9	1658
1.25	160	$301 \times 126 \times 26$	6,467,508	333.9	48	84.2	1550
2.5	80	$601 \times 251 \times 51$	35,845,983	4157.6	96	233.6	5055
5	40	$1201 \times 501 \times 101$	230,919,183	33,218.9	192	752.4	8654

guided waves can be observed in the shallow part of the wavefields given in Figure 20. Despite these difficulties, we show that the CARP-CG method is able to converge, which emphasizes its robustness.

The complexity analysis as in the last section is presented in Figure 19b. We obtain a slope approximately equal to 0.65, which means  $N_{\text{iter}} = \mathcal{O}(N^{0.65})$ . The total computational complexity for one source using CARP-CG is

$$\mathcal{O}(N_{\text{iter}} \times N^3) = \mathcal{O}(N^{3.65}), \quad (25)$$

which is close to the 3D experiment in a homogeneous model.

### CONCLUSION

In this study, we investigate the performance of the CARP-CG method for 2D and 3D frequency-domain elastic wave modeling.

Figure 19. (a) Convergence curves of the wave modeling in a 3D heterogeneous elastic model using CARP-CG. Complexity analysis of the CARP-CG method for the 3D experiments is given in panel (b), in which the number of iterations  $N_{\text{iter}}$  is plotted as a function of the average model dimension  $N$  on a log-log scale. The straight dotted-dashed line in panel (b) indicates the results obtained from line regression.

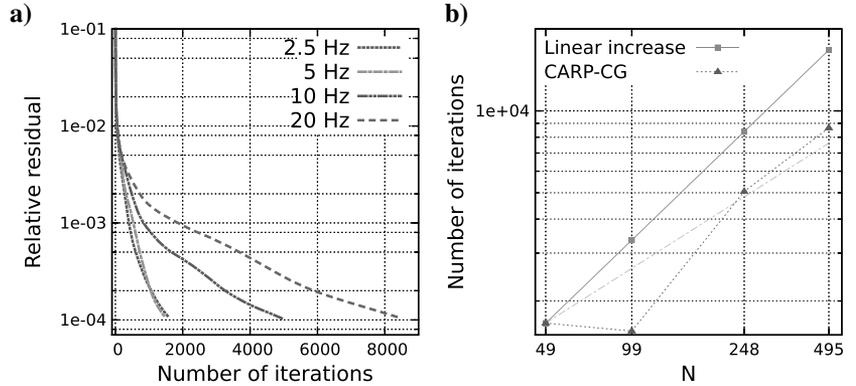
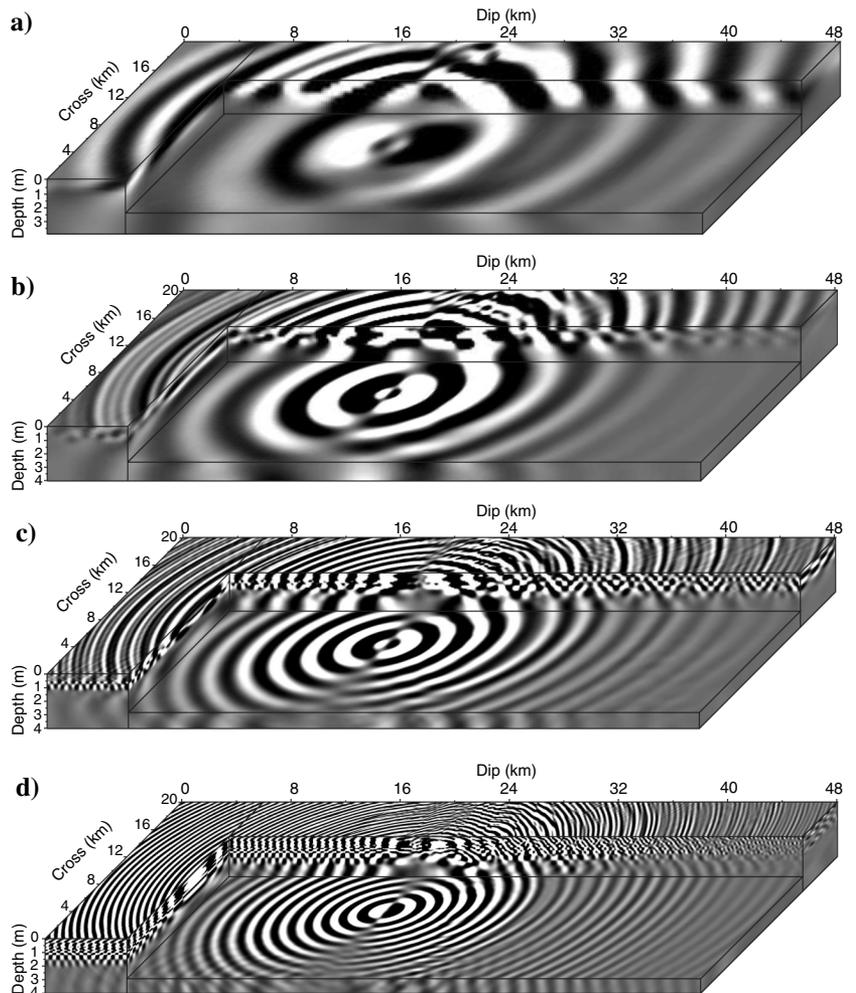


