GEOPHYSICS, VOL. 72, NO. 5 (SEPTEMBER-OCTOBER 2007); P. SM195–SM211, 14 FIGS., 3 TABLES.
10.1190/1.2759835

3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study

Stéphane Operto[1], Jean Virieux[2], Patrick Amestoy[3], Jean-Yves L'Excellent[4], Luc Giraud[3], and Hafedh Ben Hadj Ali[2]

ABSTRACT

We present a finite-difference frequency-domain method for 3D visco-acoustic wave propagation modeling. In the frequency domain, the underlying numerical problem is the resolution of a large sparse system of linear equations whose right-hand side term is the source. This system is solved with a massively parallel direct solver. We first present an optimal 3D finite-difference stencil for frequency-domain modeling. The method is based on a parsimonious staggered-grid method. Differential operators are discretized with second-order accurate staggered-grid stencils on different rotated coordinate systems to mitigate numerical anisotropy. An antilumped mass strategy is implemented to minimize numerical dispersion. The stencil incorporates 27 grid points and spans two grid intervals. Dispersion analysis shows that four grid points per wavelength provide accurate simulations in the 3D domain. To assess the feasibility of the method for frequency-domain full-waveform inversion, we computed simulations in the 3D SEG/EAGE overthrust model for frequencies 5, 7, and 10 Hz. Results confirm the huge memory requirement of the factorization (several hundred Figabytes) but also the CPU efficiency of the resolution phase (few seconds per shot). Heuristic scalability analysis suggests that the memory complexity of the factorization is $\mathcal{O}(35N^4)$ for a N^3 grid. Our method may provide a suitable tool to perform frequency-domain full-waveform inversion using a large distributed-memory platform. Further investigation is still necessary to assess more quantitatively the respective merits and drawbacks of time- and frequency-domain modeling of wave propagation to perform 3D full-waveform inversion.

INTRODUCTION

Quantitative seismic imaging of 3D crustal structures is one of the main challenges of geophysical exploration at different scales for subsurface, oil exploration, crustal, and lithospheric investigations. Frequency-domain full-waveform inversion has recently aroused increasing interest following the pioneering work of R.G. Pratt and collaborators (Pratt, 2004). A few applications of frequency-domain full-waveform inversion applied to 2D onshore and offshore wide-aperture (global-offset) seismic data have recently been presented to image complex structures such as a thrust belt or a subduction zone (Operto et al., 2004, 2007; Ravaut et al., 2004). The potential interest of such approaches is to exploit the broad aperture range spanned by global-offset geometries to image a broad and continuous range of wavelengths in the velocity model. The frequency-domain approach of full-waveform inversion has been shown to be efficient for several reasons (e.g., Pratt et al., 1996, 1998; Pratt, 1999; Brenders and Pratt, 2006). First, only a few discrete frequencies are necessary to develop a reliable image of the medium, and second, proceeding sequentially from low to high frequencies defines a multiresolution imaging strategy that helps to mitigate the nonlinearity of the inverse problem.

In 2D, the few frequency components required to solve the inverse problem can be efficiently modeled in the frequency domain using a finite-difference (FDFD) method (Marfurt, 1984; Jo et al., 1996; Štekl and Pratt, 1998; Hustedt et al., 2004). Modeling of one frequency with a finite-difference method requires solving a large, sparse system of linear equations. If one can use a direct solver to solve the system, the solution for multiple right-hand side terms (i.e.,

Manuscript received by the Editor November 29, 2006; revised manuscript received April 16, 2007; published online August 23, 2007.
[1]Géosciences Azur, CNRS, IRD, UNSA, UPMC, Villefranche-sur-mer, France. E-mail: operto@geoazur.obs-vlfr.fr.
[2]Géosciences Azur, CNRS, IRD, UNSA, UPMC, Valbonne, France. E-mail: viri@geoazur.unice.fr.
[3]ENSEEIHT-IRIT, Toulouse, France. E-mail: amestoy@enseeiht.fr; giraud@n7.fr.
[4]Laboratoire de l'informatique du parallélisme, ENS Lyon, Lyon, France. E-mail: jean-yves.l.excellent@ens-lyon.fr.
© 2007 Society of Exploration Geophysicists. All rights reserved.

multiple sources) can be obtained very efficiently and this iscritical for seismic imaging. Indeed, the matrix factorization is done once, and then multiple solutions can be rapidly obtained by forward and backward substitutions. Moreover, attenuation can be easily implemented in the frequency domain using complex velocities (Toksöz and Johnston, 1981). Another advantage of direct solvers compared to the iterative alternative is robustness in the sense that they will give highly accurate answers to very general problems after a finite number of steps (see Demmel, 1997, for a discussion on the respective merits of direct and iterative solvers).

The drawback of the direct approach with respect to the iterative counterpart or the time-domain formulation is that the LU factorization of the matrix leads to fill-in and hence requires a huge amount of RAM memory or disk space to store the LU factors. Today, modern computers with shared or distributed memory allow us to tackle 2D frequency-domain full-waveform modeling and inversion problems for representative crustal-scale problems (i.e., typically, 2D velocity models of dimension 100×25 km and frequency up to 15 Hz, [Operto et al., 2007]). In 3D, the storage requirement and the complexity of the FDFD problem may appear rather discouraging. Therefore, most of the works on resolution of the 3D Helmholtz equation focuses on iterative solvers (Riyanti et al., 2006; Plessix, 2006). However, much recent effort has been dedicated to developing massively parallel direct solvers that allow solution of problems involving several million unknowns (Amestoy et al., 2006). Therefore, we believe that it is worth investigating more quantitatively the categories of seismic imaging problems, which can be addressed with FDFD wave propagation methods based on massively parallel direct solvers. Frequency-domain full-waveform inversion of the low-frequency part of the source bandwidth may be one of these imaging problems. Direct solvers can also be implemented within hybrid direct-iterative solvers based on Schwarz-based domain decomposition method for which factorizations are performed in subdomains of limited dimensions in conjunction with a Krylov subspace method (Štekl and Pain, 2002).

In this paper, we present an optimal stencil for 3D FDFD wave propagation modeling. Our method is the 3D extension of the 2D parsimonious mixed-grid FDFD method for acoustic wave propagation developed by Hustedt et al. (2004). In the first part of the paper, we review the first-order hyperbolic velocity stress and the second-order Helmholtz-type formulations of the 3D acoustic wave equation, and we develop the discretization strategies applied to these equations. This will lead to a 27-point stencil spanning two grid intervals. In the second part, we perform a dispersion analysis to assess the accuracy of the stencil in infinite homogeneous media. This dispersion analysis suggests that four grid points per wavelength will provide accurate simulations. In the third part, we briefly review the main functionalities of the massively parallel solver that we used for the numerical resolution (Amestoy et al., 2006; MUMPS Team, 2007). In the fourth part, we illustrate the accuracy of the method with several numerical examples that confirm the conclusion of the dispersion analysis. Finally, we present a heuristic complexity analysis, which suggests that memory and time complexity of the matrix factorization are $\mathcal{O}(n^4)$ and $\mathcal{O}(n^6)$, respectively, when considering a 3D n^3 computational domain. These estimations are consis-tent with former theoretical analysis. We conclude with some comments on future developments.

THE PARSIMONIOUS MIXED-GRID FD METHOD

Principle of the method

The aim of this section is to design an accurate and spatially compact FD stencil for frequency-domain modeling based on a direct solver. We propose an extension to the 3D case of the 2D parsimonious mixed-grid method presented by Hustedt et al. (2004), which is itself an extension of the mixed-grid method developed by Jo et al. (1996) for homogeneous acoustic media and extended by Štekl and Pratt (1998) for viscoelastic heterogeneous media. Use of a spatially compact stencil is critical if a direct method (LU factorization) is used to solve the system resulting from the discretization of the Helmholtz equation. Indeed, spatially compact stencils allow limitation of the numerical bandwidth of the matrix and hence its filling during LU factorization. We implemented spatially compact stencils with second-order accurate differencing operators. Accuracy in terms of both numerical anisotropy and dispersion is achieved using the following: First, the differential operators are discretized on different rotated coordinate systems and combined linearly following the so-called mixed-grid strategy (Jo et al., 1996, Štekl and Pratt, 1998; Hustedt et al., 2004). Second, the mass term at the collocation point is replaced by its weighted average over the grid points involved in the stencil (Marfurt, 1984; Takeuchi et al., 2000).

Concerning discretization of differential operators, Hustedt et al. (2004) clarify the relationship between the original mixed-grid approach of Jo et al. (1996) and the staggered-grid methods applied to the first-order hyperbolic velocity-stress formulation of the wave equation (Virieux, 1984, 1986; Saenger et al., 2000) through a parsimonious strategy, originally developed for the time-domain wave equation (Luo and Schuster, 1990). The parsimonious strategy provides a systematic recipe for discretizing the second-order wave equation from its first-order representation. In the parsimonious approach of Hustedt et al. (2004), the wave equation is first written as a first-order hyperbolic system in pressure-particle velocity and discretized using staggered-grid stencils of second-order accuracy on different rotated coordinate systems. After discretization, particle velocity wavefields are eliminated from the system, leading to a parsimonious staggered-grid wave equation on each rotated coordinate system. Elimination of the three particle-velocity wavefields allows decreasing the order of the matrix by a factor of four for the 3D acoustic wave equation. Once the discretization and the elimination have been applied on each coordinate system, the resulting differencing operators are linearly combined into a single discrete wave equation.

In the following section, we detail the successive steps involved in FD discretization of the 3D frequency-domain wave equation introduced above.

The 3D frequency-domain visco-acoustic wave equation

We begin with a review of the 3D frequency-domain acoustic wave equation. This equation is written as a first-order hyperbolic system (Virieux, 1984) using physical quantities as the pressure, and the particle velocity is given by

$$\begin{aligned}
\frac{-i\omega}{\kappa(x,y,z)}p(x,y,z,\omega) &= \frac{1.}{\xi_x(x)}\frac{\partial v_x(x,y,z,\omega)}{\partial x} \\
&+ \frac{1.}{\xi_y(y)}\frac{\partial v_y(x,y,z,\omega)}{\partial y} \\
&+ \frac{1.}{\xi_z(z)}\frac{\partial v_z(x,y,z,\omega)}{\partial z}, \\
v_x(x,y,z,\omega) &= -\frac{b(x,y,z)}{i\omega\xi_x(x)}\frac{\partial p(x,y,z,\omega)}{\partial x} \\
&+ f_x(x,y,z,\omega), \\
v_y(x,y,z,\omega) &= -\frac{b(x,y,z)}{i\omega\xi_y(y)}\frac{\partial p(x,y,z,\omega)}{\partial y} \\
&+ f_y(x,y,z,\omega), \\
v_z(x,y,z,\omega) &= -\frac{b(x,y,z)}{i\omega\xi_z(z)}\frac{\partial p(x,y,z,\omega)}{\partial z} \\
&+ f_z(x,y,z,\omega). \qquad (1)
\end{aligned}$$

This first-order system of equations will be discretized with staggered-grid stencils in the following section. In equation 1, wavefields $v_x(x,y,z,\omega)$, $v_y(x,y,z,\omega)$, and $v_z(x,y,z,\omega)$ are components of the particle velocity vector, $p(x,y,z,\omega)$ is the pressure, ω is the angular frequency. The bulk modulus is $\kappa(x,y,z)$ and $b(x,y,z)$ is the buoyancy, the inverse of density. External forces acting as the source term are represented by f_x, f_y, and f_z. The 1D functions ξ_x, ξ_y, and ξ_z define damping functions implemented in the absorbing perfectly matched layers (PML) surrounding the model on all four sides of the computational domain (Berenger, 1994). System 1 based on unsplit PML conditions is derived from the time-domain wave equation with split PML conditions in Appendix A.

We inject the expression of the particle velocities (last three equations of equation system 1) in the first equation of system 1, which leads to the second-order elliptic wave equation in pressure,

$$\begin{aligned}
&\left[\frac{\omega^2}{\kappa(x,y,z)} + \frac{1.}{\xi_x(x)}\frac{\partial}{\partial x}\frac{b(x,y,z)}{\xi_x(x)}\frac{\partial}{\partial x} + \frac{1.}{\xi_y(y)}\frac{\partial}{\partial y}\frac{b(x,y,z)}{\xi_y(y)}\frac{\partial}{\partial y}\right. \\
&\left. + \frac{1.}{\xi_z(z)}\frac{\partial}{\partial z}\frac{b(x,y,z)}{\xi_z(z)}\frac{\partial}{\partial z}\right]p(x,y,z,\omega) = -s(x,y,z,\omega), \qquad (2)
\end{aligned}$$

where $s = \nabla\cdot\mathbf{f}$ is the pressure source. This elimination procedure will be applied after discretization of system 1 in the following section.

If f_x, f_y, and f_z are unidirectional impulses applied at the same spatial position, the pressure source corresponds to an explosion.

In matrix form, we have

$$[\mathbf{M} + \mathbf{S}]\mathbf{p} = -\mathbf{s},$$

where $\mathbf{M}$ and $\mathbf{S}$ are the mass and stiffness matrices, respectively. We implemented an explosive source in vector $\mathbf{s}$ by setting one nonzero complex coefficient at the position of the explosion. This coefficient is the Fourier transform of the source wavelet at the current frequency, normalized by the volume of the cubic cell.

