



ELSEVIER

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp



Optimal fourth-order staggered-grid finite-difference scheme for 3D frequency-domain viscoelastic wave modeling

Y. Li^a, B. Han^{a,*}, L. Métivier^{b,c}, R. Brossier^c^a Department of Mathematics, Harbin Institute of Technology, 150001, Harbin, China^b LJK, CNRS, Université Grenoble Alpes, F-38000 Grenoble, France^c ISTerre, Université Grenoble Alpes, F-38000 Grenoble, France

ARTICLE INFO

Article history:

Received 5 November 2015

Received in revised form 24 May 2016

Accepted 12 June 2016

Available online 17 June 2016

Keywords:

3D viscoelastic

Seismic wave modeling

Finite difference

Staggered grid

Anti-lumped mass

CARP-CG

ABSTRACT

We investigate an optimal fourth-order staggered-grid finite-difference scheme for 3D frequency-domain viscoelastic wave modeling. An anti-lumped mass strategy is incorporated to minimize the numerical dispersion. The optimal finite-difference coefficients and the mass weighting coefficients are obtained by minimizing the misfit between the normalized phase velocities and the unity. An iterative damped least-squares method, the Levenberg–Marquardt algorithm, is utilized for the optimization. Dispersion analysis shows that the optimal fourth-order scheme presents less grid dispersion and anisotropy than the conventional fourth-order scheme with respect to different Poisson's ratios. Moreover, only 3.7 grid-points per minimum shear wavelength are required to keep the error of the group velocities below 1%. The memory cost is then greatly reduced due to a coarser sampling. A parallel iterative method named CARP-CG is used to solve the large ill-conditioned linear system for the frequency-domain modeling. Validations are conducted with respect to both the analytic viscoacoustic and viscoelastic solutions. Compared with the conventional fourth-order scheme, the optimal scheme generates wavefields having smaller error under the same discretization setups. Profiles of the wavefields are presented to confirm better agreement between the optimal results and the analytic solutions.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Frequency-domain seismic wave modeling has been extensively investigated and applied within the framework of seismic imaging techniques, such as the full waveform inversion (FWI) [62] and the reverse time migration (RTM) [46]. Due to the evolution of techniques and the practical requirements, dealing with the anelastic properties of the real materials becomes more and more important. Therefore, developing efficient numerical schemes for the wave modeling in viscoelastic heterogeneous media is necessary. Time-domain viscoelastic modeling has been investigated using the finite-difference (FD) scheme by Carcione [10], Robertsson et al. [53], Bohlen [6], Kristek and Moczo [29], etc. The counterpart in the frequency-domain is carried out by Štekl and Pratt [63], Min et al. [35], Gosselin-Cliche and Giroux [21]. Viscoelastic wave modeling in the frequency-domain, which is equivalent to solving a large sparse linear system after discretizing the wave equation, presents several advantages over the time-domain approaches [21]: (a) the modeling is not limited by the Courant–Friedrichs–Lewy (CFL) stability condition [13]; (b) the seismic attenuation could be easily incorporated by using

* Corresponding author.

E-mail address: bohan@hit.edu.cn (B. Han).

complex-valued wave velocities [27]; (c) the full waveform inversion can be efficiently performed by limiting the inversion to a few discrete frequencies [48]; (d) conducting the multi-source modeling is inexpensive if a direct method is utilized for solving the resulting linear system, where the matrix factorization is required only once, while multiple right-hand side techniques are also available for iterative solvers [40,55,11]; (e) the modeling at different frequencies can be performed simultaneously with a parallel implementation.

As a consequence of the simplicity for incorporating seismic attenuation in the frequency domain, the viscosity of the subsurface medium can be represented more accurately. Therefore, the numerical results can provide data which fit seismic records better. Additionally, accounting for the viscosity within time-domain FWI strategies yields difficulties for the gradient computation. Efficient time-reversal techniques become unstable [12,8], and state of art techniques thus require checkpointing strategies to be considered, which may require significant amount of data to be stored in core or on disk [23,56,2]. This complexity is avoided in the frequency-domain formalism.

Among different discretization approaches, FD scheme has been gaining popularity due to its simplicity and computational efficiency. An early application was using a second-order spatial approximations for the second-order stress formulation of the acoustic wave equation, where 10 grid-points per shortest wavelength is required to ensure accurate modeling [49]. Jo et al. [25] reduced this number to 4 through using a mixed-grid stencil and an anti-lumped mass (ALM) strategy. The ALM strategy, commonly used in the finite-element method [34], considers the mass acceleration term as a linear combination of the lumped mass matrix and the consistent mass matrix, i.e., the acceleration term is distributed over the grid-points related in the stencil through certain averaging weights. Hustedt et al. [24] derived a similar scheme from the first-order stress-velocity system on a staggered-grid, and by applying a parsimonious approach [33]. Extensions of the mixed-grid scheme to 3D acoustic modeling was developed by Operto et al. [43] where a 27-point stencil is derived. The 2D time-domain elastic staggered-grid FD schemes for the first-order stress-velocity system was first studied using second-order spatial approximations [60]. A fourth-order version was later provided by Levander [30]. A second-order FD scheme for the second-order velocity-based system was used in the forward modelings for the waveform inversion [47]. Similarly, a mixed-grid approach was subsequently developed for the second-order system to reduce the required number of grid-points per shortest wavelength [63]. Extension to the 3D viscoelastic mixed-grid scheme based on the second-order system was derived in [21].

To summarize, the FD discretization for (visco)elastic wave modeling can be divided into two categories: the first is to discretize the second-order velocity-based system on the conventional grid where all the particle velocity components are located at the same position; the second is to discretize the first-order velocity-stress formulation of the (visco)elastic system on a staggered grid where the locations of the stresses and particle velocities are staggered. To decrease the numerical dispersion, one could either combine the reference difference operator and the rotated difference operator to derive the mixed-grid scheme, such as the one proposed by Štekl and Pratt [63], or apply the ALM strategy to distribute the acceleration term [25,35]. For the first kind discretization combined with the mixed-grid scheme, the numerical dispersion is highly dependent on the Poisson's ratio, despite the gain in decreasing the dispersion [63]. However, it is not the case for the staggered-grid scheme [30]. The staggered-grid scheme has several advantages over the schemes based on the conventional grid: (a) the staggered-grid scheme is stable for all values of Poisson's ratio, which means the fluid–solid interfaces could be properly handled [30]; (b) the grid-dispersion and grid-anisotropy are relatively insensitive to Poisson's ratio [30]; and (c) free surface boundary conditions can be directly implemented by manipulating the related stress components [30,54]. On the other hand, one may try to decrease the numerical dispersion of the staggered-grid scheme by using a mixed grid, whereas this is not straightforward due to the mismatch of the locations of different physical variables. Yet the ALM strategy stays as an option.

To the best of our knowledge, there are not any works on using the optimized staggered-grid scheme and the ALM strategy for 3D viscoelastic wave modeling. Therefore, in this paper we combine these two approaches and derive the optimal values of the FD coefficients and the mass weighting coefficients to minimize the numerical dispersion. Note that proper treatment of irregular topography and the free surface boundary condition are important issues for elastic wave modeling. Finite element method, discontinuous Galerkin method (DGM) [7] and spectral element method [28] could be alternatives to the FD method, but the computational cost is higher due to the complexity related to the mass matrix formation [61]. A hybrid scheme with a shallow boundary layer discretized by DGM and the deeper part by FD method may be a good option [32]. Specific techniques, such as the image method [30,54], vacuum formulation [22,36] and the immersed boundary method [17] could be applied with the FD scheme to deal with the irregular topography and the free surface boundary condition. In this research, we focus on the derivation of the optimal coefficients of the FD scheme. The modeling is restricted to flat topography and the free surface boundary condition is currently not taken into account.

The discretization of the 3D viscoelastic wave equations generates a linear system $\mathbf{Ax} = \mathbf{b}$ where A is a large, sparse and ill-conditioned matrix. Generally the linear solvers for such systems are divided into two categories: the direct methods and the iterative methods. Direct methods have received an extensive popularity in the seismic imaging community because of their high efficiency in dealing with multiple-source problems [9,43,41]. The key component of direct methods, matrix factorization, is required only once. The solution for each right-hand side can be obtained by simple computational tasks (forward and backward substitution for the LU decomposition for instance). The direct solver MUMPS [38] has been successfully used in 3D acoustic FWI [42] by means of block low-rank representations [1] which could reduce the memory demand to some extent. A successful application of MUMPS on the frequency-domain viscoelastic modeling in a $105 \times 67 \times 67$ model is presented in [21]. However, the significant in-core memory requirement, by and large, still hinders their application for

large scale problems, especially for the 3D frequency-domain viscoelastic wave modeling that we are dealing with in this research.

Alternatively, we may adopt iterative methods to solve the linear system. Iterative solvers fully benefit from the sparsity of the impedance matrix [45,52,15]. The major computational cost of iterative solvers is related to matrix–vector products. Performing these products is memory efficient and computationally cheap since sparse formats can be used to store the matrices. Moreover, the parallelization is more straightforward for iterative solvers. Nevertheless, the application of iterative solvers is constrained by the uncertainty of the convergence due to the indefiniteness and ill-conditioning of the impedance matrix. Quite recently, a robust parallel iterative solver named CARP-CG [20] has been successfully applied to the 2D and 3D frequency-domain acoustic [59,57] and elastic [31] wave modeling in complex media. The CARP-CG method is a parallelization of the CGMN (Conjugate Gradient Minimum Norm) method, which can be considered as a conjugate gradient acceleration of the symmetric Kaczmarz row-projection method [26,4,19]. The convergence of the Kaczmarz method is guaranteed for a linear system, given the existence of the solution [18]. Performing the row projections in a symmetric sense enables the incorporation of the conjugate gradient method, which accelerates the convergence rate. The parallelization is realized through a domain decomposition approach based on row-block division. The scaling property of parallelization has been shown satisfactory [31]. Due to its robustness and efficiency, the CARP-CG method is used to perform the 3D frequency-domain viscoelastic wave modeling in the following sections.