Figure 20. 3D frequency-domain wavefields of the horizontal particle velocity  $v_x$  at (a) 0.625, (b) 1.25, (c) 2.5, and (d) 5 Hz.



The CARP-CG method is based on the double-sweep Kaczmarz method, which is equivalent to the SSOR method applied to the normal equations. Solving the corresponding linear system with the CG method leads to the CGMN method. The CARP-CG method is a parallelization of the CGMN method through the row-block division and the component-averaging operations.

Several experiments are performed to investigate the influence of different physical factors on the convergence of the CARP-CG method. We compute the condition numbers corresponding to each case to illustrate how the physical factors affect the conditioning of the impedance matrix. The results indicate that a large and heterogeneous Poisson's ratio leads to a more ill-conditioned impedance matrix. More iterations are thus required to achieve the convergence. The FSBC has a similar influence on the condition number of the matrix due to the appearance of the surface waves. In both cases, CARP-CG exhibits its robustness because it presents very regular convergence. The incorporation of attenuation has a beneficial effect on the condition number. We show that the convergence rate of CARP-CG increases as more seismic attenuation is incorporated in the modeling.

The scalability of the CARP-CG method is investigated through a set of experiments performed in the original Marmousi2 model with different numbers of cores. The row-block division and the component-averaging operations lead to a satisfactory scalability for the CARP-CG method. The comparison with standard Krylov iterative solvers (GMRES, CGNR, and BiCGStab) emphasizes the good performance and robustness of CARP-CG. The CGNR method shares the same robustness of CARP-CG, but its convergence is very slow. The results show that the convergence of the GMRES cannot be guaranteed and that the BiCGStab methods fails to converge in all cases. Such nonconvergence seems to arise from the complexity of the elastic model. The CARP-CG method, however, converges in all cases.

3D elastic wave modeling is also investigated. We test the CARP-CG method in homogeneous and heterogeneous models. The CARP-CG method exhibits its capability of converging for 3D large-scale problems, involving up to hundreds of millions of unknowns. The 3D heterogeneous modeling is performed in a configuration resembling a waveguide. The source is located in a thin layer with a slow wave speed surrounded by two layers with a fast wave speed. This configuration is known to be particularly challenging for iterative solvers because the standing modes are difficult to obtain due to the prolonged propagation in the waveguide. Despite this difficult configuration, the CARP-CG method still converges.

This study suggests that CARP-CG could be a potential solver for large-scale 3D frequency-domain elastic modeling, especially for its ability to converge even in complex models. Its scalability is also encouraging, as well as the computational complexity, which is shown to be an improvement compared to the other techniques such as the direct solver, or the combination of the time-domain modeling and the DFT. However, the basic number of iterations required by CARP-CG is still large. Let us remember that because  $N_{\text{iter}} = \mathcal{O}(N^\alpha) = CN^\alpha$ , where  $C$  is a constant, a large  $C$  may generate an  $N_{\text{iter}} > N$ , leading to a larger computational cost compared to direct solvers. Thus, applying the preconditioning techniques to decrease  $C$  becomes necessary. The complexity parameter  $\alpha$  may also change with proper preconditioners. Moreover, MRHS techniques have to be investigated in the CARP-CG framework for seismic imaging problems. This will be the object of our future research. Depending on these possible accelerations, the CARP-CG method could be an efficient modeling engine for FWI applications.

## ACKNOWLEDGMENTS

This study was partially funded by the SEISCOPE Consortium (<http://seiscope2.osug.fr>), sponsored by BP, CGG, Chevron, Exxon-Mobil, JGI, Petrobras, Saudi Aramco, Schlumberger, Shell, Sinopec, Statoil, Total, and Woodside. This study was also funded by the National Natural Science Foundation of China under grant nos. 91230119 and 41474102, and the Chinese Scholarship Council. This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), which is supported by the Rhône-Alpes region (GRANT CPER07\_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56), and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme Investissements d'Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS under allocation 046091 made by GENCI. The authors would like to acknowledge the associate editor M. Sacchi and three anonymous reviewers for their instructive and detailed comments. The authors appreciate the fruitful discussions with W. Zhou (ISterre, Université Grenoble Alpes). Thanks are also due to D. Gordon (University of Haifa) for his helpful suggestions.