Differencing operators: The 3D 27-point stencil

In this section, we discretize in a finite-difference sense the Helmholtz equation, equation 2, using the parsimonious mixed-grid strategy. The successive steps of the discretization are

Step 1: 3D coordinate systems are defined such that their axes span as many directions as possible in a cubic cell (Figure 1). These coordinate systems must be consistent with 3D second-order staggered geometry (e.g., see Virieux, 1984 and Hustedt et al., 2004 for the 2D case).
Step 2: The first equation of system 1 is discretized on each of the coordinate systems using second-order centered staggered-grid stencils. The discrete equation will involve particle velocities on staggered grids.
Step 3: The particle velocities at the grid points involved in the first equation of system 1 are inferred from the last three equations of system 1 using the same staggered-grid stencils as for the first equation.
Step 4: The expressions of the particle velocities are reinjected in the first equation of system 1, leading to a second-order parsimonious staggered-grid wave equation in pressure.
Step 5: Once steps 2–4 have been performed for each coordinate systems, all the discrete wave equations are combined linearly. A necessary condition for applying this combination is that the pressure wavefield kept after elimination be discretized on the same grid as the coordinate system that was selected during step 1.

Here we identify eight coordinate systems that cover all possible directions in a cubic cell and that are consistent with the staggered-grid geometry. Spatial partial derivatives in the wave equation, equation 1, are discretized along these coordinate systems.

These eight coordinate systems are

1) the classic Cartesian coordinate system (x,y,z). The associated basis will be denoted by B_c in the following and the resulting stencil will be referred to as stencil 1 (Figure 1a).
2) three coordinate systems obtained by rotating the Cartesian system around each Cartesian axis x, y, and z. The associated basis obtained by rotation around x, y, and z will be noted by B_x, B_y, and B_z, respectively. The coordinates on each basis will be noted (x,y_x,z_x), (x_y,y,z_y), and (x_z,y_z,z), respectively (Figure 1b). Each stencil associated with the basis B_x, B_y, and B_z incorporates 11 coefficients. The stencil resulting from the averaging of the stencils developed on each basis B_x, B_y, and B_z will be referred to as stencil 2 and incorporates 19 coefficients.
3) four coordinate systems defined by the four big diagonals of a cube (Figure 1c). If we denote by $\hat{\mathbf{d}}_1$, $\hat{\mathbf{d}}_2$, $\hat{\mathbf{d}}_3$, and $\hat{\mathbf{d}}_4$ unit vectors along each of these diagonals, four bases can be formed $B_1 = (\hat{\mathbf{d}}_1,\hat{\mathbf{d}}_2,\hat{\mathbf{d}}_3)$, $B_2 = (\hat{\mathbf{d}}_1,\hat{\mathbf{d}}_2,\hat{\mathbf{d}}_4)$, $B_3 = (\hat{\mathbf{d}}_1,\hat{\mathbf{d}}_3,\hat{\mathbf{d}}_4)$, and $B_4 = (\hat{\mathbf{d}}_2,\hat{\mathbf{d}}_3,\hat{\mathbf{d}}_4)$. The associated coordinates will be denoted as d_1, d_2, d_3, and d_4 in the following. These four coordinate systems are similar to those developed by Saenger et al. (2000) for 3D time-domain elastic staggered-grid finite-difference methods. The stencil resulting from the averaging of the four stencils developed on the base B_1, B_2, B_3, and B_4 will be referred to as stencil 3. This stencil has 27 coefficients.

These coordinate systems differ from those introduced by Štekl et al. (2002) and Štekl and Pain (2002) who propose using, in addition to the Cartesian system and the three systems B_x, B_y, and B_z, six addi-

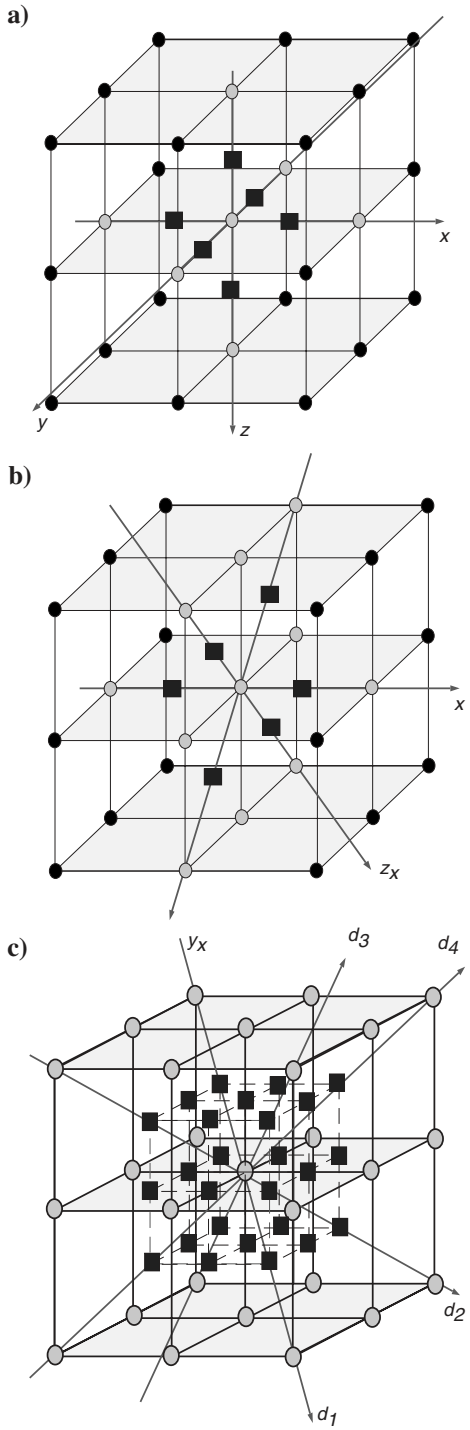


Figure 1. FD stencils. Circles are pressure grid points. Squares are positions where buoyancy needs to be interpolated because of the staggered-grid geometry. Gray circles are pressure grid points involved in the stencil. (a) Stencil on the classic Cartesian coordinate system. This stencil incorporates seven coefficients. (b) Stencil on the rotated Cartesian coordinate system. Rotation is applied around x on the figure. This stencil incorporates 11 coefficients. The same strategy can be applied by rotation around y and z. Averaging of the three resultant stencils defines a 19-coefficient stencil. (c) Stencil obtained from four coordinate systems, each associated with three big diagonals of a cubic cell. This stencil incorporates 27 coefficients.

tional coordinate systems obtained by rotation around two of the Cartesian axes. These coordinate systems are not consistent with our staggered-grid method in the sense that it would require defining more than one pressure grid.

Finally, the stiffness matrices associated with each coordinate system are combined linearly:

$$\mathbf{Sp} \Rightarrow [w_1\mathbf{S}_{B_c} + w_2/3(\mathbf{S}_{B_x} + \mathbf{S}_{B_y} + \mathbf{S}_{B_z}) + w_3/4(\mathbf{S}_{B_1} + \mathbf{S}_{B_2} + \mathbf{S}_{B_3} + \mathbf{S}_{B_4})]\mathbf{p},$$

where we introduced the weighting coefficients w_1, $w_2/3$, and $w_3/4$ associated with stencils 1, 2, and 3, respectively. The coefficients verify

$$w_1 + w_2 + w_3 = 1. \tag{3}$$

The factors 1/3 and 1/4 applied to coefficients w_2 and w_3 account for the fact that stencils 2 and 3 are the average of three elementary stencils and four elementary stencils, respectively.

Expression of the partial derivatives with respect to x, y, and z as a function of the spatial derivatives with respect to each of the above mentioned coordinates are given in Appendix B. The second-order centered staggered-grid stencils for each partial derivative of a wavefield with respect to each coordinate are given in Appendix C. These discrete expressions are used to discretize the equations in system 1 before elimination of the discrete particle velocity fields. The final expression of the eight parsimonious staggered-grid wave equations are given in Appendix D. In the appendices and below, we used the following notations for compactness: We consider a given cubic cell of the finite-difference mesh. The pressure at the nodes of the cubic cell are denoted by p_{lmn} where $l,m,n \in \{-1,0,1\}$, and p_{000} denotes the central grid point.

Indices 1/2 indicate buoyancy grid points located at intermediate positions with respect to the reference pressure grid following the staggered-grid strategy.

Mass-term averaging

The accuracy of the stencil can be greatly improved by a redistribution of the mass term over the different grid points surrounding the collocation point involved in the finite-difference stencils following an antilumped mass approach. Following standard procedure of finite element methods (Marfurt, 1984), the diagonal mass term is distributed through weighted values such that

$$\frac{\omega^2}{\kappa_{000}}p_{000} \Rightarrow \omega^2\left(w_{m1}\left[\frac{p}{\kappa}\right]_0 + w_{m2}\left[\frac{p}{\kappa}\right]_1 + w_{m3}\left[\frac{p}{\kappa}\right]_2 + w_{m4}\left[\frac{p}{\kappa}\right]_3\right), \tag{4}$$

where

$$w_{m1} + \frac{w_{m2}}{6} + \frac{w_{m3}}{12} + \frac{w_{m4}}{8} = 1. \tag{5}$$

In equation 4, we used the notations

$$\left[\frac{p}{\kappa}\right]_0 = \frac{p_{000}}{\kappa_{000}},$$

$$\left[\frac{p}{\kappa}\right]_1 = \frac{p_{100}}{\kappa_{100}} + \frac{p_{010}}{\kappa_{010}} + \frac{p_{001}}{\kappa_{001}} + \frac{p_{-100}}{\kappa_{-100}} + \frac{p_{0-10}}{\kappa_{0-10}} + \frac{p_{00-1}}{\kappa_{00-1}},$$

$$\left[\frac{p}{\kappa}\right]_2 = \frac{p_{110}}{\kappa_{110}} + \frac{p_{011}}{\kappa_{011}} + \frac{p_{101}}{\kappa_{101}} + \frac{p_{-110}}{\kappa_{-110}} + \frac{p_{0-11}}{\kappa_{0-11}} + \frac{p_{-101}}{\kappa_{-101}} + \frac{p_{1-10}}{\kappa_{1-10}} + \frac{p_{01-1}}{\kappa_{01-1}} + \frac{p_{10-1}}{\kappa_{10-1}} + \frac{p_{-1-10}}{\kappa_{-1-10}} + \frac{p_{0-1-1}}{\kappa_{0-1-1}} + \frac{p_{-10-1}}{\kappa_{-10-1}},$$

$$\left[\frac{p}{\kappa}\right]_3 = \frac{p_{111}}{\kappa_{111}} + \frac{p_{-1-1-1}}{\kappa_{-1-1-1}} + \frac{p_{-111}}{\kappa_{-111}} + \frac{p_{1-11}}{\kappa_{1-11}} + \frac{p_{11-1}}{\kappa_{11-1}} + \frac{p_{-1-11}}{\kappa_{-1-11}} + \frac{p_{1-1-1}}{\kappa_{1-1-1}} + \frac{p_{-11-1}}{\kappa_{-11-1}}.$$

In the frame of finite-element methods, this strategy is equivalent to combining the consistent and lumped mass matrices.

In summary, the discrete wave equation can be compactly written as

$$\omega^2\left(w_{m1}\left[\frac{p}{\kappa}\right]_0 + \frac{w_{m2}}{6}\left[\frac{p}{\kappa}\right]_1 + \frac{w_{m3}}{12}\left[\frac{p}{\kappa}\right]_2 + \frac{w_{m4}}{8}\left[\frac{p}{\kappa}\right]_3\right) + \{[w_1\mathbf{S}_{B_c} + w_2/3(\mathbf{S}_{B_x} + \mathbf{S}_{B_y} + \mathbf{S}_{B_z}) + w_3/4(\mathbf{S}_{B_1} + \mathbf{S}_{B_2} + \mathbf{S}_{B_3} + \mathbf{S}_{B_4})]\mathbf{p}\}_{000} = s_{000}.$$

The pattern of the resulting discrete impedance matrix is illustrated in Figure 2 for a small $8\times8\times8$ grid. The matrix is band diagonal with fringes. Their are 27 nonzero coefficients per row. The bandwidth is of the order of the product of the two smallest dimensions, $\mathcal{O}(n_i\times n_j)$, where $n_i\times n_j = \min(n_x\times n_y, n_x\times n_z, n_y\times n_z)$ (64 in this case). The impedance matrix has a symmetric pattern caused by the reciprocity of Green's functions, but is unsymmetric because of the mass-term averaging and the discretization of the PML-absorbing boundary conditions. The coefficients are complex because of the PML-absorbing boundary conditions and the use of complex velocities to introduce attenuation in the simulation (Toksöz and Johnston, 1981).

NUMERICAL DISPERSION ANALYSIS

In this section, we present a classic dispersion analysis to assess the accuracy of the stencil. We first derive the expression of the phase velocity for the infinite homogeneous scheme. Since the phase velocity depends on the weighting coefficients w_{m1}, w_{m2}, w_{m3}, w_1, and w_2, we solve in a second step an optimization problem for the estimation of these coefficients, minimizing the numerical phase velocity dispersion.

