ACCURATE 3D FREQUENCY-DOMAIN SEISMIC WAVE MODELING WITH THE WAVELENGTH-ADAPTIVE 27-POINT FINITE-DIFFERENCE STENCIL: A TOOL FOR FULL WAVEFORM INVERSION

A PREPRINT

Hossein S. Aghamiry
University Cote d’Azur - CNRS - IRD - OCA, Geoazur, Valbonne, France. aghamiry@geoazur.unice.fr

Ali Gholami
Institute of Geophysics, University of Tehran, Tehran, Iran. agholami@ut.ac.ir

Laure Combe
University Cote d’Azur - CNRS - IRD - OCA, Geoazur, Valbonne, France. combe@geoazur.unice.fr

Stéphane Operto
University Cote d’Azur - CNRS - IRD - OCA, Geoazur, Valbonne, France. operto@geoazur.unice.fr

August 20, 2021

ABSTRACT

Efficient frequency-domain Full Waveform Inversion (FWI) of long-offset/wide-azimuth node data can be designed with a few discrete frequencies. However, 3D frequency-domain seismic modeling remains challenging since it requires solving a large and sparse linear indefinite system per frequency. When such systems are solved with direct methods or hybrid direct/iterative solvers, based upon domain decomposition preconditioner, finite-difference stencils on regular Cartesian grids should be designed to conciliate compactness and accuracy, the former being necessary to mitigate the fill-in induced by the Lower-Upper (LU) factorization. Compactness is classically implemented by combining several second-order accurate stencils covering the eight cells surrounding the collocation point, leading to the so-called 27-point stencil. Accuracy is obtained by applying optimal weights on the different stiffness and consistent mass matrices such that numerical dispersion is jointly minimized for several number of grid points per wavelength ($G$). However, with this approach, the same weights are used at each collocation point, leading to suboptimal accuracy in heterogeneous media. In this study, we propose a straightforward recipe to improve the accuracy of the 27-point stencil. First, we finely tabulate the values of $G$ covering the range of wavelengths spanned by the subsurface model and the frequency. Then, we estimate with a classical dispersion analysis in homogeneous media the corresponding table of optimal weights that minimize dispersion for each $G$ treated separately. We however apply a Tikhonov regularization to guarantee smooth variation of the weights with $G$. Finally, we build the impedance matrix by selecting the optimal weights at each collocation point according to the local wavelength, hence leading to a wavelength-adaptive stencil. We validate our method against analytical solutions in homogeneous and velocity gradient models and 3D heterogeneous benchmarks with sharp contrasts and complex tectonic pattern. In the latter case, the accuracy of the adaptive stencil is checked with the highly accurate solution of the convergent Born series (CBS). Each benchmark reveals the higher accuracy of the adaptive stencil relative to the non-adaptive one. We also show that, in the presence of sharp contrasts, the adaptive stencil is more accurate than a $O(\Delta t^2, \Delta h^6)$ finite-different time-domain method.
1 Introduction

Numerical seismic wave simulation as the forward problem of Full Waveform Inversion (FWI) can be implemented in the time or frequency domains (Kostin et al., 2019; Operto et al., 2015; Plessix, 2006, 2007, 2017; Vigh & Starr, 2008). Although the time-domain formulation implemented with explicit (matrix-free) time-stepping schemes has emerged as the way forward because forward engines of reverse time migration can be readily used for FWI, frequency-domain FWI still has some potential key strengths when applied on wide-azimuth long-offset stationary-recording acquisitions such as seabed acquisitions implemented with ocean bottom cables or nodes and land acquisitions. These strengths mostly rely on the fact that reliable FWI can be performed with a few discrete frequencies by reducing the redundancy with which such acquisitions locally sample subsurface wavenumbers (Pratt, 1999; Pratt & Worthington, 1990; Sirgue & Pratt, 2004). This frequency decimation potentially leads to computationally efficient algorithms and a compact volume of data. Moreover, the frequency domain provides the most convenient framework to design frequency continuation strategies, which are useful to mitigate the FWI nonlinearity generated by cycle skipping (e.g., Görszczyk et al., 2017). Finally, implementation of attenuation is straightforward and does not generate computational overheads (Toksoz & Johnston, 1981).

The main challenge in frequency-domain FWI is the solution of the time-harmonic wave equation, a boundary-value problem that requires the solution of large and sparse indefinite unsymmetric linear system per frequency with multiple right-hand sides. Although frequency-domain FWI can also be implemented with a time-domain forward engine, when the monochromatic wavefields are extracted on the fly in the loop over time steps by phase sensitive detection (Nihei & Li, 2007) or discrete Fourier transform (Sirgue et al., 2008), the benefit of the straightforward implementation of attenuation is lost with this approach. Two main categories of linear algebra methods exist to solve the linear system resulting from the discretization of the time-harmonic wave equation: Direct methods first perform a Lower-Upper (LU) decomposition of the impedance matrix before computing the solutions by forward/backward elimination (Duff et al., 2017). The strength of this approach is predictability (the solution is obtained in a finite number of operations) and the efficiency of the solution step when a large number of right-hand sides (namely, sources) should be processed as in seismic imaging. The drawbacks are the memory overhead induced by the storage of the LU factors and the limited scalability of the LU decomposition. This approach is generally considered as intractable for 3D applications although recent developments of sparse direct solvers (Amestoy et al., 2018, 2015; Kostin et al., 2019; Wang et al., 2011) and FWI case studies (Amestoy et al., 2016a; Operto & Miniussi, 2018; Operto et al., 2015) proved precisely the opposite. The second category of linear algebra methods relies on iterative solvers (Saad, 2003), which are scalable but raise the challenge of designing a suitable preconditioner for indefinite Helmholtz problems (Ernst & Gander, 2012). Domain decomposition preconditioners based on additive Schwarz are among the most popular ones to tackle this challenge (Dolean et al., 2015), where a Krylov subspace iterative method such as the generalized minimal residual (GMRES) algorithm (Saad, 1986) is used to solve the preconditioned system while direct methods are used to solve the local problems associated with the subdomains of the computational mesh (Tournier et al., 2019).

In both approaches, the discretization scheme should be designed with care to find the best trade-off between the number of degrees of freedom in the computational mesh, accuracy, and compactness. In the frequency-domain, the second-order formulation of the wave equation is generally used to minimize the number of degrees of freedom at the expense of a first-order velocity-stress formulation, which is more popular in the time domain (Virieux, 1986). Compactness is useful to minimize the numerical bandwidth of the matrix and hence mitigate the memory burden induced by the LU factorization of the global or local matrices as well as the amount of inter-domain communication when parallel domain-decomposition preconditioner are used. Two- and three-dimensional finite-difference stencils on regular Cartesian grid (constant grid interval) have been designed in the past to fulfill these specifications (Chen et al., 2012, 2013; Gosselin-Cliche & Giroux, 2014; Hustedt et al., 2004; Jo et al., 1996; Min et al., 2000; Operto et al., 2007b, 2014; Stekl & Pratt, 1998; Turkel et al., 2013). Turkel et al. (2013) develop a 3D sixth-order accurate scheme 27-point finite-difference stencil for the Helmholtz equation with variable wavenumber. In this approach derivatives of the Helmholtz equation are used to eliminate high order derivatives in the discretization error. One difficulty with this approach, regardless its technical complexity, is related to the right-hand side of the discretized equation, which involve high-order derivatives of the physical source (Turkel et al., 2013, Their equation 27). This might raise some accuracy issues when the point sources in seismic application doesn’t coincide with a grid point (Hicks, 2002). Essentially, it seems that this approach still lacks validation against highly heterogeneous and contrasted media. The other above mentioned references rely on a different paradigm for compact discretization of the Helmholtz equation, which is more oriented toward numerical optimization. Compactness is achieved with second-order accurate stencils while accuracy is achieved by linearly combining several second-order accurate stencils built on different (rotated) coordinate systems with appropriate weights and by designing a consistent mass matrix as opposed to the lumped mass matrix. The weights used to combine the different stiffness matrices and the different entries of the consistent mass matrix are estimated by solving an optimization problem, which aims to minimize numerical phase velocity dispersion in homogeneous media for a series of plane waves of different incidence and azimuth angles. This finite difference
method leads also to a 27-point stencil involving the eight cells surrounding the central grid (or collocation) point. To deal with heterogeneous media (i.e., variable wavenumber \( k \)), these weights are generally estimated such that numerical dispersion is jointly minimized for several numbers of grid points per wavelength (\( G \)). However, the accuracy of the resulting stencil is suboptimal because the same weights are used at each grid point (i.e., row of the matrix) instead of being matched to the local wavelength. This paper proposes to overcome this accuracy issue by designing a wavelength-adaptive 27-point finite-difference stencil. The straightforward principle first consists of finely tabulating the range of \( G \) commonly found in the Earth’s crust. Then, the corresponding table of weights is built by minimizing the dispersion for each tabulated \( G \), which are processed separately although some regularization can be implemented to force smooth variations of the weights with \( G \). Then, each row of the impedance matrix is built by picking in the table the weights corresponding to the local wavelength. A similar idea was proposed for the 2D 9-point stencil by Xu & Gao (2018), who concluded from basic 2D simulations that the 9-point adaptive stencil reaches the same accuracy as the 25-point counterpart (Shin & Sohn, 1998). However, this study lacks comprehensive validation of the method against structurally complex and contrasted media.