In this paper, the fourth-order staggered-grid FD scheme is applied to the 3D frequency-domain stress-velocity formulation of the viscoelastic system. Firstly, the expressions of the normalized phase velocities are derived. Then we present the FD operators that are used to discretize the first-order viscoelastic system. The mass operator corresponding to the ALM strategy is incorporated to minimize the numerical dispersion. The formulation of the FD operators and the mass averaging operators are then explicitly derived by substituting a plane-wave solution into the discretized system. The optimal FD coefficients and ALM weighting coefficients are determined through an optimization process with respect to specific ranges of Poisson’s ratio, propagation angles and the number of grid-points per minimum shear wavelength. The optimization problem is solved by a damped least-squares method, the Levenberg–Marquardt method [37,51]. Dispersion analysis are conducted to compare the conventional fourth-order scheme and the optimal fourth-order scheme. In the numerical experiments, we validate the results of the optimal and conventional scheme with the analytic viscoacoustic and viscoelastic solutions. Profiles of the numerical and analytic results are presented to check their consistency. Quantitative comparisons through calculating the relative errors between the numerical results and analytic solutions are also performed to verify the superiority of the optimal scheme over the conventional scheme.

2. Theory

The optimal set of FD coefficients and the ALM weighting coefficients, which minimize the numerical dispersion and anisotropy, is given in this section. We first obtain the discrete dispersion equations from the 3D source-free frequency-domain system in a homogeneous and isotropic viscoelastic medium. The normalized phase velocities are obtained subsequently by calculating the ratio between the discrete phase velocities and the analytic phase velocities for the *P*- and *S*-waves. Then the misfit between the normalized phase velocities and the unity is formulated as a function of the Poisson’s ratio, the propagation angles and the number of wavelength per minimum shear wavelength. The minimization of this function is performed afterwards through a damped least-squares algorithm, where the final optimal coefficients are obtained.

We follow the framework given in [21] to develop the dispersion relations and the normalized phase and group velocities. Different from the 27-point scheme they developed directly on the 3D second-order velocity system, the discretization is performed on the first-order stress-velocity formulation of the 3D viscoelastic equation and the underlying FD stencil is based on a staggered grid. The specific FD scheme is elaborated in the following sections.

2.1. Normalized phase velocities

We start with the matrix form of the 3D frequency-domain viscoelastic equation in a homogeneous and isotropic medium,

$$\begin{bmatrix} \omega^2 \rho + v \partial_{xx}^2 + \mu(\partial_{yy}^2 + \partial_{zz}^2) & (v - \mu) \partial_{xy}^2 & (v - \mu) \partial_{zx}^2 \\ (v - \mu) \partial_{xy}^2 & \omega^2 \rho + v \partial_{yy}^2 + \mu(\partial_{zz}^2 + \partial_{xx}^2) & (v - \mu) \partial_{yz}^2 \\ (v - \mu) \partial_{zx}^2 & (v - \mu) \partial_{yz}^2 & \omega^2 \rho + v \partial_{zz}^2 + \mu(\partial_{xx}^2 + \partial_{yy}^2) \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad (1)$$

where ω is the angular frequency, ρ is the density, $\mathbf{v} = (v_x, v_y, v_z)$ is the particle velocity vector. The *P*-wave modulus v is defined as $v = \lambda + 2\mu$ where λ and μ are the Lamé parameters. Note that the viscoelastic attenuation can be incorporated easily by using complex-valued Lamé parameters generated from complex-valued wave velocities [27,16].

A vector plane-wave $(v_x, v_y, v_z) = (v_{x0}, v_{y0}, v_{z0})e^{-i(k_x x + k_y y + k_z z)}$ is substituted into equation (1), where v_{x0} , v_{y0} and v_{z0} are the amplitudes at the origin, k_x , k_y and k_z are components of the wavenumber $k = |(k_x, k_y, k_z)|$. Setting the determinant of coefficient matrix equal to zero generates a cubic polynomial of ω^2 , which has three solutions:

$$\omega_p^2 = \frac{k^2 \nu}{\rho}, \quad \omega_s^2 = \frac{k^2 \mu}{\rho} \quad (\text{degenerate}). \tag{2}$$

This is the well-known analytic dispersion relations [14]. Thus the analytic phase velocities of P- and S-waves can be expressed as

$$V_p = \frac{\omega_p}{k} = \sqrt{\frac{\nu}{\rho}}, \quad V_s = \frac{\omega_s}{k} = \sqrt{\frac{\mu}{\rho}}. \tag{3}$$

However, the analytic dispersion relations (2) no longer hold if the continuous system (1) is discretized through certain FD scheme. Assuming that the domain is uniformly discretized by a spatial step $h = \Delta x = \Delta y = \Delta z$, the matrix in system (1) hence takes the form [21]

$$\begin{bmatrix} \omega^2 \rho D_m + \nu \frac{D_{xx}}{h^2} + \mu \frac{(D_{yy} + D_{zz})}{h^2} & (v - \mu) \frac{D_{xy}}{h^2} & (v - \mu) \frac{D_{zx}}{h^2} \\ (v - \mu) \frac{D_{xy}}{h^2} & \omega^2 \rho D_m + \nu \frac{D_{yy}}{h^2} + \mu \frac{(D_{zz} + D_{xx})}{h^2} & (v - \mu) \frac{D_{yz}}{h^2} \\ (v - \mu) \frac{D_{zx}}{h^2} & (v - \mu) \frac{D_{yz}}{h^2} & \omega^2 \rho D_m + \nu \frac{D_{zz}}{h^2} + \mu \frac{(D_{xx} + D_{yy})}{h^2} \end{bmatrix}, \tag{4}$$

where D_{ij} ($i, j = x, y$ or z) are the FD operators. As mentioned in [21], the operator D_{ij} can take any form. It can be either a central, forward or backward, and second or higher order operator. We emphasize in this paper the fourth-order staggered-grid FD operator. The operator D_m is used to distribute the mass acceleration terms over the grid-points involved in the FD scheme. The specific illustration is given in the following sections.

Similarly, making the determinant of matrix (4) equal to zero generates a cubic polynomial on ω^2 , three solutions of which correspond to dispersion relations of the discrete case. Unlike the analytic case where the solution of the S-wave is degenerate, the discretization leads to two conjugate solutions of the polynomial corresponding to two different S-wave phase velocities S_1 and S_2 . Hence the discrete dispersion relations are

$$\begin{aligned} \omega_p &= \frac{1}{h} \sqrt{\frac{1}{D_m \rho} \left(-\frac{P_0}{3} + P_4 + P_5 \right)}, \\ \omega_{s1} &= \frac{1}{h} \sqrt{\frac{1}{D_m \rho} \left(-\frac{P_0}{3} - \frac{P_4 + P_5}{2} + i\sqrt{3} \left(\frac{P_4 - P_5}{2} \right) \right)}, \\ \omega_{s2} &= \frac{1}{h} \sqrt{\frac{1}{D_m \rho} \left(-\frac{P_0}{3} - \frac{P_4 + P_5}{2} - i\sqrt{3} \left(\frac{P_4 - P_5}{2} \right) \right)}, \end{aligned} \tag{5}$$

where

$$\begin{aligned} P_4 &= \left(P_2 + \sqrt{P_2^2 + P_1^3} \right)^{1/3}, \\ P_5 &= \left(P_2 - \sqrt{P_2^2 + P_1^3} \right)^{1/3}, \end{aligned} \tag{6}$$

where P_0, P_1 and P_2 are functions of the FD operators D_{ij} , the elastic moduli μ and ν and the wavenumber components k_x, k_y and k_z . The explicit formulas are given in Appendix A. One may also refer to [21] for more details.

In order to obtain the normalized phase velocities, we define G as the number of grid-points per minimum wavelength and denote the inverse of G as K (i.e., wavenumber in grid-point units), which is given by

$$K = \frac{kh}{2\pi}. \tag{7}$$

Then the discrete phase velocities can be derived through dividing (5) by the wavenumber, and the normalized phase velocities are obtained by calculating the ratio between the discrete and the analytic phase velocities. Rearranging equations (3), (5) and (7) give us

$$\begin{aligned} \frac{V_p^{ph}}{V_p} &= \frac{\omega_p/k}{\sqrt{\nu/\rho}} = \frac{1}{2\pi K} \sqrt{\frac{-\frac{P_0}{3} + P_4 + P_5}{D_m \nu}}, \\ \frac{V_{s1}^{ph}}{V_s} &= \frac{\omega_{s1}/k}{\sqrt{\mu/\rho}} = \frac{1}{2\pi K} \sqrt{\frac{-\frac{P_0}{3} - \frac{P_4 + P_5}{2} + i\sqrt{3} \left(\frac{P_4 - P_5}{2} \right)}{D_m \mu}}, \\ \frac{V_{s2}^{ph}}{V_s} &= \frac{\omega_{s2}/k}{\sqrt{\mu/\rho}} = \frac{1}{2\pi K} \sqrt{\frac{-\frac{P_0}{3} - \frac{P_4 + P_5}{2} - i\sqrt{3} \left(\frac{P_4 - P_5}{2} \right)}{D_m \mu}}. \end{aligned} \tag{8}$$

As performed in [21], substituting the Poisson’s ratio σ for the moduli ν and μ in equation (8) reduces the number of variables in equation (8). Through the relation

$$\frac{\nu}{\mu} = \frac{\lambda + 2\mu}{\mu} = \frac{1 - \sigma}{0.5 - \sigma}, \tag{9}$$

equations in (8) take the form

$$\begin{aligned} \frac{V_p^{\text{ph}}}{V_p} &= \frac{1}{2\pi K} \sqrt{\frac{-\frac{P_0}{3} + P_4 + P_5}{D_m(1 - \sigma)}}, \\ \frac{V_{s1}^{\text{ph}}}{V_s} &= \frac{1}{2\pi K} \sqrt{\frac{-\frac{P_0}{3} - \frac{P_4+P_5}{2} + i\sqrt{3}\left(\frac{P_4-P_5}{2}\right)}{D_m(0.5 - \sigma)}}, \\ \frac{V_{s2}^{\text{ph}}}{V_s} &= \frac{1}{2\pi K} \sqrt{\frac{-\frac{P_0}{3} - \frac{P_4+P_5}{2} - i\sqrt{3}\left(\frac{P_4-P_5}{2}\right)}{D_m(0.5 - \sigma)}}, \end{aligned} \tag{10}$$

where the change of variable $(\mu, \nu) \rightarrow \sigma$ in P_0, P_1 and P_2 should also be considered.