Consider an infinite homogeneous velocity model of velocity c and a constant density equal to 1. From Appendix D, the discrete wave equation (without PML conditions) reduces to

$$\frac{\omega^2}{c^2}\left(w_{m1}\Sigma_0 + \frac{w_{m2}}{6}\Sigma_1 + \frac{w_{m3}}{12}\Sigma_2 + \frac{w_{m4}}{8}\Sigma_3\right) + \frac{w_1}{h^2}(\Sigma_1 - 6\Sigma_0) + \frac{w_2}{3h^2}\left(\frac{\Sigma_2}{2} + \Sigma_1 - 12\Sigma_0\right) + \frac{w_3}{4h^2}\left(\frac{3}{2}\Sigma_3 - \Sigma_2 + 2\Sigma_1 - 12\Sigma_0\right) = 0, \quad (6)$$

where

$$\Sigma_0 = p_{000},$$

$$\Sigma_1 = p_{100} + p_{010} + p_{001} + p_{-100} + p_{0-10} + p_{00-1},$$

$$\Sigma_2 = p_{110} + p_{011} + p_{101} + p_{-110} + p_{0-11} + p_{-101} + p_{1-10} + p_{01-1} + p_{10-1} + p_{-1-10} + p_{0-1-1} + p_{-10-1},$$

$$\Sigma_3 = p_{111} + p_{-1-1-1} + p_{-111} + p_{1-11} + p_{11-1} + p_{-1-11} + p_{1-1-1} + p_{-11-1}.$$

Following a classic harmonic approach, we insert the discrete expression of a plane wave, $p_{lmn} = e^{-ihk(l\cos\phi\cos\theta + m\cos\phi\sin\theta + n\sin\phi)}$, where $i^2 = -1$, in equation 6. The phase velocity is given by ω/k. We define the normalized phase velocity by $\tilde{v}_{ph} = v_{ph}/c$ and introduce $G = \lambda/h = 2\pi/kh$, the number of points per wavelength λ.

After some straightforward manipulations, we obtain the following expression for the phase velocity:

$$\tilde{v}_{ph} = \frac{G}{\sqrt{2J}\pi}\sqrt{w_1(3 - C) + \frac{w_2}{3}(6 - C - B) + \frac{2w_3}{4}(3 - 3A + B - C)}, \quad (7)$$

where $J = (w_{m1} + 2w_{m2}C + 4w_{m3}B + 8w_{m4}A)$ with

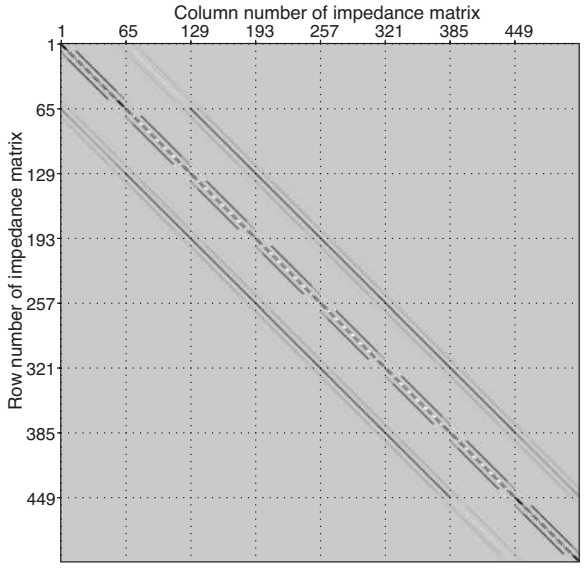


Figure 2. Pattern of the square impedance matrix discretized with the 27-point stencil. The matrix is band diagonal with fringes. The bandwidth is $\mathcal{O}(2N_1N_2)$ where N_1 and N_2 are the two smallest dimensions of the 3D grid. The number of rows/columns in the matrix is $N_1\times N_2\times N_3$. In the figure, $N_1 = N_2 = N_3 = 8$.

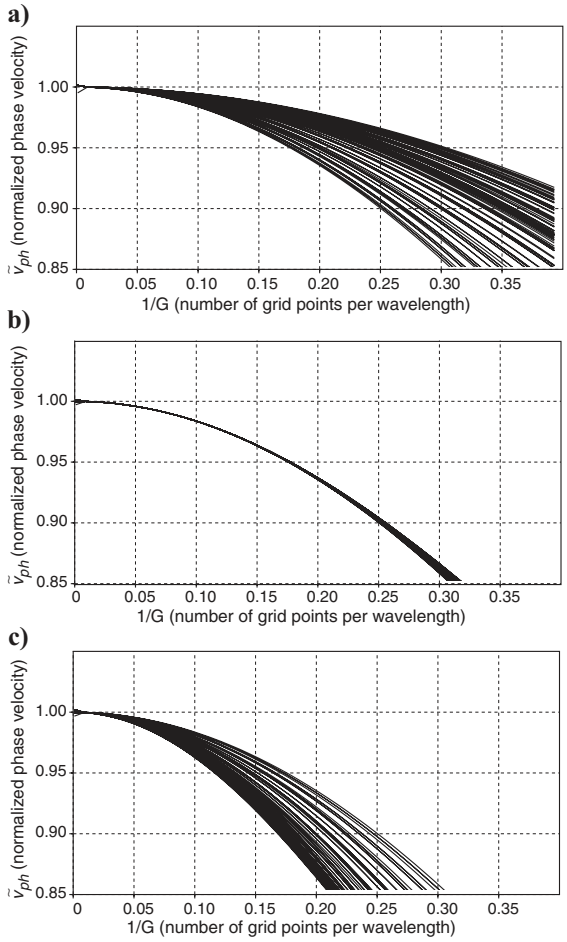


Figure 3. Dispersion curves for phase velocity. Normalized phase velocity as a function of the inverse of the number of grid points per wavelength is plotted. (a) Stencil 1 without mass averaging. (b) Stencil 2 without mass averaging. (c) Stencil 3 without mass averaging. The curves are plotted for angles θ and ϕ ranging from 0° to 45°.

$$A = \cos a \cos b \cos c,$$

$$B = \cos a \cos b + \cos a \cos c + \cos b \cos c,$$

$$C = \cos a + \cos b + \cos c,$$

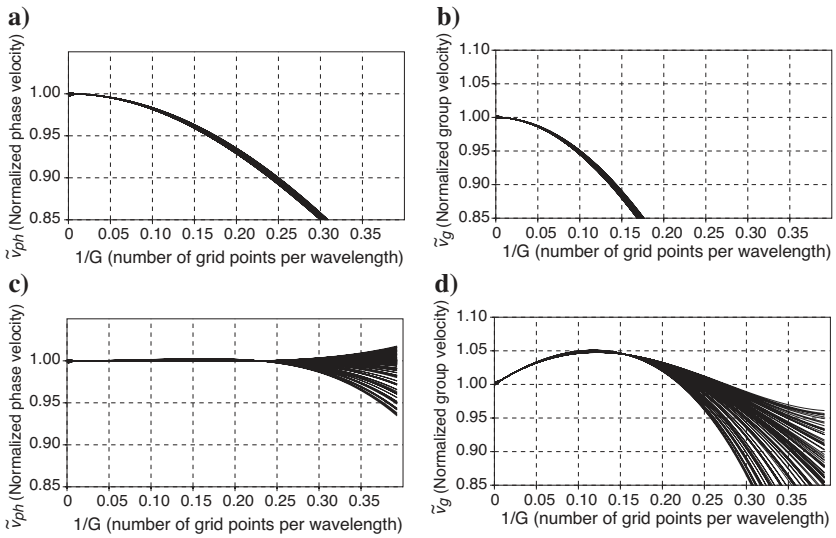
and $a = 2\pi/G \cos\phi \cos\theta, b = 2\pi/G \cos\phi \sin\theta, c = 2\pi/G \sin\phi$. One can easily check that $\tilde{v}_{ph} \rightarrow 1$ when $G \rightarrow \infty$ for $J = 1$ and for the three cases $(w_1, w_2, w_3) = (1,0,0)$, $(0,1,0)$, and $(0,0,1)$ for all ϕ and θ. This validates the expression of the phase velocity, equation 7.

We look for the five independent parameters w_{m1}, w_{m2}, w_{m3}, w_1, and w_2, which minimize the least-square norm of the misfit $(1. - \tilde{v}_{ph})$. The two remaining coefficients w_{m4} and w_3 are inferred from equations 3 and 5, respectively. We estimated these coefficients by a global optimization procedure based on a very fast simulating annealing algorithm (Sen and Stoffa, 1995). We minimize the cost function for five angles ϕ and θ spanning 0° and 45° and for four values of G ranging between four and 10.

We found $w_{m1} = 0.4964958$, $w_{m2} = 0.4510125$, $w_{m3} = 0.052487$, $w_1 = 1.8395265 \times 10^{-5}$, and $w_2 = 0.890077$, which give $w_{m4} = 0.45523 \times 10^{-5}$ and $w_3 = 0.1099046$. The coefficients show that stencils 2 and 3 are dominant contributers in the mixed-grid stencil. On the other hand, the mass coefficients show a dominant contribution of the coefficients located at the collocation point and at the nodes associated with stencil 1.

The dispersion curves for stencils 1, 2, and 3 without mass averaging are shown in Figure 3. These stencils, used individually, would require up to 40 grid points per wavelength. The phase- and group-velocity dispersion curves for the mixed stencil without mass averaging are shown in Figure 4a and b. Note how the dispersion curves for different incidence angles are focused, illustrating the isotropy of the stencil. However, the accuracy of the stencil remains poor. Combining the mixed-grid discretization strategy with mass averaging allows us to mitigate both numerical anisotropy and dispersion (Figure 4c and d). The phase-velocity dispersion curve suggests that a discretization rule of four grid-points per wavelength can be used. If the wave propagation modeling algorithm is used as a tool for full-waveform inversion, this discretization rule is optimal in the sense that theoretical resolution of full-waveform inversion at normal incidence is half the wavelength (Miller et al., 1987). The sampling theorem states that four points per wavelength is the maximum grid interval for sampling half a wavelength without aliasing.

Figure 4. Dispersion curves for phase and group velocities. Normalized phase and group velocities as a function of the inverse of the number of grid points per wavelength are plotted.(a) Phase- and (b) group-velocity dispersion curves for mixed-grid stencil without mass averaging. (c-d) Same as (a-b) but with mass averaging. The curves are plotted for angles θ and ϕ ranging from 0° to 45°.



THE MASSIVELY PARALLEL DIRECT SOLVER — MUMPS

To solve the sparse system of linear equations, we used the direct multifrontal massively parallel solver (MUMPS) designed for distributed memory platforms. The method and its underlying algorithmic aspects are extensively described by Amestoy et al. (2006), MUMPS Team (2007), and Guermouche et al. (2003). The MUMPS solver is based on a multifrontal method (Duff and Reid, 1983; Liu, 1992). In this approach, the resolution of the linear system is subdivided into three main tasks. The first is an analysis phase or symbolic factorization. This analysis is currently sequential in MUMPS. Reordering of the matrix coefficients is first performed in order to minimize fill-ins, namely, the additional nonzero coefficients introduced during the elimination process. We used the METIS algorithm, which is based on a hybrid multilevel nested-dissection and multiple minimum degree algorithm (Karypis and Kumar, 1998). Then, the dependency graph that describes the order in which the matrix can be factored is estimated as well as the memory required to perform the subsequent numerical factorization. If several simulations need to be performed in slightly different velocity models during iterative full-waveform inversion, the analysis phase needs to be performed only once per frequency. The second task is numerical factorization. The third task is the solution phase performed by forward and backward substitutions. During the resolution phase, multiple-shot solutions can be simultaneously computed from the LU factors, taking advantage of basic linear algebra subprograms (BLAS3) library, and either assembled on the host or kept distributed on the processors for subsequent parallel computations.

We performed the factorization and the resolution phases in single precision. To reduce the condition number of the matrix, row and column scaling is applied in MUMPS before factorization. The sparsity of the matrix and suitable equilibration made single precision factorization accurate enough for the 2D and 3D problems we tackled (Hustedt et al., 2004). If single precision factorization were not considered accurate enough for very large problems, an alternative approach to double precision factorization could be postprocessing of the solution by a simple and fast iterative refinement performed in double precision (Demmel, 1997, p. 60 and 61; Kurzak and Dongarra 2006; Langou et al., 2006).

NUMERICAL EXPERIMENTS

Validation in a homogeneous medium

To check this 3D mixed-grid stencil, the numerical solution is compared with the analytical one computed for homogeneous medium. The P-wave velocity is 4 km/s and the density is 2000 kg/m^3. The frequency is 4 Hz. The grid size is $161 \times 101 \times 51$ and the grid interval is 250 m, which represents one fourth of the propagated wavelength. The PML layer is discretized with four grid points and hence spans one wavelength. All the simulations presented in the following were computed on a PC cluster composed of dual-core 2.4 GHz biprocessors with 8 GB of memory per node. The interconnect is Infiniband 4X. The peak power is 19.2 Gflops per node.

We used 60 processes for this simulation. The total memory allocated for factorization was 38 GB. The real time for factorization was 409 s. The number of entries in LU factors was 2.15×10^9, corresponding to 17.2 GB of memory allocated for storage of these factors. The time for resolution was 1.67 s for one right-hand side.

Figure 5 shows a monofrequency pressure wavefield. The PML absorption is efficient, although there are only four grid points in the PML layers.

Figure 6 compares some vertical and horizontal graphs extracted from the pressure wavefield (Figure 5) with the 3D analytical solution. The agreement is quite good.