The key contribution of this study is to assess the accuracy and robustness of the adaptive 27-point stencil against 3D tectonically complex models involving sharp contrasts for FWI applications. To achieve this goal, we present several large-scale 3D numerical experiments, which show that the accuracy of the adaptive stencil is significantly improved relative to the non-adaptive counterpart without generating any computational overhead.

This paper is organized as follows. We first briefly review the principles of the 27-point stencil to introduce the different stiffness matrices involved in the stencil and the consistent mass matrix. Then, we present the dispersion analysis and the optimization algorithm that is used to estimate the tabulated weights. At the next step, we present five numerical experiments. The method is first validated against analytical solutions in homogeneous media and linear velocity models in which a large number of wavelengths are propagated. Then, we use the 3D SEG/EAGE overthrust and salt models (Aminzadeh et al., 1997) as well as the deep crustal GO_3D_OBS model (Górszczyk & Operto, 2021) to validate our method in heterogeneous media by comparison with the highly-accurate solutions of the volume-integral convergent Born series (CBS) method (Osnabrügge et al., 2016). With the 3D salt benchmark, we also show that the adaptive stencil outperforms a high-order finite-difference time-domain method in the presence of sharp contrasts.

## 2 Method

### 2.1 A review of the 27-point finite-difference frequency-domain stencil

We briefly review the main principles underlying the 27-point finite-difference frequency-domain stencil on a regular Cartesian grid. The readers are referred to Brossier et al. (2010); Operto et al. (2007b, 2014) for more details. According to the parsimonious staggered-grid formulation of Hustedt et al. (2004), we start from the 3D frequency-domain visco-acoustic velocity-stress wave equation

\[
\begin{align*}
-i\omega p(x, y, z, \omega) &= \kappa(x, y, z) \left( \frac{\partial v_x(x, y, z, \omega)}{\partial x} + \frac{\partial v_y(x, y, z, \omega)}{\partial y} + \frac{\partial v_z(x, y, z, \omega)}{\partial z} \right), \\
-i\omega v_x(x, y, z, \omega) &= b(x, y, z) \frac{\partial p(x, y, z, \omega)}{\partial x} + f_x(x, y, z, \omega), \\
-i\omega v_y(x, y, z, \omega) &= b(x, y, z) \frac{\partial p(x, y, z, \omega)}{\partial y} + f_y(x, y, z, \omega), \\
-i\omega v_z(x, y, z, \omega) &= b(x, y, z) \frac{\partial p(x, y, z, \omega)}{\partial z} + f_z(x, y, z, \omega),
\end{align*}
\]

where \( \omega \) is the angular frequency, \( \kappa(x, y, z) \) is the bulk modulus, \( b(x, y, z) \) is the buoyancy (the inverse of density), \( p(x, y, z, \omega) \) is the pressure, \( v_x(x, y, z, \omega) \), \( v_y(x, y, z, \omega) \), and \( v_z(x, y, z, \omega) \) are the components of the particle velocity vector and \( f_x \), \( f_y \) and \( f_z \) are external forces acting as the source term. We introduce perfectly-matched layer (PML) absorbing conditions in equations 2.1-2.4 through a change of coordinates in the complex space such that \( \partial_x = \frac{1}{\xi_x} \partial_z \), \( \partial_y = \frac{1}{\xi_y} \partial_z \), and \( \partial_z = \frac{1}{\xi_z} \partial_z \). We have \( \xi_x = 1 + i\frac{\gamma_x}{\omega} \), \( \xi_y = 1 + i\frac{\gamma_y}{\omega} \), and \( \xi_z = 1 + i\frac{\gamma_z}{\omega} \), where the functions \( \gamma_x \), \( \gamma_y \), and \( \gamma_z \) control the damping of the wavefield in the PMLs (Operto et al., 2007b. Their Appendix A).

Plugging \( v_x \), \( v_y \) and \( v_z \) from equations 2.2-2.4 in equation 2.1 gives a generalization of the Helmholtz equation for heterogeneous density:

\[
\frac{\omega^2}{\kappa(x)} p(x, \omega) + \frac{\partial}{\partial x} b(x) \frac{\partial p(x, \omega)}{\partial x} + \frac{\partial}{\partial y} b(x) \frac{\partial p(x, \omega)}{\partial y} + \frac{\partial}{\partial z} b(x) \frac{\partial p(x, \omega)}{\partial z} = s(x, \omega),
\]
where \( \mathbf{x} = (x, y, z) \) and \( s(\mathbf{x}, \omega) = \nabla \cdot \mathbf{f} \) denotes the pressure source. The above-mentioned elimination procedure applied on the continuous form of the wave equation can be used to discretize equation 2.5 from the well-documented finite-difference approximation of the first derivative operators involved in equations 2.1-2.4. Accordingly, the discretization procedure of equation 2.5 begins with the discretization of the spatial derivatives in equation 2.1 with classical second-order accurate staggered-grid finite-difference stencils (Saenger et al., 2000; Virieux, 1984). Then, the discretized particle velocities involved in the right-hand side of equation 2.1 are developed as a function of the pressure by discretizing equations 2.2-2.4. Then, the resulting expressions of the particle velocities are eliminated from the equation 2.1 to end up with the discrete form of equation 2.5. This parsimonious discretization procedure is applied separately on the three different rotated coordinate systems illustrated in Figure 1, namely the Cartesian coordinate system leading to the classical seven-point stencil with a cross shape (Figure 1a), three rotated coordinate systems obtained by a 45° rotation around each Cartesian axis leading to three 11-point stencils, the combination of them forming a 19-point stencil (Figure 1b), and four rotated coordinate systems defined by three diagonals of the cubic cell leading to four 27-point stencils (Figure 1c). Note also that this elimination procedure can be applied in a simpler setting by considering the approach of Min et al. (2000) extended to the 3D acoustic and elastic equations by Chen et al. (2012) and Gosselin-Cliche & Giroux (2014), respectively. In this approach, the classical second-order stencil is used to discretize the spatial derivatives in the three Cartesian directions at the collocation point but also at the neighbouring points such that the support of the mixed stencil spans the 27 points surrounding the collocation point.