## REFERENCES

- Bérenger, J.-P., 1994, A perfectly matched layer for absorption of electromagnetic waves: *Journal of Computational Physics*, **114**, 185–200, doi: [10.1006/jcph.1994.1159](https://doi.org/10.1006/jcph.1994.1159).
- Björck, Å., and T. Elfving, 1979, Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations: *BIT Numerical Mathematics*, **19**, 145–163, doi: [10.1007/BF01930845](https://doi.org/10.1007/BF01930845).
- Britt, S., S. Tsynkov, and E. Turkel, 2011, Numerical simulation of time-harmonic waves in inhomogeneous media using compact high order schemes: *Communications in Computational Physics*, **9**, 520–541, doi: [10.4208/cicp.091209.080410s](https://doi.org/10.4208/cicp.091209.080410s).
- Brossier, R., B. Pajot, L. Combe, S. Operto, L. Métivier, and J. Virieux, 2014, Time and frequency-domain FWI implementations based on time solver: Analysis of computational complexities: Presented at 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, We G102 05.
- Chan, T. F., and M. K. Ng, 1999, Galerkin projection methods for solving multiple linear systems: *SIAM Journal on Scientific Computing*, **21**, 836–850, doi: [10.1137/S1064827598310227](https://doi.org/10.1137/S1064827598310227).
- 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.
- Ernst, O. G., and M. J. Gander, 2012, Why it is difficult to solve Helmholtz problems with classical iterative methods, in I. G. Graham, T. Y. Hou, O. Lakkis, and R. Scheichl, eds., *Numerical analysis of multiscale problems*: Springer, Lecture Notes in Computational Science and Engineering 83, 325–363.
- 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, Technical report TR/PA/03/3.
- Futterman, W., 1962, Dispersive body waves: *Journal of Geophysical Research*, **67**, 5279–5291, doi: [10.1029/JZ067i013p05279](https://doi.org/10.1029/JZ067i013p05279).
- Gordon, D., and R. Gordon, 2005, Component-averaged row projections: A robust block-parallel scheme for sparse linear systems: *SIAM Journal on Scientific Computing*, **27**, 1092–1117, doi: [10.1137/040609458](https://doi.org/10.1137/040609458).
- 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, doi: [10.1145/1391989.1391991](https://doi.org/10.1145/1391989.1391991).
- Gordon, D., and R. Gordon, 2010b, Row scaling as a preconditioner for some nonsymmetric linear systems with discontinuous coefficients: *Journal of Computational and Applied Mathematics*, **234**, 3480–3495, doi: [10.1016/j.cam.2010.05.021](https://doi.org/10.1016/j.cam.2010.05.021).
- Gordon, D., and R. Gordon, 2010a, CARP-CG: A robust and efficient parallel solver for linear systems, applied to strongly convection dominated PDEs: *Parallel Computing*, **36**, 495–515, doi: [10.1016/j.parco.2010.05.004](https://doi.org/10.1016/j.parco.2010.05.004).
- Gordon, D., and R. Gordon, 2011, An overview of the CARP-CG algorithm: *Proceedings of the International Workshop on Parallel Numerics (Par-num11)*, 38–47.