2.5D corner-edge model

The second validation of the 3D FD stencil was performed by comparing the solution of the 3D FDFD code computed in a 2.5D velocity model (i.e., homogeneous along y) with a line source along y and the solution of a 2D FDFD code (Hustedt et al., 2004) computed in a slice of the 2.5D velocity model. The velocity model is composed of two homogeneous half-spaces delineated by horizontal and vertical interfaces forming a corner (the so-called corner-edge model). Velocities in the upper-left and the lower-right media are 4 and 5 km/s, respectively. The corner is at $x = 7.8$ km, $z = 2.5$ km. The source is at ($x = 10.95$ km, $z = 1.45$ km). The grid interval is 150 m corresponding to 6.7 grid points per minimum wavelength. The computed frequency is 4 Hz. The grid size is $nx = 201 \times ny = 141 \times nz = 71$. There are seven grid points in the PML layers, which corresponds to one minimum wavelength. The modeling was computed using 120 processors. The elapsed time for factorization was 1950 s. The total amount of memory allocated during factorization was 133.96 GB. The number of entries in LU factors was 7.72×10^9, corresponding to 61.76 GB in a single-precision complex. The average and maximum memory allocated to each processor during factorization was 1.11 and 1.37 GB, respectively. The time for resolution was 5.81 s for one right-hand side.

A perspective view of the 2.5D wavefield computed with the 3D FDFD code is shown in Figure 7a. A slice of the cube is compared with the solution of the 2D code in Figure 7b and c, showing good agreement between the two solutions. The difference between the 2D and 3D solutions is shown in Figure 7d with the same scale as Figure 7b and c.

3D SEG/EAGE overthrust model

In order to test the feasibility of using 3D acoustic FDFD wave propagation modeling to address realistic problems, we used the 3D SEG/EAGE overthrust velocity model designed by the oil exploration community to test 3D seismic imaging methods (Aminzadeh et al., 1994, 1995).

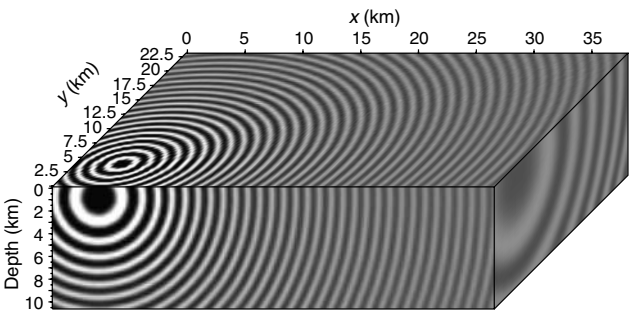


Figure 5. Example monofrequency wavefield (real part) computed in the homogeneous model.

The 3D SEG/EAGE overthrust model is a constant-density acoustic model covering an area $20 \times 20 \times 4.65$ km. It is discretized with 25-m cubic cells representing a uniform mesh of $801 \times 801 \times 187$ nodes. The minimum and maximum velocities in the overthrust model are 2.2 and 6.0 km/s, respectively.

We performed three simulations for frequencies of 5, 7, and 10 Hz. The grid interval was adapted to the modeled frequency in order to fit the dispersion condition of four grid points per minimum wavelength. This leads to grid intervals of 100, 75, and 50 m for the 5-, 7-, and 10-Hz frequencies, respectively. Before undersampling, the initial velocity model was smoothed using a 2D Gaussian filter whose correlation length was adapted to the coarse grid interval to avoid creating artificial staircases in the undersampled velocity model.

We performed the 5-Hz simulation on 32 dual-core biprocessor nodes. The 7- and 10-Hz simulations were performed using 48 nodes. The smoothed and undersampled models as well as some examples of monochromatic wavefields are shown in Figures 8–10 for 5-, 7-, and 10-Hz frequencies, respectively.

Simulation statistics are summarized in Tables 1 and 2. The computational grid spans $20 \times 20 \times 4.65$ km, $20 \times 15 \times 4.65$ km, and $20 \times 5 \times 4.65$ km at 5-, 7-, and 10-Hz frequencies, respectively, because of memory limitation. The real time for the resolution step is 3.8, 9.8, and 10.3 s for the 5-, 7-, and 10-Hz frequencies, respectively. This highlights the potential benefit of FDFD simulations based on a direct solver if enough core memory is available to store the LU factors. Resolutions for multiple shots were performed in sequence, i.e., the right-hand terms were processed sequentially. The resolution times discussed here may be significantly reduced by using the MUMPS functionality, which allows simultaneous resolution for multiple shots, taking advantage of the BLAS3 library (MUMPS team, 2007), but we have not yet attempted to do this. Simultaneous resolution for multiple shots assumes that multiple solutions can be distributed in the core memory on different processors. Part of the memory occupied by the overheads during factorization is released before resolution and can be used for this purpose.

The number of LU factors N_{LU} is $\mathcal{O}(2 \times N_x \times N_y \times N_z \times N_y \times N_z)$ for the natural ordering of the matrix, assuming that $N_x > \max(N_y, N_z)$. Using METIS ordering, we obtained for each simulation $N_{\mathrm{LU}} = 6.48 \times N_{\mathrm{LU_{ord}}}$, $3.84 \times N_{\mathrm{LU_{ord}}}$, and $2.46 \times N_{\mathrm{LU_{ord}}}$, where N_{LU} and $N_{\mathrm{LU_{ord}}}$ stand for the number of LU factors without and with METIS ordering, respectively.

We used the MUMPS analysis phase to obtain an estimate of $N_{\mathrm{LU_{ord}}}$ for a larger 3D grid with the same ratio between the three dimensions as for the overthrust model. For a grid of $340 \times 340 \times 85$, $N_{\mathrm{LU_{ord}}}$ was estimated to be 47.98×10^9, corresponding to 384 GB in a single precision complex. These grid dimensions would allow modeling of the 9.5-Hz frequency in the full overthrust model.

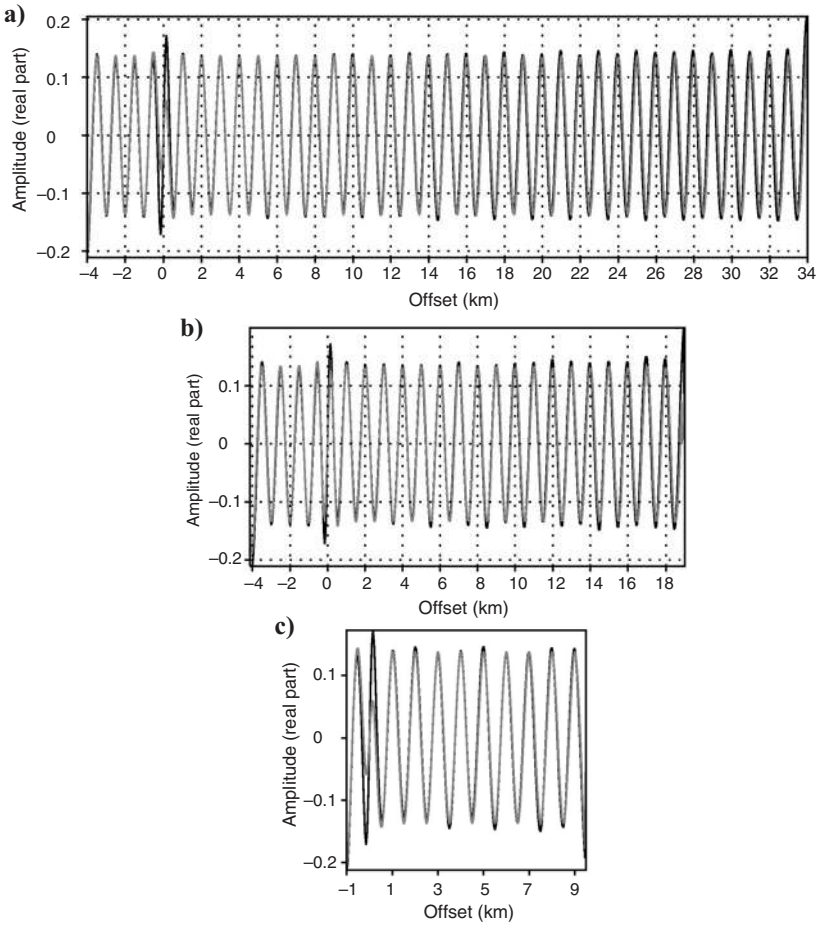


Figure 6. Comparison between (black) numerical and (gray) analytical solutions along (a) x, (b) y, and (c) z. The graphs run through the point source position. Amplitudes were corrected for 3D geometric spreading.

COMPLEXITY ANALYSIS

We performed a series of numerical tests to assess the memory and time scalability (behavior of the algorithm when the problem size augments) of the MUMPS solver. We considered cubic grids of increasing dimension N (the total number of grid points in the grid is N^3). Statistics of the simulations are summarized in Table 3. The number of processors N_{proc} was 36 for N ranging between 40 and 80, and 50 for N between 85 and 110. The theoretical memory complexity of the LU factorization is $\mathcal{O}(N^4)$ if nested-dissection ordering is used, and the number of floating-point operations during factorization is $\mathcal{O}(N^6)$ (e.g., Ashcraft and Liu, 1998). The CPU-time complexity of the resolution step is $\mathcal{O}(N^4)$. The number of LU factors ($N_{\mathrm{LU_{ord}}}$) as a function of the grid dimension N is shown in Figure 11a; the gray curve corresponds to numerical factorization, and the black curve corresponds to estimation returned by symbolic factorization performed during analysis. There is good agreement between the analysis estimation and the numerical factorization, suggesting stable behavior of the latter. These two curves are compared with different curves corresponding to some power of N. The curve showing the ratio $N_{\mathrm{LU_{ord}}}/N^4$ exhibits a reasonable horizontal pattern until $N = 140$, suggesting that the real complexity of the factorization after reordering is closed to the theoretical one (Figure 11b). This curve also reveals that a significant multiplicative factor of around 35 must be applied to N^4 to obtain the number of LU factors. This may be because of the size of the separators in the nested-dissection al-

gorithm associated with the 27-point stencil. Therefore, we conclude that the memory complexity of the factorization is $\mathcal{O}(35N^4)$.

Simulation of a 10-Hz frequency in a $10 \times 10 \times 10$ km cubic grid (corresponding to a $210 \times 210 \times 210$ grid for a minimum velocity of 2 km/s) would require around 320 GB of RAM to store the LU single precision complex factors. We speculate that the 15-Hz frequency in a $10 \times 10 \times 10$ km volume (corresponding to a $300 \times 300 \times 300$ grid) would require around 1 Tbyte of RAM or disk to store the LU factors. This 3D volume is representative of the domain spanned by one shot patch of a 3D wide-azimuth multichannel seismic experiment (see Michell et al., 2006 for a description of a 3D wide-azimuth streamer survey). Therefore, this amount of memory may be representative of the memory required to process a 3D wide-azimuth multichannel seismic reflection dataset by FDFD migration based on a two-way wave equation with a maximum frequency of 15 Hz (see Plessix and Mulder, 2004 and Mulder and Plessix, 2004, for a description of two-way wave equation FDFD migration).

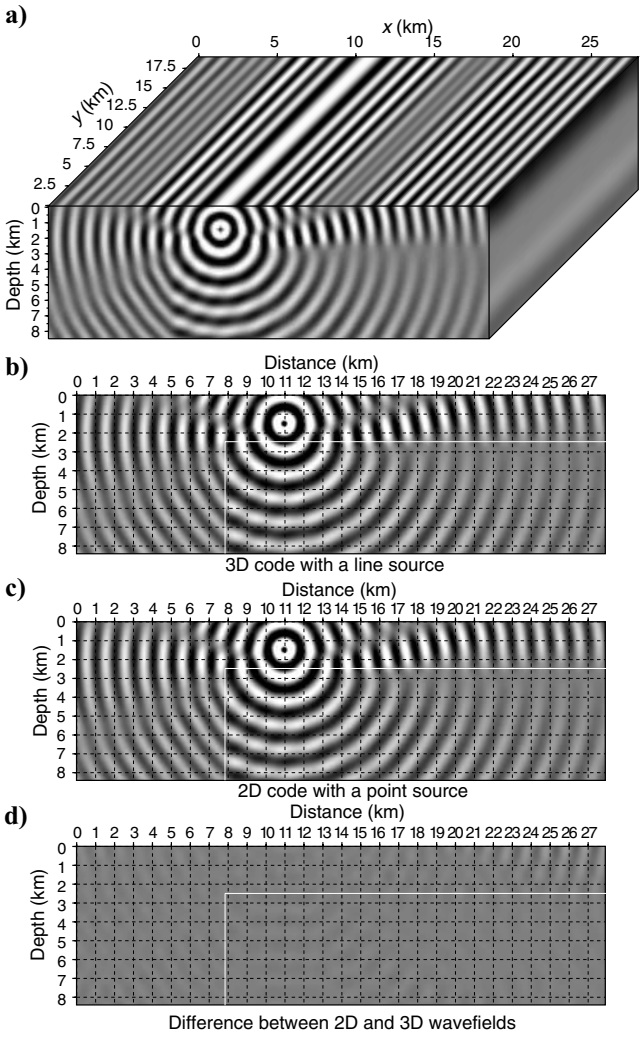


Figure 7. Monofrequency wavefield computed in the 2.5D corner-edge model. The velocity in the two layers are 4 and 5 km/s, respectively. Frequency is 4 Hz. (a) Perspective view of the 2.5D wavefield. (b) 3D solution with a line source (real part). (c) 2D solution with a point source (real part). (d) Difference between (a) and (b). (a–c) are plotted with the same amplitude scale.

The memory overheads required during factorization are shown in Figure 12. This curve suggests that the memory required by the overheads is of the order of that required to store the LU factors.