After discretization, the generalized Helmholtz, equation 2.5, can be written in matrix form as
\[
[M + S] \mathbf{p} = \mathbf{s},
\]
where \( M \) and \( S \) are the mass and stiffness matrices, respectively. The stiffness matrix \( S \) is formulated as a weighted sum of the stiffness matrices formulated on each coordinate system
\[
S = w_s S_1 + \frac{w_{s2}}{3} \sum_{i=1}^{3} S_{2,i} + \frac{w_{s3}}{4} \sum_{i=1}^{4} S_{3,i},
\]
where the weights satisfy
\[
\sum_{i=1}^{3} w_{s_i} = 1,
\]
and \( S_1, S_{2,•}, \) and \( S_{3,•} \) denote the stiffness matrices associated with the three classes of coordinate system illustrated in Figure 1.

The consistent mass matrix \( M \) is built by distributing the term \( \frac{\omega^2}{\kappa(s)} \) over the 27 points as
\[
\frac{\omega^2}{\kappa} \mathbf{p} \Rightarrow \omega^2 \left( w_{m0} p_0 + \frac{w_{m1}}{6} \sum_{i=1}^{6} p_{1,i} + \frac{w_{m2}}{12} \sum_{i=1}^{12} p_{2,i} + \frac{w_{m3}}{8} \sum_{i=1}^{8} p_{3,i} \right),
\]
where the weights satisfy
\[
\sum_{i=0}^{3} w_{m_i} = 1,
\]
and the subscript 0 denotes the collocation point (Figure 1d, green point), the subscripts \( 1, • \) denote the 6 neighbour grid points involved in the classical 7-point stencil (Figure 1d, blue points), the subscripts \( 2, • \) denote the 12 additional grid points involved in the three stencils of type 2 (Figure 1d, red points) and the subscripts \( 3, • \) denote the eight additional grid points at the corners involved in the four stencils of type 3 (Figure 1d, yellow points). Alternatively, one may use for the consistent mass
\[
\frac{\omega^2}{\kappa} \mathbf{p} \Rightarrow \omega^2 \kappa_0 \left( w_{m0} p_0 + \frac{w_{m1}}{6} \sum_{i=1}^{6} p_{1,i} + \frac{w_{m2}}{12} \sum_{i=1}^{6} p_{2,i} + \frac{w_{m3}}{8} \sum_{i=1}^{8} p_{3,i} \right),
\]
where the bulk modulus at the collocation point is used for all the 27 coefficients. This later formulation preserves the bilinearity property of the wave equation (Aghamiry et al., 2019a,b).
2.2 Numerical dispersion analysis for adaptive mass and stiffness weights

The dispersion analysis in homogeneous media of the 27-point stencil is presented in Operto et al. (2007a). Plugging a plane wave in the discretized wave equation gives the following expression of the numerical phase velocity normalized by the wavespeed $c$.

$$\tilde{v} = \frac{G}{\sqrt{2J\pi}} \sqrt{w_{s1}(3 - C) + \frac{w_{s2}}{3}(6 - C - B) + \frac{2w_{s3}}{4}(3 - 3A + B - C)},$$

(2.12)

where $G$ is the number of points per wavelength $\lambda$, $J = (w_{m0} + 2w_{m1}C + 4w_{m2}B + 8w_{m3}A)$, and the functions $A$, $B$ and $C$ are given by

$$\begin{cases} A(G, \theta, \phi) = \cos\left(\frac{2\pi \cos(\phi) \cos(\theta)}{G}\right) \cos\left(\frac{2\pi \cos(\phi) \sin(\theta)}{G}\right) \cos\left(\frac{2\pi \sin(\phi)}{G}\right), \\ B(G, \theta, \phi) = \cos\left(\frac{2\pi \cos(\phi) \cos(\theta)}{G}\right) \cos\left(\frac{2\pi \cos(\phi) \sin(\theta)}{G}\right) + \cos\left(\frac{2\pi \cos(\phi) \cos(\theta)}{G}\right) \cos\left(\frac{2\pi \sin(\phi)}{G}\right), \\ C(G, \theta, \phi) = \cos\left(\frac{2\pi \cos(\phi) \cos(\theta)}{G}\right) + \cos\left(\frac{2\pi \cos(\phi) \sin(\theta)}{G}\right) + \cos\left(\frac{2\pi \sin(\phi)}{G}\right), \end{cases}$$

(2.13)

where $\phi \in [0\,\pi/4]$ and $\theta \in [0\,\pi/2]$ are the incidence and azimuth angles of the plane wave.

Considering the relationships in equations 2.8 and 2.10, there exist only five independent weights that are to be determined for minimizing the numerical dispersion. These weights satisfy the following linear equation (obtained by requiring $\tilde{v} = 1$, eliminating $w_{s3}$ and $w_{m4}$ and rearranging the terms in equation 2.12):

$$\sum_{l=1}^{5} H_l(G, \theta, \phi) w_l = g(G, \theta, \phi),$$

(2.14)
where \( \mathbf{w} = [w_1, w_2, w_{m0}, w_{m1}, w_{m2}]^T \), \( g(G, \theta, \phi) = \frac{1}{2}((\frac{A^2}{2\pi} + 3)A - B + C - 3) \) and

\[
\begin{align*}
H_1(G, \theta, \phi) &= \frac{1}{2}(3 + 3A - B - C), \\
H_2(G, \theta, \phi) &= \frac{1}{2}(3A - \frac{7}{2}B + \frac{1}{2}C), \\
H_3(G, \theta, \phi) &= \frac{2\pi^2}{g(G)}(A - 1), \\
H_4(G, \theta, \phi) &= \frac{2\pi^2}{g(G)}(6A - 2C), \\
H_5(G, \theta, \phi) &= \frac{2\pi^2}{g(G)}(12A - 4B).
\end{align*}
\]

(2.15)

In order to solve equation 2.14, we define the following misfit function

\[
G = \sum_{i=1}^{N_G} \sum_{j=1}^{N_a} \sum_{k=1}^{N_o} \left( \sum_{l=1}^{5} H_l(ijk) w_l - g(ijk) \right)^2,
\]

(2.16)

where \( g(ijk) \equiv \bullet(G_i, \theta_j, \phi_k) \). This function can be minimized for one value of \( G \) (\( N_G = 1 \)) if an accurate monochromatic simulation should be performed in a homogeneous medium. The phase velocity dispersion curves as a function of \( 1/G \) are shown in Figures 2a, when the weights have been optimized for \( G = 4 \). We also plot phase velocity dispersion surfaces in spherical coordinates (Brossier et al., 2010). Their Figure 3) for different values of \( G \) to better assess the numerical anisotropy of the stencil (Figure 3a). A high accuracy is achieved for \( G = 4 \). However, significant errors are shown for higher values of \( G \) (lower frequencies), with the maximum error reached at \( G \approx 5.5 \) suggesting that this setting is suboptimal for simulation in heterogeneous media.

To improve accuracy in heterogeneous media, Operto et al. (2007a) and Brossier et al. (2010) minimize the equation 2.16 for a range of \( G \) representative of the wavelengths propagated in the subsurface medium during a monochromatic simulation. More precisely, they minimize the misfit function for \( N_G = 4 \) with \( G = 4, 6, 8 \) and 10. The corresponding dispersion curves in Figure 2a show that the phase velocity errors have been mitigated in an average sense, although the minimum error is higher compared to the previous case (Figure 3b).

In both approaches, the 27-point stencil is non-adaptive since the weights don’t depend on the local wavelength. To implement wavelength adaptivity, first we build the weights as functions of \( G \). The set of weight functions

\[
\{w_1(G), w_2(G), w_{m0}(G), w_{m1}(G), w_{m2}(G)\},
\]

(2.17)

can be built by using either a parametric or non-parametric regression via equation 2.16. In a non-parametric setting, equation 2.16 reads

\[
\sum_{i=1}^{N_G} \sum_{j=1}^{N_a} \sum_{k=1}^{N_o} \left( \sum_{l=1}^{5} H_l(ijk) w_l^{(i)} - g(ijk)^{(i)} \right)^2,
\]

(2.18)

where \( \mathbf{w}^{(i)} = [w_1^{(i)}, w_2^{(i)}, w_{m0}^{(i)}, w_{m1}^{(i)}, w_{m2}^{(i)}]^T \).