Note that in equations (2) and (5), we have the analytic and discrete dispersion relations with the wavenumber k . The angular frequencies ω_p, ω_s (ω_{s1}, ω_{s2} for the discrete case) can be considered as functions of k . However, in the numerical modeling where the angular frequency is usually chosen as a fixed value ω , the wavenumber k should be adjusted as k_p or k_s respectively. Hence the number of grid-points per minimum shear wavelength G_s (and its inverse K_s) can be defined consequently with respect to k_s . Therefore, in the following sections for obtaining the optimal coefficients and for the numerical wave modelings, the angular frequency is fixed as a constant value and k_p, k_s, G_s or K_s are considered appropriately.

Given the normalized phase velocities of P- and S-wave in equation (10), we need to define the FD operators D_{ij} to express the P_* terms therein, and to determine the normalized phase velocities numerically. The operators defined in the fourth-order staggered-grid FD scheme are given in the next section.

2.2. The finite-difference operators

Due to the utilization of the staggered-grid scheme, we consider the velocity-stress formulation of 3D frequency-domain elastodynamic system in heterogeneous isotropic media,

$$\begin{cases} i\omega \rho v_x = \partial_x \tau_{xx} + \partial_y \tau_{xy} + \partial_z \tau_{xz} + f_x, \\ i\omega \rho v_y = \partial_x \tau_{xy} + \partial_y \tau_{yy} + \partial_z \tau_{yz} + f_y, \\ i\omega \rho v_z = \partial_x \tau_{xz} + \partial_y \tau_{yz} + \partial_z \tau_{zz} + f_z, \\ i\omega \tau_{xx} = (\lambda + 2\mu)\partial_x v_x + \lambda\partial_y v_y + \lambda\partial_z v_z + i\omega \tau_1, \\ i\omega \tau_{yy} = \lambda\partial_x v_x + (\lambda + 2\mu)\partial_y v_y + \lambda\partial_z v_z + i\omega \tau_2, \\ i\omega \tau_{zz} = \lambda\partial_x v_x + \lambda\partial_y v_y + (\lambda + 2\mu)\partial_z v_z + i\omega \tau_3, \\ i\omega \tau_{yz} = \mu(\partial_y v_z + \partial_z v_y), \\ i\omega \tau_{zx} = \mu(\partial_x v_z + \partial_z v_x), \\ i\omega \tau_{xy} = \mu(\partial_x v_y + \partial_y v_x), \end{cases} \tag{11}$$

where i is the imaginary unit, $\mathbf{v} = (v_x, v_y, v_z)$ is the particle velocity vector, $\boldsymbol{\tau} = (\tau_{xx}, \tau_{yy}, \tau_{zz}, \tau_{yz}, \tau_{xz}, \tau_{xy})$ is the stress tensor, f_x, f_y, f_z are point forces applied along each axis, and τ_1, τ_2, τ_3 are the increments of the normal stresses. Note that given the homogeneity of the physical parameters and ignoring the source terms, equation (1) could be obtained by eliminating the stress terms in equation (11).

Throughout this study, the discretization is carried out based on equation (11). The point forces and normal stress increments are omitted in the discretization for convenience. The staggered-grid stencil is illustrated in Fig. 1. We use the fourth-order staggered-grid FD scheme [30] to approximate the spatial derivatives in system (11). For example, the x -derivative of $v_x, \partial_x v_x$, has the form:

$$[\partial_x v_x]_{i,j,k} \approx \frac{1}{h} [D_x v_x]_{i,j,k} = \frac{1}{h} \left[c_1 (v_{xi+\frac{1}{2},j,k} - v_{xi-\frac{1}{2},j,k}) + c_2 (v_{xi+\frac{3}{2},j,k} - v_{xi-\frac{3}{2},j,k}) \right], \tag{12}$$

where the indices i, j, k correspond to the discrete x, y, z grid-points. For conventional fourth-order scheme, we have $c_1 = 9/8, c_2 = -1/24$. In the dispersion analysis where the elastic moduli and the density are considered as constant, the discrete approximation of the second derivative $\partial_{xx} v_x$ takes the form

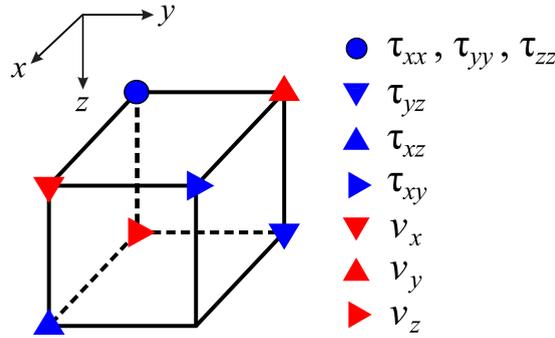


Fig. 1. 3D staggered-grid finite-difference stencil. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

$$\begin{aligned}
 [\partial_{xx} v_x]_{i+\frac{1}{2},j,k} &\approx \frac{1}{h^2} D_{xx} v_{xi+\frac{1}{2},j,k} \\
 &= \frac{1}{h^2} \left[c_1 ([D_x v_x]_{i+1,j,k} - [D_x v_x]_{i,j,k}) + c_2 ([D_x v_x]_{i+2,j,k} - [D_x v_x]_{i-1,j,k}) \right] \\
 &= \frac{1}{h^2} \left[c_1 \left([c_1 (v_{xi+\frac{3}{2},j,k} - v_{xi+\frac{1}{2},j,k}) + c_2 (v_{xi+\frac{5}{2},j,k} - v_{xi-\frac{1}{2},j,k})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{xi+\frac{1}{2},j,k} - v_{xi-\frac{1}{2},j,k}) + c_2 (v_{xi+\frac{3}{2},j,k} - v_{xi-\frac{3}{2},j,k})] \right) \right. \\
 &\quad \left. + c_2 \left([c_1 (v_{xi+\frac{5}{2},j,k} - v_{xi+\frac{3}{2},j,k}) + c_2 (v_{xi+\frac{7}{2},j,k} - v_{xi+\frac{1}{2},j,k})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{xi-\frac{1}{2},j,k} - v_{xi-\frac{3}{2},j,k}) + c_2 (v_{xi+\frac{1}{2},j,k} - v_{xi-\frac{5}{2},j,k})] \right) \right].
 \end{aligned} \tag{13}$$

The expressions of $D_{yy}, D_{zz}, D_{yz}, D_{zx}$ and D_{xy} could be obtained similarly and are given as follows

$$\begin{aligned}
 [\partial_{yy} v_y]_{i,j+\frac{1}{2},k} &\approx \frac{1}{h^2} D_{yy} v_{yi,j+\frac{1}{2},k} \\
 &= \frac{1}{h^2} \left[c_1 \left([c_1 (v_{yi,j+\frac{3}{2},k} - v_{yi,j+\frac{1}{2},k}) + c_2 (v_{yi,j+\frac{5}{2},k} - v_{yi,j-\frac{1}{2},k})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{yi,j+\frac{1}{2},k} - v_{yi,j-\frac{1}{2},k}) + c_2 (v_{yi,j+\frac{3}{2},k} - v_{yi,j-\frac{3}{2},k})] \right) \right. \\
 &\quad \left. + c_2 \left([c_1 (v_{yi,j+\frac{5}{2},k} - v_{yi,j+\frac{3}{2},k}) + c_2 (v_{yi,j+\frac{7}{2},k} - v_{yi,j+\frac{1}{2},k})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{yi,j-\frac{1}{2},k} - v_{yi,j-\frac{3}{2},k}) + c_2 (v_{yi,j+\frac{1}{2},k} - v_{yi,j-\frac{5}{2},k})] \right) \right],
 \end{aligned} \tag{14}$$

$$\begin{aligned}
 [\partial_{zz} v_z]_{i,j,k+\frac{1}{2}} &\approx \frac{1}{h^2} D_{zz} v_{zi,j,k+\frac{1}{2}} \\
 &= \frac{1}{h^2} \left[c_1 \left([c_1 (v_{zi,j,k+\frac{3}{2}} - v_{zi,j,k+\frac{1}{2}}) + c_2 (v_{zi,j,k+\frac{5}{2}} - v_{zi,j,k-\frac{1}{2}})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{zi,j,k+\frac{1}{2}} - v_{zi,j,k-\frac{1}{2}}) + c_2 (v_{zi,j,k+\frac{3}{2}} - v_{zi,j,k-\frac{3}{2}})] \right) \right. \\
 &\quad \left. + c_2 \left([c_1 (v_{zi,j,k+\frac{5}{2}} - v_{zi,j,k+\frac{3}{2}}) + c_2 (v_{zi,j,k+\frac{7}{2}} - v_{zi,j,k+\frac{1}{2}})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{zi,j,k-\frac{1}{2}} - v_{zi,j,k-\frac{3}{2}}) + c_2 (v_{zi,j,k+\frac{1}{2}} - v_{zi,j,k-\frac{5}{2}})] \right) \right],
 \end{aligned} \tag{15}$$

$$\begin{aligned}
 [\partial_{yz} v_y]_{i,j,k+\frac{1}{2}} &\approx \frac{1}{h^2} D_{yz} v_{yi,j,k+\frac{1}{2}} \\
 &= \frac{1}{h^2} \left[c_1 \left([c_1 (v_{yi,j+\frac{1}{2},k+1} - v_{yi,j+\frac{1}{2},k}) + c_2 (v_{yi,j+\frac{1}{2},k+2} - v_{yi,j+\frac{1}{2},k-1})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{yi,j-\frac{1}{2},k+1} - v_{yi,j-\frac{1}{2},k}) + c_2 (v_{yi,j-\frac{1}{2},k+2} - v_{yi,j-\frac{1}{2},k-1})] \right) \right. \\
 &\quad \left. + c_2 \left([c_1 (v_{yi,j+\frac{3}{2},k+1} - v_{yi,j+\frac{3}{2},k}) + c_2 (v_{yi,j+\frac{3}{2},k+2} - v_{yi,j+\frac{3}{2},k-1})] \right. \right. \\
 &\quad \left. \left. - [c_1 (v_{yi,j-\frac{3}{2},k+1} - v_{yi,j-\frac{3}{2},k}) + c_2 (v_{yi,j-\frac{3}{2},k+2} - v_{yi,j-\frac{3}{2},k-1})] \right) \right],
 \end{aligned} \tag{16}$$

$$\begin{aligned}
 [\partial_{zx} v_z]_{i+\frac{1}{2},j,k} &\approx \frac{1}{h^2} D_{zx} v_{zi+\frac{1}{2},j,k} \\
 &= \frac{1}{h^2} \left[c_1 \left([c_1 (v_{zi+1,j,k+\frac{1}{2}} - v_{zi+1,j,k-\frac{1}{2}}) + c_2 (v_{zi+1,j,k+\frac{3}{2}} - v_{zi+1,j,k-\frac{3}{2}})] \right) \right]
 \end{aligned}$$