The ratio between the real time for factorization over the time complexity of the factorization $\mathcal{O}(N^6)$ is plotted as a function of N in Figure 13. The reasonably horizontal pattern of the curve shows the consistency between numerical experience and theoretical complexity analysis. The sharp decrease of the ratio between $N = 80$ and $N = 85$ is caused by an increase in the number of processors from 36 to 50 (see Table 3).

The efficiency of the parallel factorization is illustrated in Figure 14. Efficiency is defined by $T_{\text{seq}}/T_{\text{par}} \cdot N_{\text{proc}}$ where T_{seq} is the effective time for sequential factorization, T_{par} is the effective time for parallel factorization, and N_{proc} is the number of processors involved in the parallel factorization. For this test, $N = 50$. Using 36 processors, we obtained an efficiency of 0.35, which leads to a speed-up of 12.6.

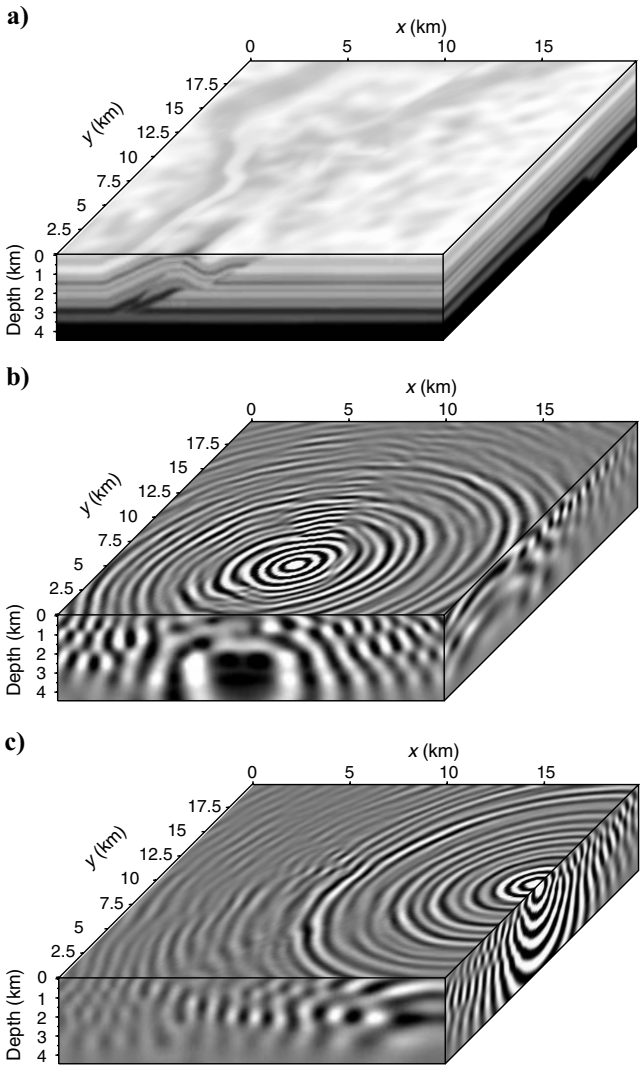


Figure 8. (a) Smoothed and undersampled overthrust model. Mesh spacing is 100 m. (b–c) Two examples of 5-Hz monofrequency wavefields in the full overthrust model.

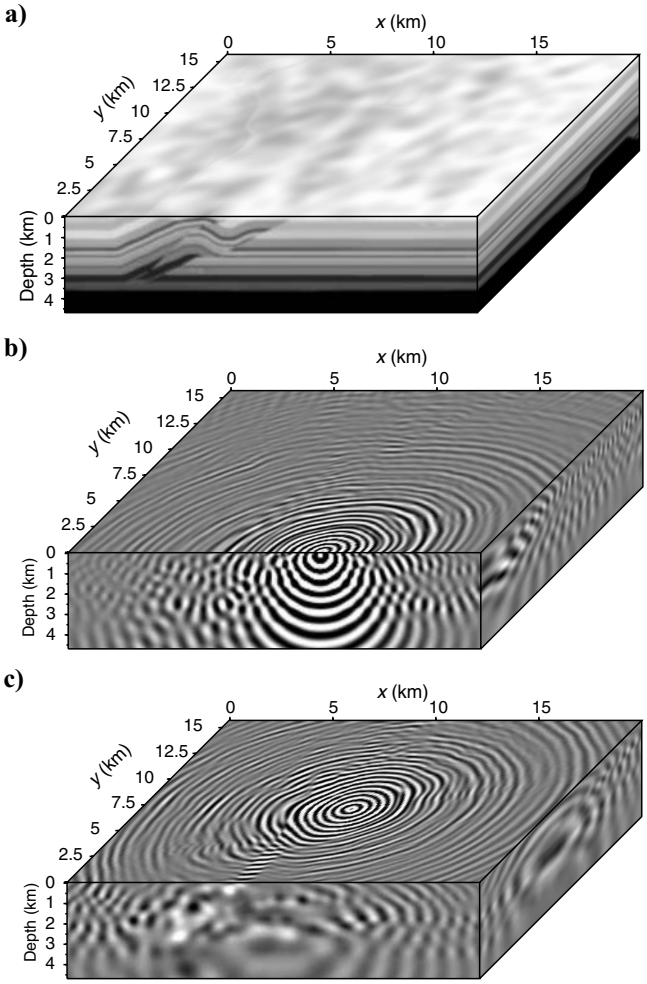


Figure 9. (a) Smoothed and undersampled overthrust model. Mesh spacing is 75 m. (b–c) Two examples of 7-Hz monofrequency wavefields in a fraction of the overthrust model.

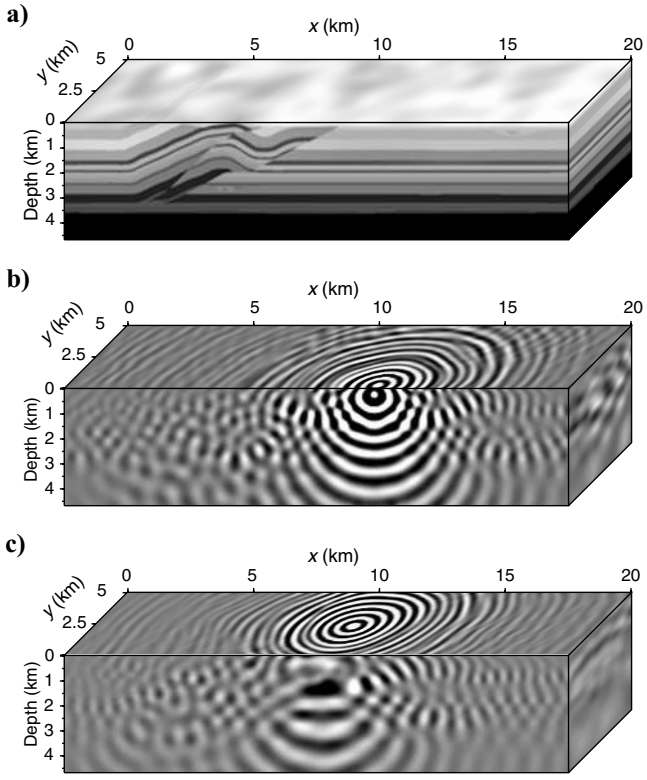


Figure 10. (a) Smoothed and undersampled overthrust model. Mesh spacing is 50 m. (b–c) Two examples of 10-Hz monofrequency wavefields in a fraction of the overthrust model.

Table 1. Model dimensions for overthrust simulations: (N_x,N_y,N_z), size of the 3D FD grid including PML layers; N_{PML}, number of grid points in the PML layers; h, grid interval; f, modeled frequency (Hz); (L_x,L_y,L_z), physical dimensions of the propagating model (i.e., without PML); O, number of unknowns; NNZ, number of nonzero coefficients.

N_x	N_y	N_z	N_{PML}	h (m)	f (Hz)	L_x (km)	L_y (km)	L_z (km)	O/1e6	NNZ//1e6
207	207	53	5	100	5	19.6	19.6	4.2	2.27	61.31
275	218	71	5	75	7	19.8	15.5	4.5	4.25	114.92
409	109	102	5	50	10	19.9	4.9	4.55	4.54	122.77

Table 2. Computational effort for overthrust simulations: N_{CPU}, number of processes; T_A(s), time for analysis; T_D(s), time for sparse matrix distribution; T_F(s), elapsed time for factorization; T_R(s), elapsed time for resolution for 1 RHS; M_P (MB), maximum space per working processor during factorization; $\widetilde{M}_P$ (MB), average space per working processor during factorization; M_F (GB), total amount of memory allocated during factorization; N_n, number of nodes in the tree; S_F, maximum frontal size; $N_{\text{LU}_{\text{ord}}}$, number of entries in LU factors; IS_F: integer space for factors.

N_{CPU}	T_A	T_D	T_F	T_R	$\widetilde{M}_P$	M_P	M_F	N_n	S_F	$N_{\text{LU}_{\text{ord}}}$/1$e$6	IS_F/1e6
128	94	16	1181	3.8	958	1249	143	139491	22886	7693	57
192	257	53	6425	9.3	1603	2891	385	258170	42814	22517	183
192	401	56	9826	10.3	1820	2406	449	280181	33693	29073	230

Table 3. Memory and time scalabilities of factorization and resolution: N_{CPU}, number of process; N, dimension of the cubic computational grid; $T_D(s)$, time for sparse matrix distribution; $T_F(s)$, elapsed time for factorization; $T_R(s)$, elapsed time for resolution for 1 RHS; M_P (MB), average space per working processor during factorization; M_F (MB), total amount of memory allocated during factorization; $N_{LU_{ord}}$, number of LU factors.

N_{CPU}	N	T_D (s)	T_F	T_R (s)	M_P (Mb)	M_F (Mb)	$N_{LU_{ord}}/1e6$
36	40	0.43	8.7957	0.17	73	2639	83.3
36	50	0.67	31.24	0.33	147	5302	207.08
36	60	0.99	81.08	0.59	255	9193	435.39
36	70	1.50	198.98	0.90	449	16177	821.62
36	80	2.08	481.80	1.71	840	30248	1423.93
50	90	3.09	628.18	1.94	908	45449	2292.02
50	100	4.10	1402.16	3.52	1508	75406	3521.87
50	110	5.34	2031.37	11.86	1871	93555	5240.76
64	115	7.52	2485.90	7.05	1819	116440	6472.24

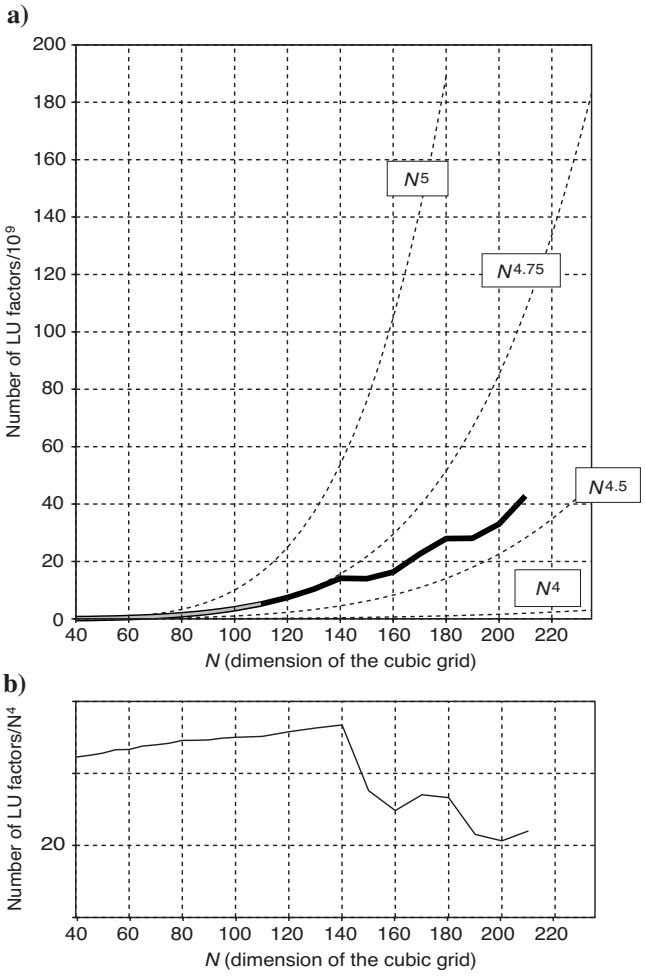


Figure 11. Evolution of the memory complexity (number of LU factors) as a function of the grid size N. (a) Comparison between the curve $N_{LU_{ord}} = f(N)$ and some power of N. The gray curve was inferred from numerical factorization and the black curve from symbolic factorization. (b) The ratio $N_{LU_{ord}}/N^4$ is shown.

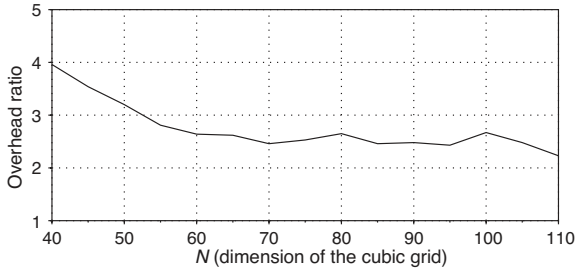


Figure 12. Memory overhead as a function of the grid dimension N. The memory overhead is the ratio between the total memory allocated during factorization and the memory to store the LU factors.