We now describe in more detail the optimization algorithm that we use to estimate the weights. Equation 2.18 is separable and may be solved for each \( i \) separately as

\[
\text{minimize} \; \left\| \mathbf{H}^{(i)} \mathbf{w}^{(i)} - \mathbf{g}^{(i)} \right\|_2^2,
\]

(2.19)

where

\[
\mathbf{H}^{(i)} = \begin{bmatrix}
H_1^{(i,1)} & H_2^{(i,1)} & H_3^{(i,1)} & H_4^{(i,1)} & H_5^{(i,1)} \\
H_1^{(i,2)} & H_2^{(i,2)} & H_3^{(i,2)} & H_4^{(i,2)} & H_5^{(i,2)} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
H_1^{(i,N_o)} & H_2^{(i,N_o)} & H_3^{(i,N_o)} & H_4^{(i,N_o)} & H_5^{(i,N_o)} \\
H_1^{(i,N_a)} & H_2^{(i,N_a)} & H_3^{(i,N_a)} & H_4^{(i,N_a)} & H_5^{(i,N_a)} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
H_1^{(i,N_o,N_a)} & H_2^{(i,N_o,N_a)} & H_3^{(i,N_o,N_a)} & H_4^{(i,N_o,N_a)} & H_5^{(i,N_o,N_a)}
\end{bmatrix}, \quad \mathbf{g}^{(i)} = \begin{bmatrix}
g^{(i,1)} \\
g^{(i,2)} \\
\vdots \\
g^{(i,N_o)} \\
g^{(i,N_o,N_a)}
\end{bmatrix}.
\]

(2.20)

However, solving these overdetermined systems independently without any prior information about the weights might cause significant oscillatory variations of each weight as a function of \( G \), which may be undesired for accurate wavefield estimation. Therefore, we estimate the \( N_G \) series of weights simultaneously by gathering the \( \mathbf{H}^{(i)} \) matrices in a block diagonal matrix \( \mathbf{H} \) of dimension \((N_o \times N_a \times N_G) \times (5 \times N_G)\) and by augmenting the misfit function with
a Tikhonov smoothing regularizer to enforce smooth variations of each weight as a function of $G$. Accordingly, the regularized optimization problem reads

$$\min_w \|Hw - g\|_2^2 + \lambda\|Lw\|_2^2, \tag{2.21}$$

where

$$H = \begin{bmatrix} H^{(1)} & \cdots & H^{(N_G)} \end{bmatrix}, \quad g = \begin{bmatrix} g^{(1)} \\ \vdots \\ g^{(N_G)} \end{bmatrix}, \quad w = \begin{bmatrix} w^{(1)} \\ \vdots \\ w^{(N_G)} \end{bmatrix},$$

and $L$ is a smoothing operator. Note that the values of not shown coefficients in matrix $H$ are zero. One may build $L$ as the product of two matrices: $L = DP$ where $D$ is a difference operator with appropriate boundary conditions and $P$ is a permutation matrix which reorders the coefficients of $w$ such that the first $N_G$ entries are $[w^{(1)}_1, \ldots, w^{(N_G)}_1]$, the second $N_G$ entries are $[w^{(1)}_2, \ldots, w^{(N_G)}_2]$ and so on. The smooth variation of the estimated adaptive weights as a function of $G$ are shown in Figure 4 while the dispersion curves and surfaces of the adaptive stencil are shown in Figures 2c and 3c. As expected, the numerical dispersion can be theoretically reduced very significantly for a wide range of $G$ by treating each $G$ separately during the dispersion analysis.

The goal of the next section is to check the robustness of this approach with large-scale numerical experiment involving complex structures and sharp material discontinuities.

## 3 Numerical examples

### 3.1 Experimental setup

We validate our method against five benchmarks (Table 1). We limit ourselves to isotropic acoustic media with a constant density equal to 1. However, the conclusions of these tests should apply to visco acoustic media with heterogeneous density. It should also apply to acoustic media with transverse isotropic (TI) effects and viscoelastic media. In the first case, the normalized phase velocity for the quasi-P mode is a function of the vertical wavespeed and the Thomsen’s parameters $\delta$ and $\epsilon$ (Operto et al., 2009). An approximation can also be used to estimate the optimal weights from the local elliptic phase velocity (Operto et al., 2014, Their equation 15) In the elastic case, the normalized $P$ and $S$ phase velocity can be formulated as a function of the Poisson ratio and the shear wavenumber $\kappa$. Thus, we can formulate the optimal weights from the local elliptic phase velocity (Operto et al., 2014, Their equation 15). The last three benchmarks involve complex velocity models corresponding to the 3D SEG/EAGE overthrust model (Figure 5a), the 3D SEG/EAGE salt model (Figure 5b) (Aminzadeh et al., 1997) and a target of the 3D GO_3D_OBS model (Figure 5c) (Görszczyk & Operto, 2021). The two SEG/EAGE models are representative of the exploration geophysics scale. At this scale, $G$ typically ranges between 4 and 12 during a monochromatic simulation. The 3D

![Figure 2: Phase velocity dispersion curves showing the ratio between the numerical phase velocity and the wavespeed versus $1/G$. (a) Weights of the 27-point mixed-grid stencil are estimated to minimize dispersion for $G=4$. (b) Same as (a) for $G=6,8,10$. (c) Adaptive weights are estimated separately for a fine table of $1/G$ ranging between 0 and 0.4 with a step of 0.001. The dip and azimuth angles of the plane waves range between 0 and 45° with a step of 10°.](image-url)
Figure 3: Three-dimensional view of the numerical phase velocity error (in percent). Numerical dispersion and anisotropy are illustrated. (a) The numerical phase velocity errors are computed with the weights minimizing dispersion for $G=4$. From left to right, error for $G=4, 6, 8$, and 10. (b) Same as (a) when the weights are computed to jointly minimize dispersion for $G=4, 6, 8$, and 10. (c) Same as (a) for the adaptive weights.

GO3D_OBS model covers a continental margin at a regional scale. In this case, $G$ typically ranges between 4 and 25 due to the wider range of wavespeeds found in the crust and upper mantle. For these last three models, the FDFD solution is assessed against a reference solution computed with the highly-accurate Convergent Born Series (CBS) method (Osnabrugge et al., 2016). Compared to finite-difference or finite-element methods that rely on discretization of differential operators, the CBS method belongs to another class of methods referred to as volume integral methods based upon the Green’s function theorem. More specifically, the solution of the Helmholtz equation in heterogeneous media is inferred from Fourier-domain analytic expression of the Green’s functions in homogeneous media by solving the scattered-field wave equation with material contrasts of arbitrary scattering strength starting from the zero-order term of the Born series and iteratively recovering the higher-order scattering terms. Therefore, we expect to achieve an accuracy of the CBS solutions at the level of machine precision while Osnabrugge et al. (2016) conclude that the CBS method is nine orders more accurate than the pseudospectral time-domain method. For all the extracted wavefields using CBS, the stopping criterion of iteration is set according to a backward error of $1e-12$. In addition to visual comparison, we assess the accuracy of the adaptive and non-adaptive FDFD stencils by computing the $\ell_1$ norm of the difference between the CBS and FDFD wavefields (Table 2) according to the formula

$$err = \frac{|W R (u_{cbs} - u_{fdfd})|}{|WR(u_{cbs})|} + \frac{|W I (u_{cbs} - u_{fdfd})|}{|WI(u_{cbs})|},$$