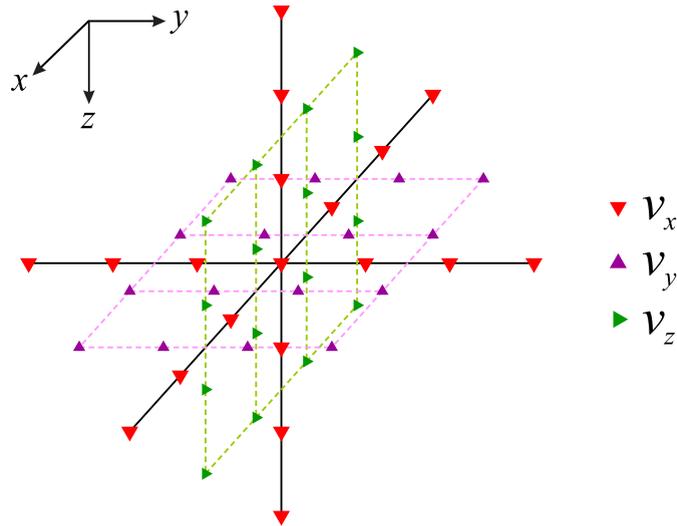


Fig. 2. Grid-points involved in the discretization of the first equation of system (1). The point $v_{xi+\frac{1}{2},j,k}$ locates at the center point. The other v_x points involved are all along the three axes through $v_{xi+\frac{1}{2},j,k}$ and are marked in red. The v_y points are on the x - y plane, marked in purple and v_z points are on the x - z plane, marked in green. The grid-points corresponding to the stress components eliminated by the parsimonious strategy are not shown here. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

$$\begin{aligned}
 & - \left[c_1(v_{zi,j,k+\frac{1}{2}} - v_{zi,j,k-\frac{1}{2}}) + c_2(v_{zi,j,k+\frac{3}{2}} - v_{zi,j,k-\frac{3}{2}}) \right] \\
 & + c_2 \left(\left[c_1(v_{zi+2,j,k+\frac{1}{2}} - v_{zi+2,j,k-\frac{1}{2}}) + c_2(v_{zi+2,j,k+\frac{3}{2}} - v_{zi+2,j,k-\frac{3}{2}}) \right] \right. \\
 & \left. - \left[c_1(v_{zi-1,j,k+\frac{1}{2}} - v_{zi-1,j,k-\frac{1}{2}}) + c_2(v_{zi-1,j,k+\frac{3}{2}} - v_{zi-1,j,k-\frac{3}{2}}) \right] \right),
 \end{aligned} \tag{17}$$

$$\begin{aligned}
 [\partial_{xy} v_x]_{i,j+\frac{1}{2},k} & \approx \frac{1}{h^2} D_{zx} v_{xi,j+\frac{1}{2},k} \\
 & = \frac{1}{h^2} \left[c_1 \left(\left[c_1(v_{xi+\frac{1}{2},j+1,k} - v_{xi+\frac{1}{2},j,k}) + c_2(v_{xi+\frac{1}{2},j+2,k} - v_{xi+\frac{1}{2},j-1,k}) \right] \right. \right. \\
 & \quad \left. \left. - \left[c_1(v_{xi-\frac{1}{2},j+1,k} - v_{xi-\frac{1}{2},j,k}) + c_2(v_{xi-\frac{1}{2},j+2,k} - v_{xi-\frac{1}{2},j-1,k}) \right] \right) \right. \\
 & \quad \left. + c_2 \left(\left[c_1(v_{xi+\frac{3}{2},j+1,k} - v_{xi+\frac{3}{2},j,k}) + c_2(v_{xi+\frac{3}{2},j+2,k} - v_{xi+\frac{3}{2},j-1,k}) \right] \right. \right. \\
 & \quad \left. \left. - \left[c_1(v_{xi-\frac{3}{2},j+1,k} - v_{xi-\frac{3}{2},j,k}) + c_2(v_{xi-\frac{3}{2},j+2,k} - v_{xi-\frac{3}{2},j-1,k}) \right] \right) \right].
 \end{aligned} \tag{18}$$

Note that equations (13)–(18) are used only in the dispersion analysis to obtain the expressions of the FD operators. The scheme used for the numerical modeling is a staggered-grid scheme developed directly from the first-order system (11), where the model parameters may be inhomogeneous and can not be put outside the derivatives.

Following the parsimonious approach proposed by Luo and Schuster [33], the stress terms in equations (11) can be eliminated to derive a discrete form of the second-order velocity formulation of the elastic system as in system (1). Thus the derived linear system, which only relates to the particle velocities, is smaller in size and is easier to solve. We illustrate the stencil corresponding to the first equation on v_x of system (1) in Fig. 2. The v_x , v_y and v_z points involved in the discretization of the first equation of system (1) are marked in red, purple and green respectively. One may notice the combination of the fourth-order staggered-grid scheme and the parsimonious approach produces a larger stencil, which is distinct from discretizing the second-order system (1) directly with the fourth-order scheme on the standard grid.

In order to reduce the numerical dispersion, an ALM strategy is applied following Marfurt [34], Jo et al. [25] and Min et al. [35]. The mass operator D_m , which distributes the acceleration terms over the grid-points involved in the stencil, is defined as a linear combination of the central grid-point and 18 adjacent grid-points. The red triangles shown in Fig. 2 indicate the grid-points related to operator D_m . The 19 grid-points are divided into four groups such that each group is equidistant from the central grid-point. Thus the acceleration term ρv_x can be approximated as

$$\begin{aligned}
 (\rho v_x)_{i+\frac{1}{2},j,k} & \approx D_m (\rho v_x)_{i+\frac{1}{2},j,k} \\
 & = a_1 (\rho v_x)_{i+\frac{1}{2},j,k} \\
 & \quad + a_2 \left[(\rho v_x)_{i+\frac{3}{2},j,k} + (\rho v_x)_{i-\frac{1}{2},j,k} + (\rho v_x)_{i+\frac{1}{2},j+1,k} \right. \\
 & \quad \left. + (\rho v_x)_{i+\frac{1}{2},j-1,k} + (\rho v_x)_{i+\frac{1}{2},j,k+1} + (\rho v_x)_{i+\frac{1}{2},j,k-1} \right]
 \end{aligned} \tag{19}$$

$$\begin{aligned}
& + a_3 [(\rho v_x)_{i+\frac{5}{2},j,k} + (\rho v_x)_{i-\frac{3}{2},j,k} + (\rho v_x)_{i+\frac{1}{2},j+2,k} \\
& \quad + (\rho v_x)_{i+\frac{1}{2},j-2,k} + (\rho v_x)_{i+\frac{1}{2},j,k+2} + (\rho v_x)_{i+\frac{1}{2},j,k-2}] \\
& + a_4 [(\rho v_x)_{i+\frac{7}{2},j,k} + (\rho v_x)_{i-\frac{5}{2},j,k} + (\rho v_x)_{i+\frac{1}{2},j+3,k} \\
& \quad + (\rho v_x)_{i+\frac{1}{2},j-3,k} + (\rho v_x)_{i+\frac{1}{2},j,k+3} + (\rho v_x)_{i+\frac{1}{2},j,k-3}],
\end{aligned}$$

where the weighting coefficients a_1 , a_2 , a_3 and a_4 determine the relative proportion of each grid-point group. The approximations of ρv_y and ρv_z can be derived similarly.

2.3. Determination of coefficients

Following a classic dispersion analysis process, a vector plane-wave solution $(v_x, v_y, v_z) = (v_{x0}, v_{y0}, v_{z0})e^{-i(k_x x + k_y y + k_z z)}$ is substituted into the discrete differential operators (13)–(18), and into the mass operator (19). Here the components of the wavenumber k_x, k_y, k_z are defined as

$$k_x = k \sin \theta \cos \phi, \quad k_y = k \sin \theta \sin \phi, \quad k_z = k \cos \theta, \quad (20)$$

where $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is the propagation angle with respect to the z -axis, and $\phi \in [-\pi, \pi]$ is the propagation angle with respect to the x -axis.