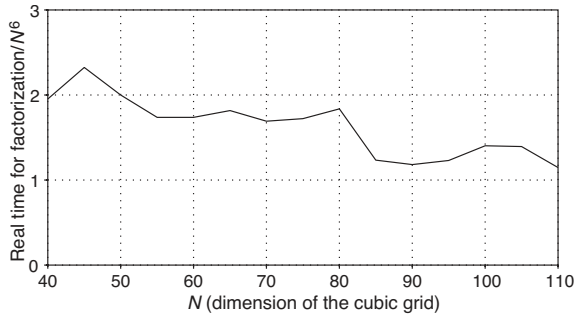


Figure 13. Ratio between the real time for factorization and the theoretical time complexity N^6 as a function of N.

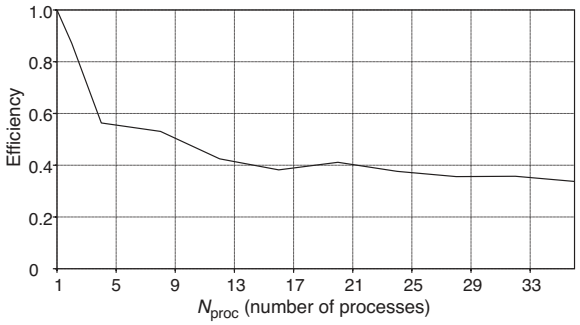


Figure 14. Efficiency of LU-factorization parallelism. The ratio between the effective time used by a sequential and a parallel execution multiplied by the number of processors is presented as a function of the number of processors N_{proc} for $N = 50$.

CONCLUSION AND PERSPECTIVES

We presented a feasibility study of 3D finite-difference frequency-domain modeling of acoustic wave propagation based on a massively parallel direct solver. This modeling algorithm was developed as a tool for frequency-domain full-waveform inversion, which requires modeling of a limited number of frequencies and a large number of sources (typically, hundreds to thousands depending on the design of the acquisition). We first designed an optimal FD stencil for the frequency-domain formulation. The key properties of this stencil are its local support, which minimizes the dependencies in the graph of the matrix, and a discretization rule of four grid points per minimum wavelength which is optimal with respect to the theoretical resolution of full-waveform inversion.

We presented several simulations in the 3D SEG/EAGE overthrust model. Our results confirm the huge memory complexity of the factorization and the CPU efficiency of the resolution. We estimated the memory complexity of the factorization to be $\mathcal{O}(35N^4)$ for an $N \times N \times N$ grid.

Future work of the MUMPS team will include the use of out-of-core memory (disk space) in the MUMPS solver and the parallelization of the analysis phase, which should allow us to address larger problems for a given distributed memory platform. The applicability of the approach presented in this paper, which is strongly related to the memory issue, will probably be dependent on the evolution of the cluster technologies in the next decade. Such technologies may evolve toward clusters with more processors and less memory per processor. Since the memory scalability of a massively parallel direct solver is not naturally high, this evolution should not be favorable to our approach. However, the out-of-core version of direct solvers should allow a more flexible management of memory for these platforms. The second possible evolution is the increasing development of hybrid architectures combining shared and distributed memory models. For these architectures, one can take advantage of threaded BLAS libraries by allowing fewer MPI processes than actual processors. This should avoid duplication of symbolic information and improve memory scalability.

Another perspective of a methodological nature is the evolution toward hybrid direct/iterative methods relying on domain decomposition methods. One possiblity is to use an iterative approach after incomplete/partial direct factorization. Another is to apply the direct factorization to subdomains of limited dimensions and to use a Schwarz domain decomposition method to retrieve the solution in the full domain. In both cases, the method presented in this paper will provide a valuable tool to tackle the direct part in the hybrid approach.

The modeling code presented in this paper will be implemented in a 3D massively parallel frequency-domain full-waveform inversion program. Implementation of the frequency-domain full-waveform inversion will contribute to a number of applications as well as toward the comparison with time-domain inversions.

ACKNOWLEDGMENTS

This work was funded by the SEISCOPE consortium sponsored by BP, CGG-Veritas, ExxonMobil, Shell, and Total. Partial support is acknowledged for the use of computer facilities at the MESOCENTRE SIGAMM computer center. We are grateful to A. Miniussi for his assistance during the installation of the modeling and MUMPS programs on the MESOCENTRE SIGAMM. We thank the anonymous reviewers and associate editor Clement Kostov for critical comments and careful reading of the manuscript.

APPENDIX A

3D FREQUENCY-DOMAIN WAVE EQUATION WITH UNSPLIT PML CONDITION

In this appendix, we develop the 3D frequency-domain acoustic wave equation with unsplit PML conditions from the time-domain acoustic equation with split PML conditions. The velocity-stress system of hyperbolic equations for the 3D damped wave equation in the time domain is given by

$$
\begin{aligned}
&\frac{\partial p_x(x,y,z,t)}{\partial t} + \gamma_x(x) p_x(x,y,z,t) = \kappa(x,y,z) \cdot \frac{\partial v_x(x,y,z,t)}{\partial x},\\
&\frac{\partial p_y(x,y,z,t)}{\partial t} + \gamma_y(y) p_y(x,y,z,t) = \kappa(x,y,z) \cdot \frac{\partial v_y(x,y,z,t)}{\partial x},\\
&\frac{\partial p_z(x,y,z,t)}{\partial t} + \gamma_z(z) p_z(x,y,z,t) = \kappa(x,y,z) \cdot \frac{\partial v_z(x,y,z,t)}{\partial z},\\
&\frac{\partial v_x(x,y,z,t)}{\partial t} + \gamma_x(x) v_x(x,y,z,t) = b(x,y,z) \cdot \frac{\partial p(x,y,z,t)}{\partial x}\\
&\qquad\qquad + f_x(x,y,z,t),\\
&\frac{\partial v_y(x,y,z,t)}{\partial t} + \gamma_y(y) v_y(x,y,z,t) = b(x,y,z) \cdot \frac{\partial p(x,y,z,t)}{\partial x}\\
&\qquad\qquad + f_x(x,y,z,t).\\
&\frac{\partial v_z(x,y,z,t)}{\partial t} + \gamma_z(z) v_z(x,y,z,t) = b(x,y,z) \cdot \frac{\partial p(x,y,z,t)}{\partial z}\\
&\qquad\qquad + f_x(x,y,z,t). \qquad \text{(A-1)}
\end{aligned}
$$

The pressure $p(x,y,z,t)$ is split in to three components, $p_x(x,y,z,t)$, $p_y(x,y,z,t)$, and $p_z(x,y,z,t)$, for separation of the horizontal and vertical derivatives (Berenger, 1994).

The pressure wavefield is deduced through the simple addition of three unphysical acoustic fields p_x, p_y, and p_z: $p(x,y,z,t) = p_x(x,y,z,t) + p_y(x,y,z,t) + p_z(x,y,z,t)$. These unphysical acoustic fields p_x, p_y and p_z are used to account for the PML-absorbing boundary conditions (see Zhang and Ballmann, 1997, or Operto et al., 2002, for the application of PML boundary conditions to the SH and P-SV wave equations). The 1D functions γ_x, γ_y, and γ_z define the PML damping behavior in the PML layers surrounding the model on all four sides. These functions differ from zero only inside the PML layers. In the PML layers, we used $\gamma(x) = c_{\mathrm{PML}} \cos\left(\frac{\pi}{2}\frac{x}{L}\right)$ where L denotes the width of the PML layer and x is a local coordinate in the PML layer whose origin is located at the outer edges of the model. The scalar c_{PML} is defined by trial and error depending on the width of the PML layer (i.e., several values of c_{PML} are tested for representative PML widths; the best value is the one for which the reflections coming from the edges of the model have the smallest amplitude).

We transform the system of equations to the Fourier domain and introduce the new functions $\xi_x(x) = 1 + i\gamma_x(x)/\omega$, $\xi_y(y) = 1 + i\gamma_y(y)/\omega$, and $\xi_z(z) = 1 + i\gamma_z(z)/\omega$ in the right-hand sides of the equations:

$$
\begin{aligned}
\frac{-i\omega}{\kappa(x,y,z)}p_x(x,y,z,\omega) &= \frac{1}{\xi_x(x)}\frac{\partial v_x(x,y,z,\omega)}{\partial x},\\
\frac{-i\omega}{\kappa(x,y,z)}p_y(x,y,z,\omega) &= \frac{1}{\xi_y(y)}\frac{\partial v_y(x,y,z,\omega)}{\partial y},\\
\frac{-i\omega\xi_z(z)}{\kappa(x,y,z)}p_z(x,y,z,\omega) &= \frac{1}{\xi_z(z)}\frac{\partial v_z(x,y,z,\omega)}{\partial z},\\
-i\omega v_x(x,y,z,\omega) &= \frac{b(x,y,z)}{\xi_x(x)}\frac{\partial p(x,y,z,\omega)}{\partial x}\\
&\quad + f_x(x,y,z,\omega),\\
-i\omega v_y(x,y,z,\omega) &= \frac{b(x,y,z)}{\xi_y(y)}\frac{\partial p(x,y,z,\omega)}{\partial y}\\
&\quad + f_y(x,y,z,\omega),\\
-i\omega v_z(x,y,z,\omega) &= \frac{b(x,y,z)}{\xi_z(z)}\frac{\partial p(x,y,z,\omega)}{\partial z}\\
&\quad + f_z(x,y,z,\omega). \qquad \text{(A-2)}
\end{aligned}
$$

The ξ_x, ξ_y, and ξ_z functions are not applied to the external forces because we assume these forces are located outside the PML layers where these functions are equal to one.

The first three equations of equation system A-2 can be added,

$$
\begin{aligned}
\frac{-i\omega}{\kappa(x,y,z)}p(x,y,z,\omega) &= \frac{1.}{\xi_x(x)}\frac{\partial v_x(x,y,z,\omega)}{\partial x}\\
&\quad + \frac{1.}{\xi_y(y)}\frac{\partial v_y(x,y,z,\omega)}{\partial y}\\
&\quad + \frac{1.}{\xi_z(z)}\frac{\partial v_z(x,y,z,\omega)}{\partial z},\\
v_x(x,y,z,\omega) &= -\frac{b(x,y,z)}{i\omega\xi_x(x)}\frac{\partial p(x,y,z,\omega)}{\partial x}\\
&\quad + f_x(x,y,z,\omega),\\
v_y(x,y,z,\omega) &= -\frac{b(x,y,z)}{i\omega\xi_y(y)}\frac{\partial p(x,y,z,\omega)}{\partial y}\\
&\quad + f_y(x,y,z,\omega),\\
v_z(x,y,z,\omega) &= -\frac{b(x,y,z)}{i\omega\xi_z(z)}\frac{\partial p(x,y,z,\omega)}{\partial z}\\
&\quad + f_z(x,y,z,\omega), \qquad \text{(A-3)}
\end{aligned}
$$

which provides equation 1 in the text.

After Fourier transform, introduction of the ξ functions and summation of the first three equations of system A-2, the physical pressure wavefield $p(x,y,z,\omega)$ was reintroduced in equation 1. This highlights the benefit of the frequency-domain formulation for designing unsplit PML-absorbing boundary conditions.

APPENDIX B

ROTATED DIFFERENTIAL OPERATORS

We give below the expression of the spatial derivatives with respect to x, y, and z as a function of the spatial derivatives with respect to the coordinates of each basis (B_x, B_y, B_z, B_1, B_2, B_3, B_4). These expressions are reinjected into equation system 1 before discretization.

$$
\begin{aligned}
B_x\text{:}\ \frac{\partial}{\partial y} &= \frac{\sqrt{2}}{2}\left(\frac{\partial}{\partial y_x} - \frac{\partial}{\partial z_x}\right), \quad \frac{\partial}{\partial z} = \frac{\sqrt{2}}{2}\left(\frac{\partial}{\partial y_x} + \frac{\partial}{\partial z_x}\right).\\
B_y\text{:}\ \frac{\partial}{\partial x} &= \frac{\sqrt{2}}{2}\left(\frac{\partial}{\partial x_y} + \frac{\partial}{\partial z_y}\right), \quad \frac{\partial}{\partial z} = \frac{\sqrt{2}}{2}\left(-\frac{\partial}{\partial x_y} + \frac{\partial}{\partial z_y}\right).\\
B_z\text{:}\ \frac{\partial}{\partial x} &= \frac{\sqrt{2}}{2}\left(\frac{\partial}{\partial x_z} - \frac{\partial}{\partial y_z}\right), \quad \frac{\partial}{\partial y} = \frac{\sqrt{2}}{2}\left(\frac{\partial}{\partial x_z} + \frac{\partial}{\partial y_z}\right).
\end{aligned}
\qquad \text{(B-1)}
$$

$$
\begin{aligned}
B_1\text{:}\ \frac{\partial}{\partial x} &= \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_2} + \frac{\partial}{\partial d_3}\right), \quad \frac{\partial}{\partial y} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_1} - \frac{\partial}{\partial d_3}\right),\\
&\quad \frac{\partial}{\partial z} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_1} - \frac{\partial}{\partial d_2}\right).\\
B_2\text{:}\ \frac{\partial}{\partial x} &= \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_1} + \frac{\partial}{\partial d_4}\right), \quad \frac{\partial}{\partial y} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_2} - \frac{\partial}{\partial d_4}\right),\\
&\quad \frac{\partial}{\partial z} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_1} - \frac{\partial}{\partial d_2}\right).\\
B_3\text{:}\ \frac{\partial}{\partial x} &= \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_1} + \frac{\partial}{\partial d_4}\right), \quad \frac{\partial}{\partial y} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_1} - \frac{\partial}{\partial d_3}\right),\\
&\quad \frac{\partial}{\partial z} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_3} - \frac{\partial}{\partial d_4}\right).\\
B_4\text{:}\ \frac{\partial}{\partial x} &= \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_2} + \frac{\partial}{\partial d_3}\right), \quad \frac{\partial}{\partial y} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_2} - \frac{\partial}{\partial d_4}\right),\\
&\quad \frac{\partial}{\partial z} = \frac{\sqrt{3}}{2}\left(\frac{\partial}{\partial d_3} - \frac{\partial}{\partial d_4}\right).
\end{aligned}
\qquad \text{(B-2)}
$$