(3.1)

where $u_{cbs}$ and $u_{fdfd}$ denotes the CBS and FDFD wavefields, respectively, $R$ and $I$ denote the real and imaginary parts of a complex number, respectively, and $W$ is a diagonal matrix applying a linear gain with distance from the source to the wavefield for amplitude balancing.
For each benchmark, we perform the FDFD simulation with non-adaptive and adaptive weights. In the non-adaptive case, we use two sets of weights that have been obtained by minimizing dispersion for $G = 4$ ($N_G = 1$) and $G = 4, 6, 8, 10$ ($N_G = 4$). In the following, the wavefields computed with such weights are referred to as $G4$ and $Gm$ wavefields, respectively. We test two implementations of the adaptive stencil. The first implementation builds each row of the impedance matrix with the weights corresponding to the local wavelength at the collocation point. The second one builds each row of the impedance matrix with average weights, the averaging being performed over the 27 points of the stencil. In the following, the wavefields computed with these two adaptive approaches are referred to as $GA$ and $GAm$ wavefields, respectively. The source is a point source located on a grid point where the temporal source signature is a delta function, unless otherwise mentioned: $s(x, \omega) = \delta(x - x_s)$, where $\delta$ denotes the delta function and $x_s$ denotes the source coordinates. We solve the discretized Helmholtz equation in single precision with the full-rank version of the massively-parallel sparse direct solver MUMPS (MUMPS team, 2021) based upon the multifrontal method (Amestoy et al., 2001, 2003; Duff & Reid, 1983; Duff et al., 1986). We implement PMLs along each face of the grid. We perform the simulation on the Occigen supercomputer of CINES (https://www.cines.fr). The computer nodes contain two Haswell E5-2690V3@2.6 GHz processors with 128 Giga bytes of shared memory and 12 cores per processor. The high-speed network is Infiniband FDR 56 Gbit/s. For the simulation, we assign one Message Passing Interface (MPI) per node and use multithreading with 24 threads per MPI process to minimize memory overheads during Lower-Upper (LU) decomposition while optimizing computational time.

3.2 Homogeneous velocity model

We first consider a homogeneous medium of size $10 \text{ km} \times 20 \text{ km} \times 10 \text{ km}$ with a wavespeed of 1500 m/s (Table 1). The medium is discretized with a grid interval of 50 m, and the frequency is 7.5 Hz, leading to $G=4$. The number of grid points in the PML is as high as 16 to minimize the footprint of parasite reflections in the accuracy assessment of the stencil. The source is positioned at $(5 \text{ km}, 2 \text{ km}, 5 \text{ km})$. We perform the simulation with 40 nodes of the Occigen
Figure 5: Perspective view of the (a) overthrust, (b) salt, and (c) GO_3D_OBS models.
computer. The elapsed time and the memory required to perform the LU factorization is 837 s and around 2 Tera bytes. The analytical and FDFD wavefields in the (x,y) plane across the source, as well as the difference between the two solutions, are shown in Figure 6. The direct comparison between the two wavefields and the error along the y-profile across the source are shown in Figure 7. As expected, the $G_4$ wavefield is highly accurate since we discretize the homogeneous medium with $G=4$ to perform the simulation. Consistently with the dispersion curves, the accuracy of the $G_A$ and $G_A m$ wavefields are similar to that of the $G_4$ counterpart, while the $G_m$ wavefield is significantly less accurate since the $G_m$ weights result from a compromise to jointly minimize the dispersion for $G=4,6,8$ and 10.

3.3 Linear velocity model

The second benchmark involves a laterally homogeneous medium in which the velocity linearly increases in the y direction. The size of the grid and the frequency are the same as that used for the homogeneous medium. The wavespeed increases from 1500 m/s at y=0 km to 8500 m/s at y=20 km hence covering the wavespeeds encountered in the earth's crust and upper mantle (the velocity gradient is 0.35 s$^{-1}$). Accordingly, $G$ increases from 4 to $\sim$23 (Table 1). The results are shown in Figures 8 and 9 with the same showing as for the homogeneous case. As expected, the $G_4$ wavefield accumulates phase error as $G$ increases in the y direction. The $G_m$ wavefield shows a better accuracy as expected, while the $G_A$ and $G_A m$ wavefields show a higher accuracy without visible phase error and small amplitude errors, hence providing a first validation of the adaptive stencil in smooth media.

3.4 3D SEG/EAGE Overthrust model

We now benchmark the adaptive stencil against more heterogeneous models involving either complex compressive tectonics (overthrust) or sharp contrasts (salt). We start with the overthrust model of dimensions 20 km $\times$ 20 km $\times$ 4.65 km. It represents a complex thrusted sedimentary succession on top of a structurally decoupled extensional and rift basement block (Aminzadeh et al., 1997) (Figure 5a). A complex weathered zone also characterizes it in the near-surface with sharp lateral velocity variations and several sand channels. The velocities range between 2179 m/s at the surface and 6000 m/s at the basement. We re-mesh the model with a grid interval of 50 m and 8 grid points in the PMls along each face of the 3D grid of dimensions 417 $\times$ 417 $\times$ 109. The source is positioned at (2500 m, 2500 m, 500 m). The frequency is 10 Hz, and hence $G$ varies between 4.4 and 12 (Table 1). We perform the simulation with 40 nodes of the Occigen computer. The elapsed time and the memory required to perform the LU factorization are 395 s and 1.2 Tera bytes. The wavefields are shown in Figures 10 and 11. In Figure 10, we show two depth slices of the wavefields at 500 m depth across the source and 3500 m depth just above the basement and two vertical sections at y=2500 m across the source and y=15000 m. In Figure 11, we show a direct comparison between the CBS and the FDFD wavefields along four horizontal profiles at (x,z)=(2000 m, 500 m), (x,z)=(15000 m, 3500 m), (y,z)=(2500 m, 500 m), (y,z)=(15000 m, 3500 m). As expected, the $G_m$ wavefield is more accurate than the $G_4$ counterpart since the values of $G$ covered by the propagated wavelengths are consistent with the values of $G$ for which phase velocity dispersion was minimized during the estimation of the $G_m$ weights (Table 2). The accuracy of the $G_A$ and $G_A m$
Figure 7: Homogeneous model. Direct comparison between analytical (blue) and FDFD (red) solutions (real part) along an horizontal profile across the source oriented in the $y$ direction. The black line is the difference. Amplitudes are corrected for geometrical spreading with a linear gain with offset. From top to bottom, real part of the $G_A$, $G_m$, $G_A$, $G_A m$ wavefields.

Figure 8: Linear velocity model. Top row: Real part of the analytical solution. Middle row from left to right: Real part of the wavefield computed with $G_A$, $G_m$, $G_A$, and $G_A m$ weights. Bottom row: Differences between analytical and FDFD wavefields. The vertical and horizontal axis represents the $x$ and $y$ directions, respectively.
Figure 9: Linear velocity model. Direct comparison between analytical (blue) and FDFD (red) wavefield solutions (real part) along an horizontal profile across the source oriented in the y—direction. The black line is the difference. Amplitudes are corrected for geometrical spreading with a linear gain with offset. From top to bottom, $G^4$, $Gm$, $GA$, $GAm$ wavefields.

wavefields is significantly higher than that of the $Gm$ wavefield, hence providing a first validation of the adaptive stencil in heterogeneous media with the complex tectonic trend. The accuracies of the $GA$ and $GAm$ stencil are almost identical for this benchmark.

3.5 3D SEG/EAGE Salt model

We continue with the 3D SEG/EAGE salt model, which is representative of a typical Gulf coast salt structure (Aminzadeh et al., 1997) (Figure 5b). Here, we want to check the ability of the adaptive stencil to deal with sharp velocity contrasts between the sediments and the salt. The velocities range between 1500 m/s in the water and 4481 m/s in the salt. We re-mesh the velocity model with a grid interval of 40 m and 8 grid points in the PMIs along each face of the 3D grid of dimensions $354 \times 354 \times 131$. The source is positioned at (4000 m, 4000 m, 400 m). The frequency is 9.375 Hz, and hence $G$ varies between 4 and 12 (Table 1). We perform the simulation with 40 nodes of the Occigen computer. The elapsed time and the memory required to perform the LU factorization are 448 s and 1.1 Tera bytes.