- 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**, 182–196, doi: [10.1016/j.cam.2012.07.024](https://doi.org/10.1016/j.cam.2012.07.024).
- Gosselin-Cliche, B., and B. Giroux, 2014, 3D frequency-domain finite-difference viscoelastic-wave modeling using weighted average 27-point operators with optimal coefficients: *Geophysics*, **79**, no. 3, T169–T188, doi: [10.1190/geo2013-0368.1](https://doi.org/10.1190/geo2013-0368.1).
- Graves, R., 1996, Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences: *Bulletin of the Seismological Society of America*, **86**, 1091–1106.
- Jo, C. H., C. Shin, and J. H. Suh, 1996, An optimal 9-point, finite-difference, frequency-space 2D scalar extrapolator: *Geophysics*, **61**, 529–537, doi: [10.1190/1.1443979](https://doi.org/10.1190/1.1443979).
- 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.
- Kolsky, H., 1956, The propagation of stress pulses in viscoelastic solids: *Philosophical Magazine*, **1**, 693–710, doi: [10.1080/14786435608238144](https://doi.org/10.1080/14786435608238144).
- Levander, A. R., 1988, Fourth-order finite-difference P-SV seismograms: *Geophysics*, **53**, 1425–1436, doi: [10.1190/1.1442422](https://doi.org/10.1190/1.1442422).
- Li, Y., L. Métivier, R. Brossier, B. Han, and J. Virieux, 2014a, CARP-CG: A robust parallel iterative solver for frequency-domain elastic wave modeling, application to the Marmousi2 model: 84th Annual International Meeting, SEG, 3487–3492.
- Li, Y., L. Métivier, R. Brossier, B. Han, and J. Virieux, 2014b, A robust parallel iterative solver for frequency-domain elastic wave modeling: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, Th P10 01.
- Luo, Y., and G. T. Schuster, 1990, Parsimonious staggered grid finite-differencing of the wave equation: *Geophysical Research Letters*, **17**, 155–158, doi: [10.1029/GL017i002p00155](https://doi.org/10.1029/GL017i002p00155).
- Marfurt, K., 1984, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations: *Geophysics*, **49**, 533–549, doi: [10.1190/1.1441689](https://doi.org/10.1190/1.1441689).
- Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An elastic upgrade for Marmousi: *The Leading Edge*, **25**, 156–166, doi: [10.1190/1.2172306](https://doi.org/10.1190/1.2172306).
- 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, doi: [10.1190/1.1444785](https://doi.org/10.1190/1.1444785).
- Mittet, R., 2002, Free-surface boundary conditions for elastic staggered-grid modeling schemes: *Geophysics*, **67**, 1616–1623, doi: [10.1190/1.1512752](https://doi.org/10.1190/1.1512752).
- Nihei, K. T., and X. Li, 2007, Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD): *Geophysical Journal International*, **169**, 1069–1078, doi: [10.1111/j.1365-246X.2006.03262.x](https://doi.org/10.1111/j.1365-246X.2006.03262.x).
- O'Leary, D. P., 1980, The block conjugate gradient algorithm and related methods: *Linear Algebra and Its Applications*, **29**, 293–322, doi: [10.1016/0024-3795\(80\)90247-5](https://doi.org/10.1016/0024-3795(80)90247-5).
- Operto, S., J. Virieux, P. Amestoy, J.-Y. L'Excellent, L. Giraud, and H. Ben Hadj Ali, 2007, 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study: *Geophysics*, **72**, no. 5, SM195–SM211, doi: [10.1190/1.2759835](https://doi.org/10.1190/1.2759835).
- Plessix, R. E., 2007, A Helmholtz iterative solver for 3D seismic-imaging problems: *Geophysics*, **72**, no. 5, E185–E191, doi: [10.1190/1.2738849](https://doi.org/10.1190/1.2738849).
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, part I: Theory and verification in a physics scale model: *Geophysics*, **64**, 888–901, doi: [10.1190/1.1444597](https://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**, 431–448, doi: [10.1016/j.jcp.2007.03.033](https://doi.org/10.1016/j.jcp.2007.03.033).
- Saad, Y., 1986, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems: *SIAM Journal on Scientific and Statistical Computing*, **7**, 856–869, doi: [10.1137/0907058](https://doi.org/10.1137/0907058).
- 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**, 457–469, doi: [10.1016/0377-0427\(95\)00198-0](https://doi.org/10.1016/0377-0427(95)00198-0).
- Sirgue, L., J. T. Etgen, and U. Albertin, 2008, 3D frequency domain waveform inversion using time domain finite difference methods: 70th Annual International Conference and Exhibition, EAGE, Extended Abstracts, F022.
- Sirgue, L., T. J. Etgen, U. Albertin, and S. Brandsberg-Dahl, 2007, System and method for 3D frequency-domain waveform inversion based on 3D time-domain forward modeling: U.S. Patent Application, US2007/0282535 A1.
- Sleijpen, G. L. G., and D. R. Fokkema, 1993, BiCGSTAB(L) for linear equations involving unsymmetric matrices with complex spectrum: *Electronic Transactions on Numerical Analysis*, **1**, 11–32.
- Turkel, E., D. Gordon, R. Gordon, and S. Tsynkov, 2013, Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number: *Journal of Computational Physics*, **232**, 272–287, doi: [10.1016/j.jcp.2012.08.016](https://doi.org/10.1016/j.jcp.2012.08.016).
- Van der Vorst, H. A., 1992, BI-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems: *SIAM Journal on Scientific and Statistical Computing*, **13**, 631–644, doi: [10.1137/0913035](https://doi.org/10.1137/0913035).
- 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.
- van Leeuwen, T., and F. J. Herrmann, 2014, 3D frequency-domain seismic inversion with controlled sloppiness: *SIAM Journal on Scientific Computing*, **36**, S192–S217, doi: [10.1137/130918629](https://doi.org/10.1137/130918629).
- Virieux, J., 1986, P-SV wave propagation in heterogeneous media, velocity-stress finite difference method: *Geophysics*, **51**, 889–901, doi: [10.1190/1.1442147](https://doi.org/10.1190/1.1442147).
- Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, doi: [10.1190/1.3238367](https://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, doi: [10.1190/geo2011-0163.1](https://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: 83rd Annual International Meeting, SEG, Expanded Abstracts, 3411–3416.