APPENDIX C

SECOND-ORDER CENTERED STAGGERED-GRID STENCILS

We give below the expression of the second-order accurate staggered-grid FD stencil of the first-order spatial derivative on each coordinate system. Indices 1/2 indicate intermediate positions with respect to the reference grid i,j,k. In the example below, X stands for the particle velocity wavefields in the sense that their spatial derivatives need to be estimated on the reference pressure grid.

$$\left[\frac{\partial X}{\partial x}\right]_{i,j,k} = \frac{1}{h}(X_{i+1/2,j,k} - X_{i-1/2,j,k}),$$
$$\left[\frac{\partial X}{\partial y}\right]_{i,j,k} = \frac{1}{h}(X_{i,j+1/2,k} - X_{i,j-1/2,k}),$$
$$\left[\frac{\partial X}{\partial z}\right]_{i,j,k} = \frac{1}{h}(X_{i,j,k+1/2} - X_{i,j,k-1/2}). \quad \text{(C-1)}$$

$$B_x\text{: } \left[\frac{\partial X}{\partial y_x}\right]_{i,j,k} = \frac{1}{\sqrt{2}h}(X_{i,j+1/2,k+1/2} - X_{i,j-1/2,k-1/2}),$$
$$\left[\frac{\partial X}{\partial z_x}\right]_{i,j,k} = \frac{1}{\sqrt{2}h}(X_{i,j-1/2,k+1/2} - X_{i,j+1/2,k-1/2}).$$
$$B_y\text{: } \left[\frac{\partial X}{\partial x_y}\right]_{i,j,k} = \frac{1}{\sqrt{2}h}(X_{i+1/2,j,k-1/2} - X_{i-1/2,j,k+1/2}),$$
$$\left[\frac{\partial X}{\partial z_y}\right]_{i,j,k} = \frac{1}{\sqrt{2}h}(X_{i+1/2,j,k+1/2} - X_{i-1/2,j,k-1/2}).$$
$$B_z\text{: } \left[\frac{\partial X}{\partial x_z}\right]_{i,j,k} = \frac{1}{\sqrt{2}h}(X_{i+1/2,j+1/2,k} - X_{i-1/2,j-1/2,k}),$$
$$\left[\frac{\partial X}{\partial y_z}\right]_{i,j,k} = \frac{1}{\sqrt{2}h}(X_{i-1/2,j+1/2,k} - X_{i+1/2,j-1/2,k}).$$
$$B_1\text{: } \left[\frac{\partial X}{\partial d_1}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j+1/2,k+1/2} - X_{i-1/2,j-1/2,k-1/2}),$$
$$\left[\frac{\partial X}{\partial d_2}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j+1/2,k-1/2} - X_{i-1/2,j-1/2,k+1/2}),$$
$$\left[\frac{\partial X}{\partial d_3}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j-1/2,k+1/2} - X_{i-1/2,j+1/2,k-1/2}).$$
$$B_2\text{: } \left[\frac{\partial X}{\partial d_1}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j+1/2,k+1/2} - X_{i-1/2,j-1/2,k-1/2}),$$
$$\left[\frac{\partial X}{\partial d_2}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j+1/2,k-1/2} - X_{i-1/2,j-1/2,k+1/2}),$$
$$\left[\frac{\partial X}{\partial d_4}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j-1/2,k-1/2} - X_{i-1/2,j+1/2,k+1/2}).$$
$$B_3\text{: } \left[\frac{\partial X}{\partial d_1}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j+1/2,k+1/2} - X_{i-1/2,j-1/2,k-1/2}),$$
$$\left[\frac{\partial X}{\partial d_3}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j-1/2,k+1/2} - X_{i-1/2,j+1/2,k-1/2}),$$
$$\left[\frac{\partial X}{\partial d_4}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j-1/2,k-1/2} - X_{i-1/2,j+1/2,k+1/2}).$$
$$B_4\text{: } \left[\frac{\partial X}{\partial d_2}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j+1/2,k-1/2} - X_{i-1/2,j-1/2,k+1/2}),$$
$$\left[\frac{\partial X}{\partial d_3}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j-1/2,k+1/2} - X_{i-1/2,j+1/2,k-1/2}),$$
$$\left[\frac{\partial X}{\partial d_4}\right]_{i,j,k} = \frac{1}{\sqrt{3}h}(X_{i+1/2,j-1/2,k-1/2} - X_{i-1/2,j+1/2,k+1/2}), \quad \text{(C-2)}$$

where h is the mesh spacing in the uniform grid.

APPENDIX D

PARSIMONIOUS STAGGERED-GRID FINITE-DIFFERENCE WAVE EQUATIONS

Implementation of the stencils given in Appendix B in equation system 1 and elimination of the particle velocity wavefields lead to the following second-order discrete wave equations in pressure. The buoyancy needs to be interpolated at intermediate positions with respect to the reference grid.

In this study, we used the following averaging

$$b_{1/200} = 0.5(b_{000} + b_{100}),$$
$$b_{1/21/20} = 0.25(b_{000} + b_{100} + b_{010} + b_{110}),$$
$$b_{1/21/21/2} = 0.125(b_{000} + b_{100} + b_{010} + b_{001} + b_{110} + b_{011} + b_{101} + b_{111}). \quad \text{(D-1)}$$

$$B_c\text{: } \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 h^2}\left[\frac{b_{1/200}}{\xi x_{1/2}}(P_{100} - P_{000}) - \frac{b_{-1/200}}{\xi x_{-1/2}}(P_{000} - P_{-100})\right] + \frac{1}{\xi y_0 h^2}\left[\frac{b_{01/20}}{\xi y_{1/2}}(P_{010} - P_{000}) - \frac{b_{0-1/20}}{\xi y_{-1/2}}(P_{000} - P_{0-10})\right] + \frac{1}{\xi z_0 h^2}\left[\frac{b_{001/2}}{\xi z_{1/2}}(P_{001} - P_{000}) - \frac{b_{00-1/2}}{\xi z_{-1/2}}(P_{000} - P_{00-1})\right] = S_{000}. \quad \text{(D-2)}$$

$$B_x\colon \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 h^2}\left[\frac{b_{1/200}}{\xi x_{1/2}}(P_{100} - P_{000}) - \frac{b_{-1/200}}{\xi x_{-1/2}}(P_{000} - P_{-100})\right] + \frac{1}{\xi y_0 4h^2}\left[\frac{b_{01/21/2}}{\xi y_{1/2}}(P_{011} - P_{000} - P_{001} + P_{010}) - \frac{b_{0-1/2-1/2}}{\xi y_{-1/2}}(P_{000} - P_{0-1-1} - P_{0-10} + P_{00-1}) - \frac{b_{0-1/21/2}}{\xi y_{-1/2}}(P_{001} - P_{0-10} - P_{0-11} + P_{000}) + \frac{b_{01/2-1/2}}{\xi y_{1/2}}(P_{010} - P_{00-1} - P_{000} + P_{01-1})\right] + \frac{1}{\xi z_0 4h^2} \times \left[\frac{b_{01/21/2}}{\xi z_{1/2}}(P_{011} - P_{000} + P_{001} - P_{010}) - \frac{b_{0-1/2-1/2}}{\xi z_{-1/2}}(P_{000} - P_{0-1-1} + P_{0-10} - P_{00-1}) + \frac{b_{0-1/21/2}}{\xi z_{1/2}}(P_{001} - P_{0-10} + P_{0-11} - P_{000}) - \frac{b_{01/2-1/2}}{\xi z_{-1/2}}(P_{010} - P_{00-1} + P_{000} - P_{01-1})\right] = S_{000}. \tag{D-3}$$

$$B_y\colon \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 4h^2}\left[\frac{b_{1/20-1/2}}{\xi x_{1/2}}(P_{10-1} - P_{000} + P_{100} - P_{00-1}) - \frac{b_{-1/201/2}}{\xi x_{-1/2}}(P_{000} - P_{-101} - P_{001} + P_{-100}) + \frac{b_{1/201/2}}{\xi x_{1/2}}(P_{100} - P_{001} + P_{101} - P_{000}) - \frac{b_{-1/20-1/2}}{\xi x_{-1/2}}(P_{00-1} - P_{-100} + P_{000} - P_{-10-1})\right] + \frac{1}{\xi y_0 h^2}\left[\frac{b_{01/20}}{\xi y_{1/2}}(P_{010} - P_{000}) - \frac{b_{0-1/20}}{\xi y_{-1/2}}(P_{000} - P_{0-10})\right] + \frac{1}{\xi z_0 4h^2}\left[-\frac{b_{1/20-1/2}}{\xi z_{-1/2}}(-P_{10-1} + P_{000} + P_{100} - P_{00-1}) + \frac{b_{1/201/2}}{\xi z_{1/2}}(-P_{000} + P_{-101} + P_{001} - P_{-100}) + \frac{b_{1/201/2}}{\xi z_{1/2}}(-P_{100} + P_{001} + P_{101} - P_{000}) - \frac{b_{-1/20-1/2}}{\xi z_{-1/2}}(-P_{00-1} - P_{-100} + P_{000} - P_{-10-1})\right] = S_{000}. \tag{D-4}$$

$$B_z\colon \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 4h^2}\left[\frac{b_{1/21/20}}{\xi x_{1/2}}(P_{110} - P_{000} - P_{010} + P_{100}) - \frac{b_{-1/2-1/20}}{\xi x_{-1/2}}(P_{000} - P_{-1-10} - P_{-100} + P_{0-10}) - \frac{b_{-1/21/20}}{\xi x_{-1/2}}(P_{010} - P_{-100} - P_{-110} + P_{000}) - \frac{b_{1/2-1/20}}{\xi x_{1/2}}(P_{100} - P_{0-10} - P_{000} + P_{1-10})\right] + \frac{1}{\xi y_0 4h^2} \times \left[\frac{b_{1/21/20}}{\xi y_{1/2}}(P_{110} - P_{000} + P_{010} - P_{100}) - \frac{b_{-1/2-1/20}}{\xi y_{-1/2}}(P_{000} - P_{-1-10} + P_{-100} - P_{0-10}) + \frac{b_{-1/21/20}}{\xi y_{1/2}}(P_{010} - P_{-100} + P_{-110} - P_{000}) - \frac{b_{1/2-1/20}}{\xi y_{-1/2}}(P_{100} - P_{0-10} + P_{000} - P_{1-10})\right] + \frac{1}{\xi z_0 h^2}\left[\frac{b_{001/2}}{\xi z_{1/2}}(P_{001} - P_{000}) - \frac{b_{00-1/2}}{\xi z_{-1/2}}(P_{000} - P_{00-1})\right] = S_{000}. \tag{D-5}$$

$$B_1\colon \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 4h^2}\left[\frac{b_{1/21/2-1/2}}{\xi x_{1/2}}(P_{11-1} - P_{000} + P_{100} - P_{01-1}) - \frac{b_{-1/2-1/21/2}}{\xi x_{-1/2}}(P_{000} - P_{-1-11} + P_{0-11} - P_{-100}) + \frac{b_{1/2-1/21/2}}{\xi x_{1/2}}(P_{100} - P_{0-11} + P_{1-11} - P_{000}) - \frac{b_{-1/21/2-1/2}}{\xi x_{-1/2}}(P_{01-1} - P_{-100} + P_{000} - P_{-11-1})\right] + \frac{1}{\xi y_0 4h^2}\left[\frac{b_{1/21/21/2}}{\xi y_{1/2}}(P_{111} - P_{000} - P_{101} + P_{010}) - \frac{b_{-1/2-1/2-1/2}}{\xi y_{-1/2}}(P_{000} - P_{-1-1-1} - P_{0-10} - P_{-10-1}) - \frac{b_{1/2-1/21/2}}{\xi y_{-1/2}}(P_{101} - P_{0-10} - P_{1-11} + P_{000}) + \frac{b_{-1/21/2-1/2}}{\xi y_{1/2}}(P_{010} - P_{-10-1} - P_{000} + P_{-11-1})\right] + \frac{1}{\xi z_0 4h^2}\left[\frac{b_{1/21/21/2}}{\xi z_{1/2}}(P_{111} - P_{000} - P_{110} + P_{001}) - \frac{b_{-1/2-1/2-1/2}}{\xi z_{-1/2}}(P_{000} - P_{-1-1-1} - P_{00-1} + P_{-1-10})$$

$$- \frac{b_{1/21/2-1/2}}{\xi z_{-1/2}}(P_{110} - P_{00-1} - P_{11-1} + P_{000}) + \frac{b_{-1/2-1/21/2}}{\xi z_{1/2}}(P_{001} - P_{-1-10} - P_{000} + P_{-1-11})\Bigg] = S_{000}. \quad \text{(D-6)}$$