The wavefields are shown in Figures 12 and 13. Similar to the overthrust model, the $G^4$ wavefield shows the worst accuracy. The $GA$ and $GAm$ wavefields show an impressive agreement with the CBS wavefield suggesting that the adaptive stencil is very valuable to deal with sharp contrasts. This is highlighted by the fact that the relative accuracy improvement achieved by the adaptive stencils related to the $Gm$ counterpart is much more significant for the salt model compared to the overthrust model (Table 2). Unlike the overthrust model, the $GA$ scheme slightly outperforms the $GAm$ counterpart, further supporting the above statement.

We also check how the stencil behaves at lower frequencies. We re-mesh the model with a grid interval of 80 m and perform the simulation at the 3 Hz frequency leading to $G$ ranging between 6.25 and 18.7. Again the adaptive stencils achieve an impressive accuracy (Figures 14 and 15).
Figure 10: Wavefields computed in the 3D SEG/EAGE Overthrust model. The real part is shown. The figure can be read as a $F_{2 \times 5}$ matrix where the first and second subscripts denote row and column, respectively. From top to bottom, one entry of the matrix (as an example, $F_{1,1}$ is delineated by the thick line in the Figure) shows two depth sections at $z=0.5$ km depth and $z=3.5$ km depth and two vertical sections at $y=2.5$ km and $y=15$ km. ($F_{1,1}$: CBS wavefield. ($F_{1,2}$): $G_4$ wavefield. ($F_{1,4}$): $Gm$ wavefield. ($F_{2,2}$): $GA$ wavefield. ($F_{2,4}$): $GAm$ wavefield. ($F_{1,3}$, $F_{1,5}$, $F_{2,3}$, $F_{2,5}$): Differences between CBS wavefield and $G4$, $Gm$, $GA$, $GAm$ wavefields.
Figure 11: 3D SEG/EAGE Overthrust model. Direct comparison between horizontal profiles across the CBS (blue) and the FDFD (red) wavefields. The difference is plotted with black lines. Amplitudes are corrected from geometrical spreading through a linear gain with offset. In (a-d), from top to bottom, CBS versus $G_4$, $G_{m}$, $G_A$, and $G_{Am}$ wavefields. (a) $(x,z)=(2.5 \text{ km},0.5 \text{ km})$. (b) $(x,z)=(15 \text{ km},3.5 \text{ km})$. (c) (a) $(y,z)=(2.5 \text{ km},0.5 \text{ km})$. (b) $(y,z)=(15 \text{ km},3.5 \text{ km})$
Figure 12: Wavefields computed in the 3D SEG/EAGE Salt model. Same showing as that of Figure 10 is used. The depth sections are at 0.4 km and 2 km depth and two vertical sections at y=4 km and y=8 km.
Figure 13: Same as Figure 11 for the Salt model. Positions of the horizontal profiles are (a) (x,z)=(4 km,0.4 km). (b) (x,z)=(8 km,0.4 km). (c) (y,z)=(4 km,0.4 km). (d) (y,z)=(8 km,2 km).
Figure 14: Wavefields computed in the 3D SEG/EAGE Salt model. Same showing as that of Figure 12 is used but frequency is 3 Hz.
Figure 15: Same as Figure 13 for the 3Hz frequency.
3.6 3D GO,3D_OBS crustal model

The last benchmark corresponds to the 3D GO,3D_OBS crustal model (Górszczyk & Operto, 2021). This geomodel has been designed to assess seismic imaging techniques for regional deep crustal exploration. It embeds the main structural factors that characterize subduction zones and has been inspired by the FWI case study performed in the eastern Nankai trough by Górszczyk et al. (2017). We select a target of the model of dimensions 20 km × 102 km × 28.4 km. We discretize this target with a grid interval of 100 m with eight grid points in the PMLs, leading to a finite-difference grid of dimensions 217 × 1037 × 300, hence 67.5 millions degrees of freedom. We show the result of the simulation for a source positioned on the sea bottom at (10500 m, 12300 m, 900 m). The simulated frequency is 3.75 Hz. The wavespeeds range between 1500 m/s in the water and 8639.1 m/s in the upper mantle, and hence $G$ ranges between 4 and 23 (Table 1). We perform the simulation using 120 nodes of Occigen computer. The elapsed time for factorization and the elapsed time to compute 130 wavefield solutions associated with a sparse layout of ocean bottom nodes is 2615 s and 50 s, respectively. The used peak memory is 9.17 Tera bytes. Note that much improved performances can be obtained in such kinds of large computational domains with the block-low rank version of MUMPS (Amestoy et al., 2016a,b, 2018). This topic will be addressed in the next study. Note also that this simulation could have been performed at the 5 Hz frequency with good accuracy. For this frequency, $G=3$ in the water, which is undersampled according to the FWI resolution. However, the water layer can be kept fixed during FWI since the wavespeed in water is know with good accuracy. Comparisons between the CBS and FDFD wavefields are shown in Figures 16 and 17. In Figure 17, a direct comparison between the CBS and wavefields is performed along with three horizontal profiles in the $y$ direction running across the source position and cross-cutting the accretionary wedge at 6 km depth, the subduction megathrust at 10 km depth and the Moho at 16 km depth. The $G4$ wavefield lacks accuracy since this value of $G$ is met in the water only. The $GA$ and $GAm$ wavefields are significantly more accurate than the $Gm$ counterpart. The $GA$ wavefield is slightly more accurate than the $GAm$ counterpart, further supporting that the adaptive stencil doesn’t require averaging of the weights. Note that compared to the previous tests, we see more significant errors in the bottom-right part of the computational domain for the $Gm$, $GA$, and $DAm$ wavefields. These errors may result from the higher number of propagated wavelengths in the GO,3D_OBS crustal model. They may also result from some inaccuracies of the PMLs at grazing incidences.

3.7 Comparison with high-order finite-difference time-domain method

It is also worth benchmarking the adaptive FDFD method against a high-order staggered-grid velocity-stress finite-difference time-domain method (FDTD) (Graves, 1996; Levander, 1988; Virieux, 1984). In the latter case,
monochromatic solution is computed on the fly in the loop over time steps by discrete Fourier transform (Nihei & Li, 2007; Sirgue et al., 2008). The FDTD stencil is second-order accurate in time and eight-order accurate in space. We use a Ricker wavelet for these simulations instead of a delta signature to prevent numerical artifacts during the FDTD simulations. We first compare the CBS wavefield, the GA wavefield and the monochromatic wavefield inferred from the FDTD simulation for the overthrust model (Figures 18-19 and Table 3). The wavefields computed with the two finite-difference methods match pretty well the CBS wavefield. Using the global metric given by equation 3.1, the FDTD solution is in average more accurate than the FDFD counterpart. However, a close examination of the profiles in Figure 19 shows that the FDFD wavefield is clearly more accurate than the FDTD counterpart in the shallow part, while the FDTD wavefield is more accurate in the deep part. This results because the accuracy of the FDTD and FDFD methods is managed with different paradigms: the error increases monotonically with $1/G$ from 0 at the zero frequency to higher values near the Nyquist frequency in FDTD where the derivative is approximated by a Taylor polynomial of a given degree (Brac, 1997), while the accuracy of the adaptive FDFD is optimized for each value of $G$ through the weight estimation. Accordingly, the error doesn’t necessarily increase monotonically with $1/G$ in FDFD, as illustrated for example in Figure 2(a-b). This explains why the FDFD wavefields are more accurate for small values of $G$, while high-order FDTD is more accurate for high values of $G$ as the error tends to 0.