With proper algebraic rearrangement, we obtain the following expressions of the discrete differential operators:

$$\begin{aligned}
D_{xx} &= 2 \left(-c_1^2 - c_2^2 + c_1^2 \cos(\Delta k \sin \theta \cos \phi) + c_2^2 \cos(3hk \sin \theta \cos \phi) \right. \\
&\quad \left. + 2c_1 c_2 [\cos(2hk \sin \theta \cos \phi) - \cos(hk \sin \theta \cos \phi)] \right), \\
D_{yy} &= 2 \left(-c_1^2 - c_2^2 + c_1^2 \cos(hk \sin \theta \sin \phi) + c_2^2 \cos(3hk \sin \theta \sin \phi) \right. \\
&\quad \left. + 2c_1 c_2 [\cos(2hk \sin \theta \sin \phi) - \cos(hk \sin \theta \sin \phi)] \right), \\
D_{zz} &= 2 \left(-c_1^2 - c_2^2 + c_1^2 \cos(hk \cos \theta) + c_2^2 \cos(3hk \cos \theta) \right. \\
&\quad \left. + 2c_1 c_2 [\cos(2hk \cos \theta) - \cos(hk \cos \theta)] \right), \\
D_{yz} &= -4 \left(c_1^2 [\sin(\frac{1}{2}hk \sin \theta \sin \phi) \sin(\frac{1}{2}hk \cos \theta)] \right. \\
&\quad + c_2^2 [\sin(\frac{3}{2}hk \sin \theta \sin \phi) \sin(\frac{3}{2}hk \cos \theta)] \\
&\quad + c_1 c_2 [\sin(\frac{1}{2}hk \sin \theta \sin \phi) \sin(\frac{3}{2}hk \cos \theta) \\
&\quad \quad \left. + \sin(\frac{3}{2}hk \sin \theta \sin \phi) \sin(\frac{1}{2}hk \cos \theta)] \right), \\
D_{zx} &= -4 \left(c_1^2 [\sin(\frac{1}{2}hk \sin \theta \cos \phi) \sin(\frac{1}{2}hk \cos \theta)] \right. \\
&\quad + c_2^2 [\sin(\frac{3}{2}hk \sin \theta \cos \phi) \sin(\frac{3}{2}hk \cos \theta)] \\
&\quad + c_1 c_2 [\sin(\frac{1}{2}hk \sin \theta \cos \phi) \sin(\frac{3}{2}hk \cos \theta) \\
&\quad \quad \left. + \sin(\frac{3}{2}hk \sin \theta \cos \phi) \sin(\frac{1}{2}hk \cos \theta)] \right), \\
D_{xy} &= -4 \left(c_1^2 [\sin(\frac{1}{2}hk \sin \theta \cos \phi) \sin(\frac{1}{2}hk \sin \theta \sin \phi)] \right. \\
&\quad + c_2^2 [\sin(\frac{3}{2}hk \sin \theta \cos \phi) \sin(\frac{3}{2}hk \sin \theta \sin \phi)] \\
&\quad + c_1 c_2 [\sin(\frac{1}{2}hk \sin \theta \cos \phi) \sin(\frac{3}{2}hk \sin \theta \sin \phi) \\
&\quad \quad \left. + \sin(\frac{3}{2}hk \sin \theta \cos \phi) \sin(\frac{1}{2}hk \sin \theta \sin \phi)] \right),
\end{aligned} \quad (21)$$

and the mass operator:

$$\begin{aligned}
D_m &= a_1 + 2a_2 [\cos(hk \sin \theta \cos \phi) + \cos(hk \sin \theta \sin \phi) + \cos(hk \cos \theta)] \\
&\quad + 2a_3 [\cos(2hk \sin \theta \cos \phi) + \cos(2hk \sin \theta \sin \phi) + \cos(2hk \cos \theta)] \\
&\quad + 2a_4 [\cos(3hk \sin \theta \cos \phi) + \cos(3hk \sin \theta \sin \phi) + \cos(3hk \cos \theta)].
\end{aligned} \quad (22)$$

Substituting equation (7) in equations (21) and (22) renders these operators depending only on K_s (shear wavenumber in grid-point units) and the propagation angles θ and ϕ . Then, the normalized phase velocities can be expressed as functions of K_s, θ, ϕ , Poisson's ratio σ and the FD coefficients $(c_1, c_2, a_0, a_1, a_2, a_3)$.

As mentioned in [35], the coefficients which give the best phase velocities also lead to the best group velocities. Therefore, we consider only the phase velocities to determine the optimal coefficients. The numerical dispersion of group velocities will also be investigated. We first define a vector function

$$\mathbf{F}(\boldsymbol{\beta}, \mathbf{m}) = \left(\frac{V_p^{\text{ph}}}{V_p}, \frac{V_{s1}^{\text{ph}}}{V_s}, \frac{V_{s2}^{\text{ph}}}{V_s} \right)^T, \tag{23}$$

which returns the normalized phase velocities of *P*- and *S*-wave. The argument $\boldsymbol{\beta} = (c_1, c_2, a_0, a_1, a_2, a_3)$ is the set of coefficients to optimize and \mathbf{m} contains all selected values of K_s, θ, ϕ and σ . Hence the error vector is defined as

$$\mathbf{e}(\boldsymbol{\beta}) = \mathbf{d} - \mathbf{F}(\boldsymbol{\beta}, \mathbf{m}), \tag{24}$$

where \mathbf{d} is a vector of unities. The misfit between the normalized phase velocities and the unity can be cast in a least-square sense as

$$E(\boldsymbol{\beta}) = \mathbf{e}^T \mathbf{e} = \|\mathbf{d} - \mathbf{F}(\boldsymbol{\beta}, \mathbf{m})\|_2^2, \tag{25}$$

where $E(\boldsymbol{\beta})$ is a scalar function of the target coefficients. By using Taylor expansion, function $\mathbf{F}(\boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta}, \mathbf{m})$ can be approximated as

$$\mathbf{F}(\boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta}, \mathbf{m}) = \mathbf{F}(\boldsymbol{\beta}_0, \mathbf{m}) + \mathbf{J}(\boldsymbol{\beta}_0, \mathbf{m})\Delta\boldsymbol{\beta}, \tag{26}$$

where $\mathbf{J}(\boldsymbol{\beta}_0, \mathbf{m}) = \left. \frac{\partial \mathbf{F}(\boldsymbol{\beta}, \mathbf{m})}{\partial \boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}$ is the Jacobian matrix with respect to $\boldsymbol{\beta}$. The expression of $E(\boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta})$ could be obtained subsequently by substituting equation (26). We try to seek an optimal $\Delta\boldsymbol{\beta}$ which minimizes $E(\boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta})$. Taking the derivative of E with respect to $\Delta\boldsymbol{\beta}$ and setting the result to zero gives

$$\Delta\boldsymbol{\beta} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T (\mathbf{d} - \mathbf{F}(\boldsymbol{\beta}_0, \mathbf{m})). \tag{27}$$

Updating the coefficients with equation (27) gives the Gauss–Newton algorithm. However, the convergence of the iteration may suffer from the ill-conditioning of matrix $\mathbf{J}^T \mathbf{J}$. We therefore resort to the more robust Levenberg–Marquardt algorithm [37,51], which replaces equation (27) by a “damped version”

$$\Delta\boldsymbol{\beta} = (\mathbf{J}^T \mathbf{J} + \alpha \mathbf{I})^{-1} \mathbf{J}^T (\mathbf{d} - \mathbf{F}(\boldsymbol{\beta}_0, \mathbf{m})), \tag{28}$$

where \mathbf{I} is the identity matrix and α is the damping factor. Introducing this damped identity matrix could guarantee the invertibility of matrix $\mathbf{J}^T \mathbf{J} + \alpha \mathbf{I}$ and prevent severe divergence during the iteration. Thus the set of coefficients is updated in an iterative manner

$$\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta}, \tag{29}$$

and the optimal set of coefficients is determined after the convergence is achieved.

For the sake of discretization, we denote a chosen sets of parameters as $\mathbf{m}_i, i = 1, \dots, N$, where N depends on the range and the interval length of Poisson’s ratio, propagation angles, and the number of grid-points per minimum shear wavelength. Specifically we have $N = N_\sigma \times N_\theta \times N_\phi \times N_{K_s}$, where $N_\sigma, N_\theta, N_\phi, N_{K_s}$ are respectively the number of selected values of the corresponding parameter. Thus the vector function $\mathbf{F}(\boldsymbol{\beta}, \mathbf{m})$ takes the form

$$\mathbf{F}(\boldsymbol{\beta}, \mathbf{m}) = [\mathbf{F}_1(\boldsymbol{\beta}, \mathbf{m}_1)^T, \dots, \mathbf{F}_N(\boldsymbol{\beta}, \mathbf{m}_N)^T]^T, \tag{30}$$

where \mathbf{F}_i is a three-valued function with respect to a given set of $K_{si}, \theta_i, \phi_i, \sigma_i$, making \mathbf{F} a vector function of $3N$ components. The size of the Jacobi matrix \mathbf{J} is thus $3N \times 6$, with 6 being the number of coefficients. In this work, K_s and σ vary from 0.01 to 0.33 in steps of 0.01 and the angles θ and ϕ range from 0 to $\pi/4$ in steps of $\pi/12$. Although the maximum σ is set to 0.33, the numerical dispersion stays small for σ up to 0.49 (see Figs. 3 and 4). Thus we have $N = 17424 (= 33 \times 33 \times 4 \times 4)$ and the size of \mathbf{J} becomes 52272×6 . Despite the large number of rows of \mathbf{J} , the size of $\mathbf{J}^T \mathbf{J}$ is merely 6×6 , making equation (28) easy to calculate. The initial set of coefficients $\boldsymbol{\beta}_0$ is chosen as for the conventional fourth-order staggered-grid scheme and without using the ALM strategy, i.e., $\boldsymbol{\beta}_0 = (1, 0, 0, 0, 9/8, -1/24)$. The damped factor α is set to 0.1 and the iteration stops at the fifth step. Due to the fast convergence of the Levenberg–Marquardt algorithm and the small number of coefficients to optimize, the misfit function becomes small enough after only 5 iterations and additional iterations do not bring much improvement. The optimal values of these coefficients are given in Table 1. The corresponding values of misfit function $E(\boldsymbol{\beta})$ are also listed. As seen in the table, $E(\boldsymbol{\beta}_{\text{opt}})$ is about two order of magnitude smaller than $E(\boldsymbol{\beta}_{\text{con}})$, where $\boldsymbol{\beta}_{\text{con}} = \boldsymbol{\beta}_0$. Because the optimal coefficients are obtained by considering a vast range of the propagation angles and nearly a full range of Poisson’s ratio, the numerical dispersion is expected to be relatively insensitive to the physical properties of the medium.