$$B_2: \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 4h^2}\Bigg[\frac{b_{1/21/21/2}}{\xi x_{1/2}}(P_{111} - P_{000} + P_{100} - P_{011}) - \frac{b_{-1/2-1/2-1/2}}{\xi x_{-1/2}}(P_{000} - P_{-1-1-1} + P_{0-1-1} - P_{-100}) + \frac{b_{1/2-1/2-1/2}}{\xi x_{1/2}}(P_{100} - P_{0-1-1} + P_{1-1-1} - P_{000}) - \frac{b_{-1/21/21/2}}{\xi x_{-1/2}}(P_{011} - P_{-100} + P_{000} - P_{-111})\Bigg] + \frac{1}{\xi y_0 4h^2}\Bigg[\frac{b_{1/21/2-1/2}}{\xi y_{1/2}}(P_{11-1} - P_{000} - P_{10-1} + P_{010}) - \frac{b_{-1/2-1/21/2}}{\xi y_{-1/2}}(P_{000} - P_{-1-11} - P_{0-10} + P_{-101}) - \frac{b_{1/2-1/2-1/2}}{\xi y_{-1/2}}(P_{10-1} - P_{0-10} - P_{1-1-1} + P_{000}) + \frac{b_{-1/21/21/2}}{\xi y_{1/2}}(P_{010} - P_{-101} - P_{000} + P_{-111})\Bigg] + \frac{1}{\xi z_0 4h^2}\Bigg[\frac{b_{1/21/21/2}}{\xi z_{1/2}}(P_{111} - P_{000} - P_{110} + P_{001}) - \frac{b_{-1/2-1/2-1/2}}{\xi z_{-1/2}}(P_{000} - P_{-1-1-1} - P_{00-1} + P_{-1-10}) - \frac{b_{1/21/2-1/2}}{\xi z_{-1/2}}(P_{110} - P_{00-1} - P_{11-1} + P_{000}) + \frac{b_{-1/2-1/21/2}}{\xi z_{1/2}}(P_{001} - P_{-1-10} - P_{000} + P_{-1-11})\Bigg] = S_{000}. \quad \text{(D-7)}$$

$$B_3: \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 4h^2}\Bigg[\frac{b_{1/21/21/2}}{\xi x_{1/2}}(P_{111} - P_{000} + P_{100} - P_{011}) - \frac{b_{-1/2-1/2-1/2}}{\xi x_{-1/2}}(P_{000} - P_{-1-1-1} + P_{0-1-1} - P_{-100}) + \frac{b_{1/2-1/2-1/2}}{\xi x_{1/2}}(P_{100} - P_{0-1-1} + P_{1-1-1} - P_{000}) - \frac{b_{-1/21/21/2}}{\xi x_{-1/2}}(P_{011} - P_{-100} + P_{000} - P_{-111})\Bigg] + \frac{1}{\xi y_0 4h^2}\Bigg[\frac{b_{1/21/21/2}}{\xi y_{1/2}}(P_{111} - P_{000} - P_{101} + P_{010}) - \frac{b_{-1/2-1/2-1/2}}{\xi y_{-1/2}}(P_{000} - P_{-1-1-1} - P_{0-10} + P_{-10-1}) - \frac{b_{1/2-1/21/2}}{\xi y_{-1/2}}(P_{101} - P_{0-10} - P_{1-11} + P_{000}) + \frac{b_{-1/21/2-1/2}}{\xi y_{1/2}}(P_{010} - P_{-10-1} - P_{000} + P_{-11-1})\Bigg] + \frac{1}{\xi z_0 4h^2}\Bigg[\frac{b_{1/2-1/21/2}}{\xi z_{1/2}}(P_{1-11} - P_{000} - P_{1-10} + P_{001}) - \frac{b_{-1/21/2-1/2}}{\xi z_{-1/2}}(P_{000} - P_{-11-1} - P_{00-1} + P_{-110}) - \frac{b_{1/2-1/2-1/2}}{\xi z_{-1/2}}(P_{1-10} - P_{00-1} - P_{1-1-1} + P_{000}) + \frac{b_{-1/21/21/2}}{\xi z_{1/2}}(P_{001} - P_{-110} - P_{000} + P_{-111})\Bigg] = S_{000}. \quad \text{(D-8)}$$

$$B_4: \frac{\omega^2}{K_{000}}P_{000} + \frac{1}{\xi x_0 4h^2}\Bigg[\frac{b_{1/21/2-1/2}}{\xi x_{1/2}}(P_{11-1} - P_{000} + P_{100} - P_{01-1}) - \frac{b_{-1/2-1/21/2}}{\xi x_{-1/2}}(P_{000} - P_{-1-11} + P_{0-11} - P_{-100}) + \frac{b_{1/2-1/21/2}}{\xi x_{1/2}}(P_{100} - P_{0-11} + P_{1-11} - P_{000}) - \frac{b_{-1/21/2-1/2}}{\xi x_{-1/2}}(P_{01-1} - P_{-100} + P_{000} - P_{-11-1})\Bigg] + \frac{1}{\xi y_0 4h^2}\Bigg[\frac{b_{1/21/2-1/2}}{\xi y_{1/2}}(P_{11-1} - P_{000} - P_{10-1} + P_{010}) - \frac{b_{-1/2-1/21/2}}{\xi y_{-1/2}}(P_{000} - P_{-1-11} - P_{0-10} + P_{-101}) - \frac{b_{1/2-1/2-1/2}}{\xi y_{-1/2}}(P_{10-1} - P_{0-10} - P_{1-1-1} + P_{000}) + \frac{b_{-1/21/21/2}}{\xi y_{1/2}}(P_{010} - P_{-101} - P_{000} + P_{-111})\Bigg] + \frac{1}{\xi z_0 4h^2}\Bigg[\frac{b_{1/2-1/21/2}}{\xi z_{1/2}}(P_{1-11} - P_{000} - P_{1-10} + P_{001}) - \frac{b_{-1/21/2-1/2}}{\xi z_{-1/2}}(P_{000} - P_{-11-1} - P_{00-1} + P_{-110}) - \frac{b_{1/2-1/2-1/2}}{\xi z_{-1/2}}(P_{1-10} - P_{00-1} - P_{1-1-1} + P_{000}) + \frac{b_{-1/21/21/2}}{\xi z_{1/2}}(P_{001} - P_{-110} - P_{000} + P_{-111})\Bigg] = S_{000}. \quad \text{(D-9)}$$

REFERENCES

Amestoy, P. R., A. Guermouche, J. Y. L'Excellent, and S. Pralet, 2006, Hybrid scheduling for the parallel solution of linear systems: Parallel computing, **32**, 136–156.

Aminzadeh, F., N. Burehard, L. Nicoletis, F. Rocca, and K. Wyatt, 1994, SEG/EAEG 3-D modeling project: Second report: The Leading Edge, **13**, 949–952.

——– 1995, SEG/EAEG 3-D modeling project: Third report: The Leading Edge, **14**, 125–128.

Ashcraft, C., and J. W. H. Liu, 1998, Robust ordering of sparse matrices using multisection: SIAM Journal on Matrix Analysis and Applications, **19**, 816–832.

Berenger, J.-P., 1994, A perfectly matched layer for absorption of electromagnetic waves: Journal of Computational Physics, **114**, 185–200.

Brenders, A. J., and R. G. Pratt, 2006, Efficient waveform tomography for lithospheric imaging: Implications for realistic 2D acquisition geometries and low frequency data: Geophysical Journal International, **168**, 152–170.

Demmel, J. W., 1997, Applied numerical linear algebra: SIAM.

Duff, I. S., and J. K. Reid, 1983, The multifrontal solution of indefinite sparse symmetric linear systems: ACM Transactions on Mathematical Software, **9**, 302–325.

Guermouche, A., J. Y. L'Excellent, and G. Utard, 2003, Impact of reordering on the memory of a multifrontal solver: Parallel computing, **29**, 1191–1218.

Hustedt, B., S. Operto, and J. Virieux, 2004, Mixed-grid and staggered-grid finite difference methods for frequency-domain acoustic wave modelling: Geophysical Journal International, **157**, 1269–1296.

Jo, C.-H., C. Shin, and J. H. Suh, 1996, An optimal 9-point, finite-difference, frequency-space, 2-D scalar wave extrapolator: Geophysics, **61**, 529–537.

Karypis, G., and V. Kumar, 1998, METIS—A software package for partitioning unstructured graphs, partitioning meshes and computing fill-reducing orderings of sparse matrices: Version 4.0, University of Minnesota.

Kurzak, J., and J. Dongarra, 2006, Implementation of the mixed-precision high performance LINPACK benchmark on the CELL processor: Technical report ut-cs-06–580, University of Tennessee; http://icl.cs.utk.edu/iter-ref/.

Langou, J., J. Langou, P. Luszczek, J. Kurzak, A. Buttari, and J. Dongarra, 2006, LAPACK working note 175: Exploiting the performance of 32 bit floating point arithmetic in obtaining 64 bit accuracy (revisiting iterative refinement for linear systems): Technical report, University of Tennessee; http://icl.cs.utk.edu/iter-ref/.

Liu, J. W. H., 1992, The multifrontal method for sparse matrix solution: Theory and practice: SIAM Review, **34**, 82–109.

Luo, Y., and G. Schuster, 1990, Parsimonious staggered grid finite differencing of the wave equation: Geophysical Research Letters, **17**, 155–158.

Marfurt, K., 1984, Accuracy of finite-difference and finite-elements modeling of the scalar and elastic wave equation: Geophysics, **49**, 533–549.

Michell, S., J. Sharp, and D. Chergotis, 2007, Dual-azimuth versus wide-azimuth technology as applied in subsalt imaging of Mad Dog Field: A case study: The Leading Edge, 470–478.

Miller, D., M. Oristaglio, and G. Beylkin, 1987, A new slant on seismic imaging: Migration and integral geometry: Geophysics, **52**, 943–964.

Mulder, W. A., and R. E. Plessix, 2004, How to choose a subset of frequencies in frequency-domain finite-difference migration: Geophysical Journal International, **158**, 801–812.

MUMPS Team, 2007, Multifrontal massively parallel solver users' guide: ENSEEIHT-ENS Lyon, http://graal.ens-lyon.fr/MUMPS, accessed June 15, 2007.

Operto, S., C. Ravaut, L. Improta, J. Virieux, A. Herrero, and P. dell'Aversana, 2004, Quantitative imaging of complex structures from multi-fold wide-aperture data: A case study: Geophysical Prospecting, **52**, 625–651.

Operto, S., J. Virieux, J.-X. Dessa, and G. Pascal, 2007, Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography: Application to the eastern Nankai trough: Journal of Geophysical Research, 111, http://dx.doi.org/10.1029/2005JB003835.

Operto, S., J. Virieux, B. Hustedt, and F. Malfanti, 2002, Adaptive wavelet-based finite-difference modelling of SH-wave propagation: Geophysical Journal International, **148**, 477–499.

Plessix, R.-E., 2006, A robust three-dimensional iterative solver for the time-harmonic wave equation: 68th Annual Conference and Exhibition, EAGE, Extended Abstracts, H047.

Plessix, R.-E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

——–, 2004, Velocity models from frequency-domain waveform tomography: Past, present and future: Presented at the 66th Annual conference and Exhibition, EAGE.

Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Pratt, R. G., Z. Song, and M. Warner, 1996, Two-dimensional velocity models from wide-angle seismic data by wavefield inversion: Geophysical Journal International, **124**, 323–340.

Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. dell'Aversana, 2004, Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions: Application to a thrust belt: Geophysical Journal International, **159**, 1032–1056.

Riyanti, C. D., Y. A. Erlangga, R. E. Plessix, W. A. Mulder, C. Vuik, and C. Oosterlee, 2006, A new iterative solver for the time-harmonic wave equation: Geophysics, **71**, no. 5, E57–E63.

Saenger, E. H., N. Gold, and A. Shapiro, 2000, Modeling the propagation of elastic waves using a modified finite-difference grid: Wave motion, **31**, 77–92.

Sen, M., and P. Stoffa, 1995, Global optimization methods in geophysical inversion: Elsevier Science Publ-Co, Inc.

Štekl, I., and C. Pain, 2002, 3D frequency domain visco-acoustic modeling using rotated finite difference operators: 64th Annual Conference and Exhibition, EAGE, Expanded Abstracts, C27.

Štekl, I., C. Pain, and M. R. Warner, 2002, Frequency-domain 3D modelling and waveform inversion: What can be done today?: Sub-basalt Imaging, Expanded Abstracts, 190.

Štekl, I., and R. G. Pratt, 1998, Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators: Geophysics, **63**, 1779–1794.

Takeuchi, N., R. J. Geller, and P. R. Cummins, 2000, Complete synthetic seismograms for 3-D heterogeneous earth models computed using modified DSM operators and their applicability to inversion for earth structure: Physics of the Earth and Planetary Interiors, **119**, 1191–1218.

Toksöz, M., and D. Johnston, 1981, Seismic wave attenuation: SEG.

Virieux, J., 1984, SH wave propagation in heterogeneous media, velocity-stress finite difference method: Geophysics, **49**, 1933–1957.

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

Zhang, Y.-G., and J. Ballmann, 1997, Two techniques for the absorption of elastic waves using an artificial transition layer: Wave motion, **25**, 15–33.