We repeat this comparison for the salt model. The results are shown in Figures 20 and 21 and in Table 3. Interestingly, the GA wavefield is significantly more accurate than the FDTD counterpart for this benchmark involving sharp contrasts. This highlights how large-support high-order stencils implemented with high-degree Taylor polynomials implemented in FDTD tend to smooth out the effects of the contrasts, unlike the compact-support adaptive FDFD stencil.

## 4 Conclusion

We significantly improve the accuracy of the 27-point finite-difference stencil for 3D frequency-domain (visco-)acoustic seismic wave modeling by adapting the weights through which dispersion is minimized to the local wave-
Figure 18: Wavefields computed in the 3D SEG/EAGE Overthrust model. The real part is shown. The figure can be read as a $F_{1 \times 5}$ matrix where the first and second subscript denote row and column, respectively. From top to bottom, one entry of the matrix shows two depth sections at $z=0.5$ km and $z=3.5$ km and two vertical sections at $x=2.5$ km and $x=15$ km. ($F_{1,1}$) CBS. ($F_{1,2}$): $O(\Delta t^2, \Delta x^8)$ FDTD solution. ($F_{1,4}$) FDFD-Ga. ($F_{1,3}, F_{1,5}$): Differences between CBS and FDTD, FDFD-Ga.

Figure 19: 3D SEG/EAGE Overthrust model. Direct comparison between horizontal profiles across the CBS (blue), FDTD and FDFD (red) wavefield solutions. The difference is plotted with black lines. Amplitudes are corrected from geometrical spreading through a linear gain with offset. In (a-d), from top to bottom, CBS versus FDTD and CBS versus FDFD-Ga. (a) $(x,z)=(2.5$ km,0.5 km). (b) $(x,z)=(15$ km,3.5 km). (c) $(a,y,z)=(2.5$ km,0.5 km). (b) $(y,z)=(15$ km,3.5 km).
Figure 20: Wavefields computed in the 3D SEG/EAGE salt model. The real part is shown. Same as Figure 18 but for the salt model.

Figure 21: 3D SEG/EAGE salt model. Direct comparison between horizontal profiles across the CBS (blue), FDTD and FDFD (red) wavefield solutions. The difference is plotted with black lines. Amplitudes are corrected from geometrical spreading through a linear gain with offset. In (a-d), from top to bottom, CBS versus FDTD and CBS versus FDFD-Ga. (a) (x,z)=(4 km,0.4 km). (b) (x,z)=(8 km,2 km). (c) (y,z)=(4 km,0.4 km). (b) (y,z)=(8 km,2 km).
Table 2: Accuracy measurement of the 27-point FDFD stencil for the different benchmarks. For the homogeneous and linear models, the reference solutions are the analytical ones. For the other models, the reference solution is the CBS one. The errors between the reference wavefield and the $G_4$, $Gm$, $GA$ and $GAm$ wavefields (equation 3.1) are shown in the table.

| Models        | $G_4$  | $Gm$  | $GA$  | $GAm$ |
|---------------|--------|-------|-------|-------|
| Homogeneous   | 0.0242 | 0.8065| 0.0252| 0.0252|
| Linear        | 0.4822 | 0.2285| 0.013 | 0.0135|
| Overthrust    | 0.8619 | 0.3123| 0.129 | 0.1289|
| Salt          | 0.7689 | 0.2582| 0.0298| 0.031 |
| GO_3D_OBS     | 0.4573 | 0.3022| 0.2114| 0.2125|

Table 3: Accuracy of the $GA$ stencil versus the $O(\Delta t^2, \Delta h^8)$ staggered-grid finite-difference time-domain stencil for the overthrust and salt benchmarks. The error given by equation 3.1 is provided in the table.

| Models        | FDTD error | FDFD-GA error |
|---------------|------------|---------------|
| Overthrust    | 0.0876     | 0.1278        |
| Salt          | 0.1977     | 0.0297        |

length. This wavelength-adaptivity doesn’t introduce computational overhead because the shape (position and number of coefficients) of the stencil is unchanged, and the weights are tabulated once and for all for a wide range of the number of grid points per wavelength. Therefore, any impedance matrix associated with a particular simulation is easily built by picking the appropriate weights in the table for each of its rows. We benchmark the method with different subsurface models of increasing complexity. In all cases, the adaptive stencil outperforms the non-adaptive stencil in terms of accuracy. The accuracy improvement is more significant when the subsurface model contains sharp contrasts. The compact adaptive stencil also shows improved accuracy compared to the high-order finite-difference time-domain method in the presence of sharp contrasts. The method should also apply to acoustic media with transverse isotropic effects and elastic media.

Acknowledgments

This study was partially funded by the WIND consortium (https://www.geoazur.fr/WIND), sponsored by Chevron, Shell and Total. The authors are grateful to the OPAL infrastructure from Observatoire de la Côte d’Azur (CRIMSON) for providing resources and support. This work was granted access to the HPC resources of IDRIS under the allocation A5005410596 made by GENCI. We thank the MUMPS developers, P. Amestoy (Mumps Technologies - ENS Lyon), A. Buttari (CNRS-IRIT), J-Y. L’Excellent (Mumps Technologies - ENS Lyon), T. Mary (LPI-CNRS), C. Puglisi (Mumps Technologies - ENS Lyon), for providing the MUMPS solver (http://mumps-solver.org) and assistance for optimal tuning.

References

Aghamiry, H., Gholami, A., & Operto, S., 2019a. ADMM-based multi-parameter wavefield reconstruction inversion in VTI acoustic media with TV regularization, Geophysical Journal International, 219(2), 1316–1333.

Aghamiry, H., Gholami, A., & Operto, S., 2019b. Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers, Geophysics, 84(1), R139–R162.

Amestoy, P., Brossier, R., Buttari, A., L’Excellent, J.-Y., Mary, T., Métiévier, L., Miniussi, A., & Operto, S., 2016a. Fast 3D frequency-domain FWI with a parallel Block Low-Rank multifrontal direct solver: application to OBC data from the North Sea, Geophysics, 81(6), R363 – R383.

Amestoy, P., Buttari, A., L’Excellent, J.-Y., & Mary, T., 2016b. Complexity and performance of the Block Low-Rank multifrontal factorization, in SIAM Conference on Parallel Processing (SIAM PP16), Paris, France.

Amestoy, P., Buttari, A., L’Excellent, J.-Y., & Mary, T., 2018. On the complexity of the block low-rank multifrontal factorization, SIAM Journal on Scientific Computing, 49(4), A1710–A1740.

Amestoy, P. R., Duff, I. S., Koster, J., & L’Excellent, J. Y., 2001. A fully asynchronous multifrontal solver using distributed dynamic scheduling, SIAM Journal of Matrix Analysis and Applications, 23(1), 15–41.
Amestoy, P. R., Duff, I. S., Pralet, S., & Vomel, C., 2003. Adapting a parallel sparse direct solver to architectures with clusters of SMPs, *Parallel computing*, **29**, 1645–1668.

Amestoy, P. R., Ashcraft, C., Boiteau, O., Buttari, A., L’Excellent, J.-Y., & Weisbecker, C., 2015. Improving multifrontal methods by means of block low-rank representations, *SIAM Journal on Scientific Computing*, **37**(3), 1451–1474.

Aminzadeh, F., Brac, J., & Kunz, T., 1997. *3-D Salt and Overthrust models*, SEG/EAGE 3-D Modeling Series No.1.

Brac, J., 1997. Simulation de la propagation des ondes acoustiques 2D et 3D, élastiques 2D. Etude des schémas d’ordre élevé. Méthode pseudo-spectrale tronquée. Logiciel C2D, C3D, E2D, Tech. rep., Institut Francais du Pétrole (IFP), Rapport 44 220.

Brossier, R., Etienne, V., Operto, S., & Virieux, J., 2010. Frequency-domain numerical modelling of visco-acoustic waves based on finite-difference and finite-element discontinuous galerkin methods, in *Acoustic Waves*, pp. 125–158, ed. Dissanayake, D. W., SCIYO.