An alternative way to minimize the misfit function is to utilize the global optimization scheme, such as the simulated annealing algorithm [58], the genetic algorithm [64], etc. Unlike the gradient-based Levenberg–Marquardt algorithm, the global optimization scheme could prevent the inversion from getting trapped in local minima. We perform the simulated

Table 1

Conventional coefficients and optimal coefficients obtained by minimizing the misfit between the normalized phase velocities and unity, using the Levenberg–Marquardt method and the simulated annealing algorithm. The corresponding values of the misfit functions are given in the last row.

| | β_{con} | $\beta_{\text{opt}}^{\text{LM}}$ | $\beta_{\text{opt}}^{\text{SA}}$ |
|------------|----------------------|----------------------------------|----------------------------------|
| a_1 | 1 | 0.926585 | 0.923527 |
| a_2 | 0 | 0.069999 | 0.069755 |
| a_3 | 0 | −0.032869 | −0.032755 |
| a_4 | 0 | 0.006283 | 0.006260 |
| c_1 | 9/8 | 1.252740 | 1.250650 |
| c_2 | −1/24 | −0.054599 | −0.054510 |
| $E(\beta)$ | 1.430142E−2 | 7.451415E−4 | 7.400696E−4 |

annealing minimization on the continuous coefficient range by using a modification of the downhill simplex method [50]. The derived coefficients and the corresponding misfit function value are also listed in Table 1. The two sets of coefficients and the misfit function values are very close to each other. The dispersion curves not presented here lead to a conclusion which is completely in conformity with the comments given by Gosselin-Cliche and Giroux [21]: although the value of the misfit function obtained from the global optimization scheme may be slightly smaller and the optimal coefficients may be different, the dispersion curves correspond to the least-squares inversion and the global minimization presents great similarity, both in shape and amplitude. We rely on the coefficients given by the Levenberg–Marquardt algorithm in this work.

2.4. Dispersion analysis

With the optimal coefficients given in Table 1, we now investigate the numerical dispersion of the proposed scheme through checking the consistency between the normalized velocities and the unity. The normalized phase and group velocities with conventional and optimal coefficients are presented in Figs. 3 and 4 as functions of the shear wavenumber in grid-point units K_s , for propagation angles being 0, $\pi/12$, $\pi/6$ and $\pi/4$, and for Poisson's ratios of 0.25 and 0.49. Normalized group velocities of P - and S -wave can be obtained by calculating $\partial\omega/\partial k$ from equation (5). The expressions of the normalized group velocities are given in Appendix B. As seen in the figures, the normalized phase and group velocities of P - and S -wave given by the optimized scheme are less dispersive and the dispersion is isotropic in terms of different propagation angles (i.e., the curves for different propagation angles are more clustered). Moreover, the grid dispersion and grid anisotropy are relatively insensitive to Poisson's ratio. Slight difference between $\sigma = 0.25$ and $\sigma = 0.49$ only appears for the P -wave phase and group velocities. Other values of Poisson's ratio give similar dispersion curves. Thus we exemplify the analysis by the results of $\sigma = 0.25$ and $\sigma = 0.49$. Note that even for a rather high Poisson's ratio 0.49, the dispersion remains well-conditioned, with the P -wave phase and group velocities even closer to the unity. This interesting phenomenon stems from the good properties of the staggered-grid scheme, which is known to be favorable for wave modeling with high Poisson's ratio [60,30].

The error between the normalized velocities and the unity is greatly reduced using the optimal coefficients. Indeed, the error of the conventional scheme becomes larger than 4.8% for the phase velocities V_{s1}^{ph} and V_{s2}^{ph} as K_s reaches 0.3, while the counterpart with the optimal coefficients remains below 0.12%, which is a great improvement. Although the situation for the group velocities V_{s1}^{gr} and V_{s2}^{gr} is slightly degraded, the error produced by the optimal coefficients is still lower than 2.2%, while the curves of the conventional scheme have already gone beyond the frame of the figures (error of 21%). The correspondence between the group velocities with the optimal coefficients and the unity implies the consistency to the assumption that the coefficients which give the best phase velocities also lead to the best group velocities. Therefore, if the error is set to certain level, the K_s required by the optimal scheme is larger, which means that we could sample fewer grid-points per minimum shear wavelength for the same wave modeling problem. In other words, we are capable of performing the frequency-domain wave modeling at a higher frequency with fixed geometry settings. The reduction in the number of grid-points per minimum shear wavelength entitles a coarser sampling of the model, which generates a smaller linear system where the impedance matrix has the same sparsity pattern as the one derived from the conventional scheme. Thus the computational cost of solving the linear system could be dramatically reduced. Table 2 presents the number of grid-points per minimum shear wavelength required by the optimal scheme (N_{opt}) and by the conventional scheme (N_{con}), with the upper limit of dispersion error (e_{lim}) of 1%, 2% and 3% respectively (with respect to more dispersive group velocities). As seen in the table, if the upper limit of error is set to 1%, only 3.7 grid-points per minimum shear wavelength is required using the optimal coefficients, while the number increases to 7.69 for the conventional scheme. We define ratio $R_N = N_{\text{opt}}/N_{\text{con}}$ as the relative reduction of the sampling requirement in one dimension. Suppose that the model is discretized into a cube with identical number of grid-points in each dimension, then the size of the impedance matrix is $3N^3 \times 3N^3$. Thus the linear system size is reduced by ratio

$$R_S = \frac{S_{\text{opt}}}{S_{\text{con}}} = \frac{3N_{\text{opt}}^3}{3N_{\text{con}}^3} = R_N^3. \quad (31)$$

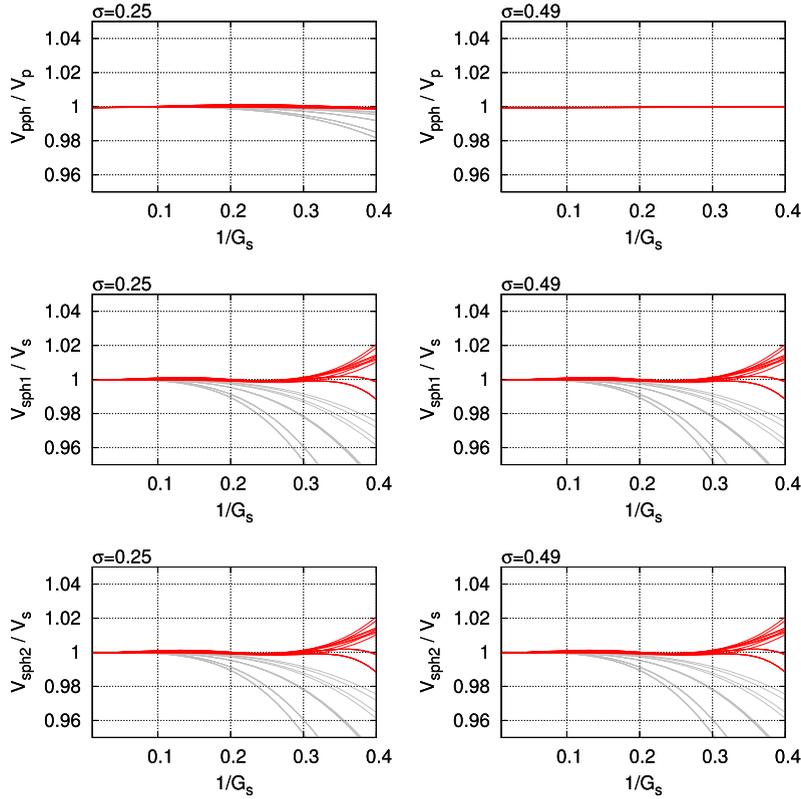


Fig. 3. Dispersion curves of normalized P-wave and S-wave **phase** velocities using the conventional coefficients (in gray) and the optimal coefficients (in red). The curves are plotted for propagation angles with respect to the x -axis and the z -axis both ranging from 0 to $\pi/4$, and for the Poisson's ratios equal to $\sigma = 0.25$ (left) and 0.49 (right). G_s is the number of grid-points per minimum S-wavelength. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

However, the reduction in system size does not imply equivalent reduction in the memory cost required for solving the system due to the usage of specific sparse storage format of the matrix. According to the analysis given in [39,31], the memory cost of using a direct method (via nested dissection reordering) or an iterative method to solve the linear system is $\mathcal{O}(N^4)$ and $\mathcal{O}(N^3)$ respectively. Thus the ratio of the reduction in storage requirement is

$$R_{ST} = \begin{cases} R_N^4, & \text{Direct method,} \\ R_N^3, & \text{Iterative method.} \end{cases} \quad (32)$$

For instance, with the upper limit of error being 1%, using an iterative solver leads to a ratio $R_{ST} = 0.48^3 = 0.11$, i.e., the memory cost of the optimal scheme is 11% of the conventional scheme. Table 2 summarizes the ratio of memory cost for each case. Compared with using the conventional scheme, the memory cost of using the optimized scheme reduces to a percentage between 5.37% and 10.3% with a direct linear solver, and for an iterative solver the ratio ranges from 11.2% to 18.2%. Note that the computational complexity also decreases, because extra algebraic operations related to solving the system are saved due to the smaller number of unknowns from a coarser sampling. Such improvements would significantly mitigate the computational burden of the 3D viscoelastic wave modeling.

2.5. Implementation of seismic attenuation

Compared with the time-domain wave modeling, incorporating the intrinsic attenuation of the medium is straightforward in the frequency-domain by using complex-valued wave velocities. In this study, we use the Kolsky–Futterman model without the dispersion term [27,16] such that

$$\bar{c}(\omega) = c \left(1 - i \operatorname{sign}(\omega) \frac{1}{2Q} \right), \quad (33)$$

where c is the velocity of P- or S- waves, i denotes the imaginary unit, $\operatorname{sign}(\cdot)$ is the sign function, ω is the angular frequency, and Q is the quality factor. Note that the smaller Q is, the more attenuating the medium is, whereas using great values for Q corresponds to almost non-attenuating medium.