Chen, Z., Cheng, D., & Wu, T., 2012. A dispersion minimizing finite difference scheme and preconditioned solver for the 3D Helmholtz equation, *Journal of Computational Physics*, **231**, 8152–8175.

Chen, Z., Cheng, D., Feng, W., & Wu, T., 2013. An optimal 9-point finite difference scheme for the Helmholtz equation with PML, *International Journal of Numerical Analysis & Modeling*, **10**(2).

Dolean, V., Jolivet, P., & Nataf, F., 2015. *An introduction to domain decomposition methods - Algorithms, theory, and parallel implementation*, SIAM.

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

Duff, I. S., Erisman, A. M., & Reid, J. K., 1986. *Direct methods for sparse matrices*, second edition, Oxford Science Publications, Oxford, U. K.

Duff, I. S., Erisman, A. M., & Reid, J. K., 2017. *Direct methods for sparse matrices*, Clarendon Press, Oxford, U. K., second edition.

Ernst, O. G. & Gander, M. J., 2012. Why it is Difficult to Solve Helmholtz Problems with Classical Iterative Methods, in *Numerical Analysis of Multiscale Problems*, vol. 83 of *Lecture Notes in Computational Science and Engineering*, pp. 325–363, eds Graham, I. G., Hou, T. Y., Lakkis, O., & Scheichl, R., Springer Berlin Heidelberg.

Górszczyk, A. & Operto, S., 2021. GO3D_OBS: the multi-parameter benchmark geomodel for seismic imaging method assessment and next-generation 3d survey design (version 1.0), *Geoscientific Model Development*, **14**, 1773–1799.

Górszczyk, A., Operto, S., & Malinowski, M., 2017. Toward a robust workflow for deep crustal imaging by FWI of OBS data: The eastern nankai trough revisited, *Journal of Geophysical Research: Solid Earth*, **122**(6), 4601–4630.

Gosselin-Cliche, B. & Giroux, B., 2014. 3D frequency-domain finite-difference viscoelastic-wave modeling using weighted average 27-point operators with optimal coefficients, *Geophysics*, **79**(3), T169–T188.

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.

Hicks, G. J., 2002. Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions, *Geophysics*, **67**, 156–166.

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

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

Kostin, V., Solovyev, S., Bakulin, A., & Dmitriev, M., 2019. Direct frequency-domain 3D acoustic solver with intermediate data compression benchmarked against time-domain modeling for full-waveform inversion applications, *Geophysics*, **84**(4), T207–T219.

Kuvshinov, B. N. & Mulder, W. A., 2006. The exact solution of the time-harmonic wave equation for a linear profile, *Geophysical Journal International*, **167**, 659–662.

Levander, A. R., 1988. Fourth-order finite-difference P-SV seismograms, *Geophysics*, **53**(11), 1425–1436.

Min, D. J., Sin, C., Kwon, B.-D., & Chung, S., 2000. Improved frequency-domain elastic wave modeling using weighted-averaging difference operators, *Geophysics*, **65**, 884–895.
Nihei, K. T. & Li, X., 2007. Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD), *Geophysical Journal International*, **169**, 1069–1078.

Operto, S. & Miniussi, A., 2018. On the role of density and attenuation in 3D multi-parameter visco-acoustic VTI frequency-domain FWI: an OBC case study from the North Sea, *Geophysical Journal International*, **213**, 2037–2059.

Operto, S., Virieux, J., Amestoy, P., L’Élèctent, J.-Y., Giraud, L., & Ben Hadj Ali, H., 2007a. 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study, *Geophysics*, **72**(5), SM195–SM211.

Operto, S., Virieux, J., Amestoy, P., L’Élèctent, J.-Y., Giraud, L., & Ben Hadj Ali, H., 2007b. 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study, *Geophysics*, **72**(5), SM195–SM211.

Operto, S., Virieux, J., Ribodetti, A., & Anderson, J. E., 2009. Finite-difference frequency-domain modeling of visco-acoustic wave propagation in two-dimensional TTI media, *Geophysics*, **74**(5), T75–T95.

Operto, S., Brossier, R., Combe, L., Métivier, L., Ribodetti, A., & Virieux, J., 2014. Computationally-efficient three-dimensional visco-acoustic finite-difference frequency-domain seismic modeling in vertical transversely isotropic media with sparse direct solver, *Geophysics*, **79**(5), T257–T275.

Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., Ribodetti, A., & Virieux, J., 2015. Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation, *Geophysical Journal International*, **202**(2), 1362–1391.

Osnabruggge, G., Leedumrongwatthanakun, S., & Vellekoop, I. M., 2016. A convergent born series for solving the inhomogeneous helmholtz equation in arbitrarily large media, *Journal of computational physics*, **322**, 113–124.

Plessix, R. E., 2006. A robust three-dimensional iterative solver for the time-harmonic wave equation, in *Extended Abstracts*, Eur. Ass. Expl.Geophys.

Plessix, R. E., 2007. A Helmholtz iterative solver for 3D seismic-imaging problems, *Geophysics*, **72**(5), SM185–SM194.

Plessix, R.-E., 2017. Some computational aspects of the time and frequency domain formulations of seismic waveform inversion, pp. 159–187, Springer.

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

Pratt, R. G. & Worthington, M. H., 1990. Inverse theory applied to multi-source cross-hole tomography. Part I: acoustic wave-equation method, *Geophysical Prospecting*, **38**, 287–310.

Saad, Y., 1986. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM Journal on Scientific and Statistical Computing*, **7**(3), 856–869.

Saad, Y., 2003. *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia.

Shin, C. & Sohn, H., 1998. A frequency-space 2-D scalar wave extrapolator using extended 25-point finite-difference operator, *Geophysics*, **63**, 289–296.

Sirgue, L. & Pratt, R. G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics*, **69**(1), 231–248.

Sirgue, L., Etgen, J. T., & Albertin, U., 2008. 3D frequency domain waveform inversion using time domain finite difference methods, in *Proceedings 70th EAGE, Conference and Exhibition*, Roma, Italy, p. F022.

Stekl, I. & Pratt, R. G., 1998. Accurate viscoelastic modeling by frequency-domain finite difference using rotated operators, *Geophysics*, **63**, 1779–1794.

Toksöz, M. N. & Johnston, D. H., 1981. *Geophysics reprint series, No. 2: Seismic wave attenuation*, Society of exploration geophysicists, Tulsa, OK.

Tournier, P. H., Aliferis, I., Bonazzoli, M., Buhan, M. D., Darbas, M., Dolean, V., Hecht, F., Jolivet, P., Kanfoud, I. E., Miglaccio, C., Nataf, F., Pichot, C., & Semenoce, S., 2019. Microwave Tomographic Imaging of Cerebrovascular Accidents by Using High-Performance Computing, *Parallel computing*, **85**, 88–97.
Turkel, E., Gordon, D., Gordon, R., & Tsynkov, S., 2013. Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number, *Journal of Computational Physics*, **232**, 272–287.

Vigh, D. & Starr, E. W., 2008. Comparisons for Waveform Inversion, Time domain or Frequency domain?, in *Extended Abstracts*, pp. 1890–1894.

Virieux, J., 1984. SH wave propagation in heterogeneous media: Velocity-stress finite difference method, *Geophysics*, **49**, 1259–1266.

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

Wang, S., de Hoop, M. V., & Xia, J., 2011. On 3D modeling of seismic wave propagation via a structured parallel multifrontal direct Helmholtz solver, *Geophysical Prospecting*, **59**(5), 857–873.

Xu, W. & Gao, J., 2018. Adaptive 9-point frequency-domain finite-difference scheme for wavefield modeling of 2D acoustic wave equation, *Journal of Geophysics and Engineering*, **15**, 1432–1445.