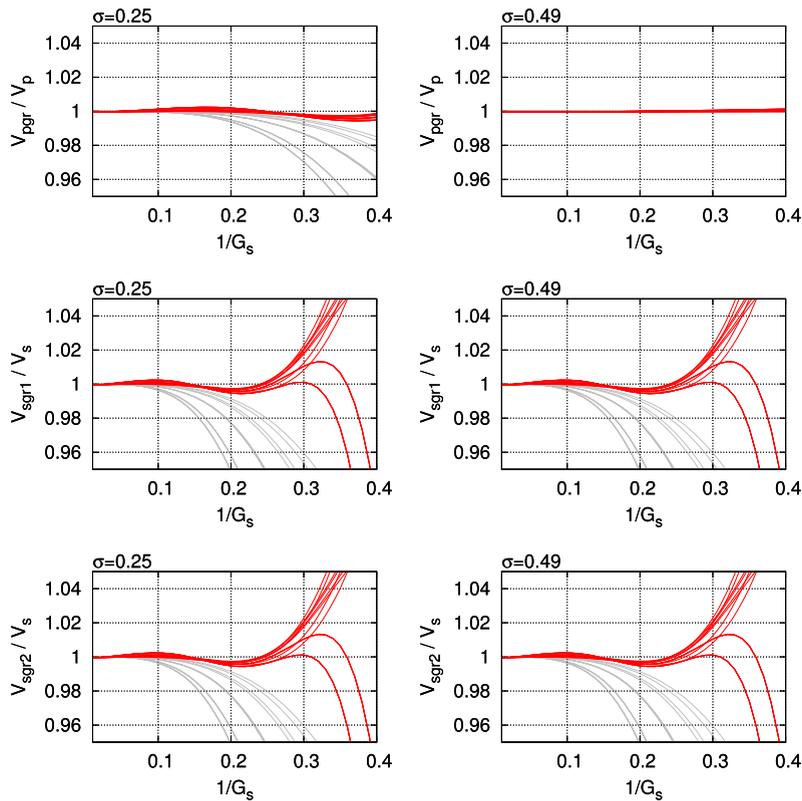


Fig. 4. Dispersion curves of normalized P-wave and S-wave **group** velocities using the conventional coefficients (in gray) and the optimal coefficients (in red). The curves are plotted for propagation angles with respect to the x -axis and the z -axis both ranging from 0 to $\pi/4$, and for the Poisson's ratios equal to $\sigma = 0.25$ (left) and 0.49 (right). G_S is the number of grid-points per minimum S-wavelength. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

Number of grid-points per minimum shear wavelength N_{opt} and N_{con} required by different upper error limits, the corresponding reduction in sampling R_N , and the percentage of reduction in memory cost using direct solvers R_{ST} (Dir) or iterative solvers R_{ST} (Itr).

| ϵ_{lim} | N_{opt} | N_{con} | R_N | R_{ST} (Dir) | R_{ST} (Itr) |
|-------------------------|------------------|------------------|-------|-----------------------|-----------------------|
| 1% | 3.70 | 7.69 | 0.48 | 5.37% | 11.2% |
| 2% | 3.45 | 6.67 | 0.52 | 7.16% | 13.8% |
| 3% | 3.33 | 5.88 | 0.57 | 10.3% | 18.2% |

3. Numerical results

In this section, we first give a brief introduction of the CARP-CG method. Then we validate the accuracy of the optimal scheme by comparing the numerical results with the analytic solutions in homogeneous media. The superiority over the conventional scheme is also verified through presenting the profiles of the wavefields and through the error analysis with different experimental settings.

3.1. The CARP-CG method

CARP-CG method is a conjugate gradient (CG) acceleration of the component-averaged row projections (CARP) algorithm [18,20]. The acceleration procedure follows the ideas of the CGMN method [4,19], where the symmetric Kaczmarz row projections [26] is used to transform the indefinite ill-conditioned linear system into a symmetric positive semi-definite system. CG is thus applicable to solve the resultant system. With the component averaging operations, CARP acts as a block-parallel version of the Kaczmarz method (which is inherently sequential).

As mentioned in [18], the component averaging operations are equivalent to performing orthogonal projections onto certain hyperplanes, so by representing these hyperplanes as equations and adding these so-called “averaging equations” to the system, the projections can be done as part of the Kaczmarz iterations. Thus CARP-CG in the regular space is mathematically equivalent to CGMN in the extended space. As a sequence, CARP-CG shares the same robustness as the CGMN method.

Because the impedance matrix is sparse and the bandwidth is small enough, the components are only shared by adjacent matrix blocks, which means the averaging operations are limited to the subdomains that share a boundary. Therefore, the communication is limited to two cores, which leads to a good scalability of CARP-CG [31].

In the following experiments, the stopping criterion for the iterations is based on the relative residual:

$$\frac{\|\mathbf{b} - \mathbf{A}\mathbf{x}^k\|}{\|\mathbf{b} - \mathbf{A}\mathbf{x}^0\|} < 10^{-5}, \tag{34}$$

where \mathbf{x}^0 is the starting point, which is set to $\mathbf{0}$. The relaxation parameter $\gamma = 1$ is selected by trial and error. Note that the number of blocks divided from the matrix is equal to the number of processors used in the experiments. In our particular implementation, the grid-points are ordered in the z, y, x manner. Hence the partitioning of the matrix is equivalent to a division in the x -axis in terms of the physical domain, i.e., the cube is divided into several slabs along x -axis. Because most of the 3D model has smaller y and z dimensions, such a division yields an impedance matrix with a smaller bandwidth.

3.2. Comparison with analytic solutions

3.2.1. The acoustic case

By setting the velocity of shear wave to zero and using an explosive source, the propagation of acoustic waves inside an infinite homogeneous medium can be simulated with the elastic setups. In this case, any normal stress component (τ_{xx} , τ_{yy} and τ_{zz}) is equivalent to the acoustic pressure P . With $V_s = 0$ and an explosive source (a same increment τ for τ_{xx} , τ_{yy} and τ_{zz}), the elastic system (11) reduces to the following acoustic case

$$\begin{cases} -\omega^2 \tau_{xx} = V_p^2 (\partial_{xx}^2 \tau_{xx} + \partial_{yy}^2 \tau_{yy} + \partial_{zz}^2 \tau_{zz}) - \omega^2 \tau, \\ -\omega^2 \tau_{yy} = V_p^2 (\partial_{xx}^2 \tau_{xx} + \partial_{yy}^2 \tau_{yy} + \partial_{zz}^2 \tau_{zz}) - \omega^2 \tau, \\ -\omega^2 \tau_{zz} = V_p^2 (\partial_{xx}^2 \tau_{xx} + \partial_{yy}^2 \tau_{yy} + \partial_{zz}^2 \tau_{zz}) - \omega^2 \tau, \end{cases} \tag{35}$$

where $\tau_{xx} = \tau_{yy} = \tau_{zz} = P(x, y, z)$. With proper rearrangement and by setting

$$\tau = -\frac{V_p^2}{\omega^2} \delta(x - x_s)\delta(y - y_s)\delta(z - z_s), \tag{36}$$

we obtain the Helmholtz equation

$$\Delta P + k^2 P = -\delta(x - x_s)\delta(y - y_s)\delta(z - z_s), \tag{37}$$

where Δ is the Laplace operator, $k = \omega/V_p$ is the wavenumber, and $\mathbf{x}_s = (x_s, y_s, z_s)$ is the location of the point source. Thus we have the Green's function [5]

$$P(\mathbf{x}) = \frac{e^{ik|\mathbf{x}-\mathbf{x}_s|}}{4\pi|\mathbf{x}-\mathbf{x}_s|}, \tag{38}$$

as the analytic solution of equation (37).

For the numerical experiments, we keep the same discretization process of the elastic equations as described previously. The source is adapted according to equation (36). The pressure field is obtained from the calculated particle velocities. A computational domain of size $2\text{ km} \times 2\text{ km} \times 2\text{ km}$ is uniformly discretized into a $81 \times 81 \times 81$ grid with a spatial interval of 25 m. The perfectly matched layers (PML) are attached around the computational domain to absorb the outgoing waves and to avoid the artificial reflections [3]. The number of PML grid-points is set to 10 and the absorbing coefficient is given by trial and error. Thus the grid ends up to be of size $101 \times 101 \times 101$ and the degree of freedom of this linear system is 3,090,903. Solving such a linear system may consume hundreds of gigabytes of RAM for a direct solver based on the LU decomposition [21,41,42]. Using the CARP-CG method, however, only requires around 2 GB of RAM in total, and this amount is further shared by different processors if parallelization is taken into account.

An explosive point source is excited at the center of the domain. The P-wave velocity V_p is 2500 m/s, the S-wave velocity V_s is set to zero and the density is 1000 kg/m^3 . Several wave modelings are carried out at four frequencies: 10 Hz, 20 Hz, 25 Hz and 30 Hz. Given the fixed grid size and spatial interval, the corresponding number of grid-points per minimum wavelength varies from 10, 5, 4 to 3.3 respectively. Different levels of attenuation $Q = 5, 10, 20, 200, \infty$ are incorporated, where smaller value of Q generates more attenuating medium and by infinite Q we mean no attenuation. Considering the elapsed time, it requires about 728.18 seconds with 4 cores to perform a modeling at 10 Hz without attenuation, which is the most time-consuming test among different frequencies and different levels of attenuation. The number of iterations of CARP-CG is 976.

Fig. 5 presents the real part of the 3D acoustic pressure wavefields obtained from the analytic solution, the conventional scheme and the optimal scheme with $Q = \infty$ and 20 at 30 Hz (i.e., 3.3 grid-points per minimum wavelength). The real

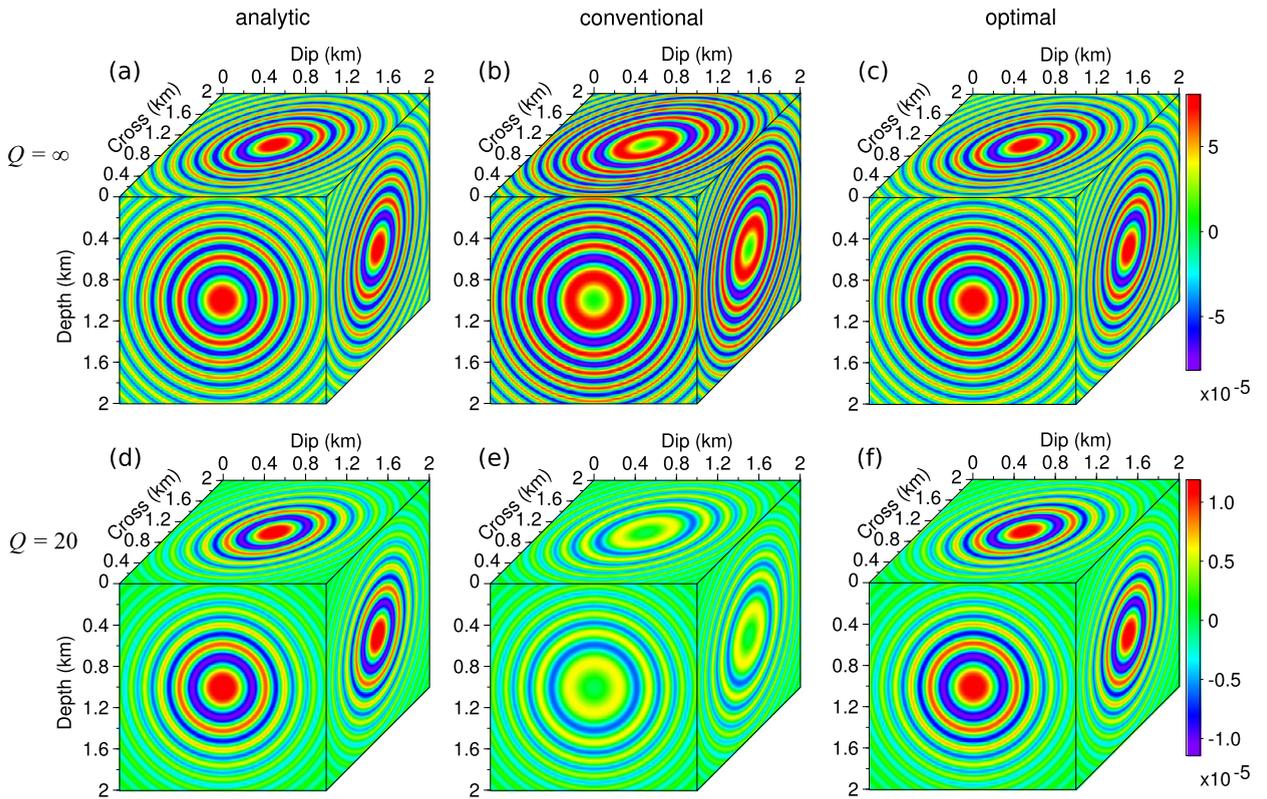


Fig. 5. 3D acoustic pressure wavefields obtained from analytic solution (first column), conventional scheme (second column) and optimal scheme (third column) with $Q = \infty$ (first row) and $Q = 20$ (second row) at 30Hz. The figures in the same row are shown with the same color scale given on the right side. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

part of the wavefields are selected to present the numerical results in the rest of this work if not specified otherwise. The wavefields with the same level of attenuation are plotted with the same color scale. Discrepancies between the analytic and the conventional wavefields can be obviously observed, while the optimal results are shown to be perfectly consistent with the analytic solutions. The attenuation effects can be also seen between the first and second rows. Compared with the wavefields with $Q = \infty$, those with $Q = 20$ are entirely attenuated and the amplitudes far from the source are relatively weaker than that at the center.

Further comparisons are carried out by presenting the profiles of the pressure P for $x = 1000\text{m}$ and $x = 675\text{m}$ in the x - z plane sliced at $y = 1000\text{m}$. Fig. 6 shows these profiles of P with $Q = 20$ at 10Hz, 20Hz, 25Hz and 30Hz. As can be seen, both the conventional and optimal results agree well at 10Hz (i.e., 10 grid-points per minimum wavelength). However, the conventional profile starts to deviate from the analytic solution as the frequency increases. The deviation can be clearly observed at 25Hz and 30Hz, especially for the profiles of $x = 675\text{m}$, while the optimal profiles fit very well. The attenuation effects can be seen more clearly in Fig. 7, where the profiles at 30Hz with $Q = 5, 10, 20, 200$ and ∞ are presented. The profile amplitude becomes increasingly attenuated as Q decreases. Still, the optimal profiles agree very well with the analytic solutions, but the conventional ones are distorted in shape.

Quantitative comparison is conducted by calculating the difference between the analytic and numerical wavefields. Taking the pressure wavefield P as example, the relative error is defined as follows

$$e_{\text{rel}} = \frac{|P_{\text{analytic}} - P_{\text{numerical}}|}{|P_{\text{analytic}}|} \quad (39)$$

The source point is skipped during the computation because the analytic solution is not defined at this point. Fig. 8 presents the relative errors of the conventional and optimal results as a function of frequency with different levels of attenuation. The first x -axis (bottom) represents the frequency and the second x -axis (top) contains the corresponding value of K_s , which is the inverse of the number of grid-points per minimum wavelength. As shown in the figure, the errors of the optimal results are generally smaller than those of the conventional ones. We see that the relative errors of the optimal results remain small between 10Hz and 30Hz, which is consistent with the range of K_s , $[0, 0.33]$, selected for the optimization. The counterpart of the conventional results appears to be relatively unsatisfactory. Note that the errors become smaller if more attenuation is incorporated. This can be explained by the fact that with more attenuation, the amplitude decays faster as the distance from the source point increases. Hence, the error also diminishes due to the attenuation.

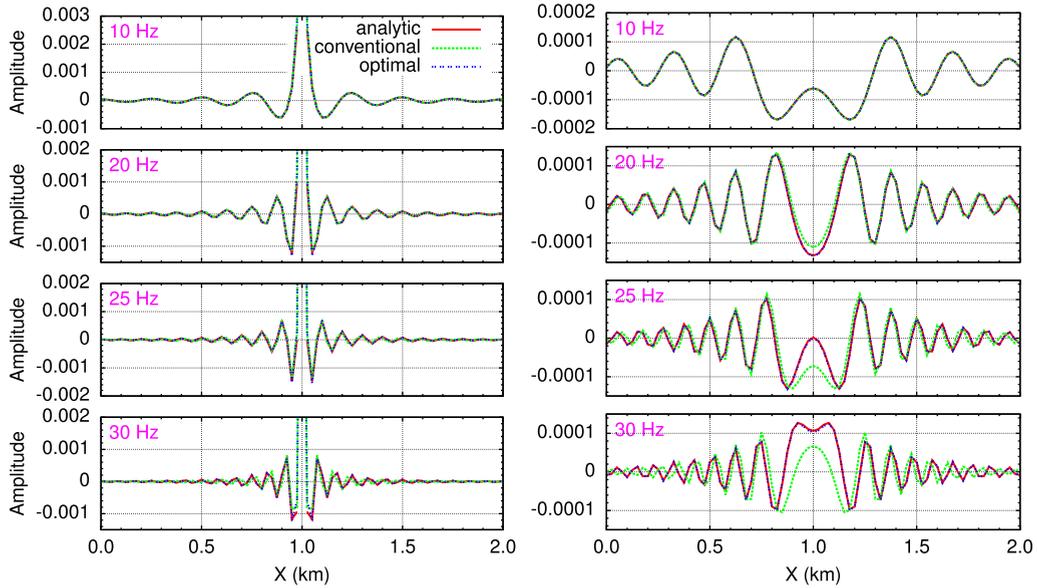


Fig. 6. Profiles of the pressure wavefield P for $x = 1000$ m (left) and $x = 675$ m (right) in the x - z plane sliced at $y = 1000$ m with $Q = 20$ at different frequencies.

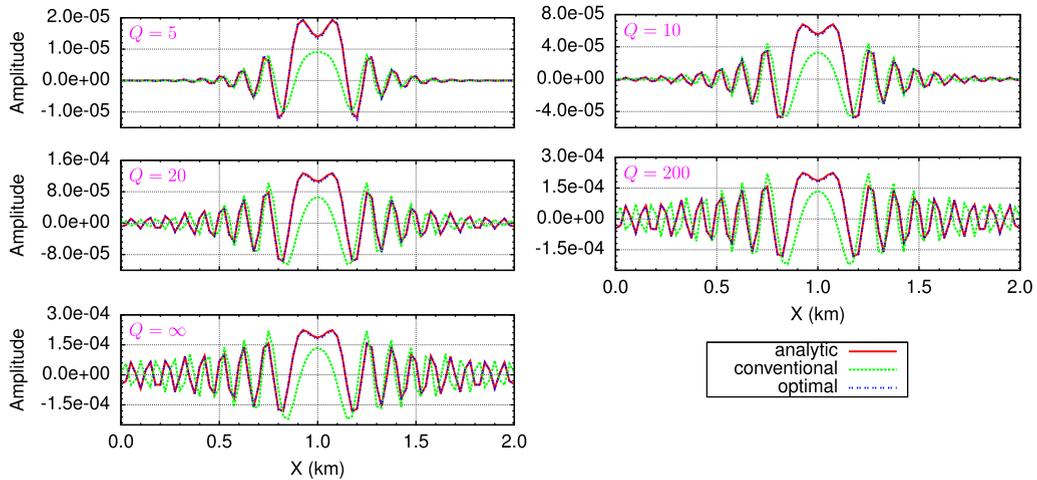


Fig. 7. Profiles of the pressure wavefield P for $x = 675$ m in the x - z plane sliced at $y = 1000$ m at 30 Hz with different levels of attenuations.

3.2.2. The elastic case

The analytic solution of the elastodynamic system is computed following the method given in [44,21]. The time-domain elastic equation with a point force along the z -axis in an infinite homogeneous medium can be written as

$$\left(\frac{\lambda + 2\mu}{\rho}\right)\nabla\nabla \cdot \mathbf{u} - \frac{\mu}{\rho}\nabla \times \nabla \times \mathbf{u} - \frac{\partial^2 \mathbf{u}}{\partial t^2} = -\mathbf{e}_z \frac{f(t)}{\rho} \frac{\delta(R)}{R^2}, \tag{40}$$

where $\mathbf{u} = (u_x, u_y, u_z)$ is the particle displacement, $R = |\mathbf{x} - \mathbf{x}_s|$ is the distance from the source and δ is the Dirac function. Note that the attenuation can be implicitly embedded in the Lamé parameters through using complex velocities (33). Transforming equation (40) into the frequency domain, Pilant [44] derives the following expression of the analytic solution after proper developments

$$\mathbf{u} = \frac{-F(\omega)}{4\pi\rho\omega^2} \left[\nabla\nabla \cdot \left(\frac{e^{-ik_p R}}{R} \mathbf{e}_z - \frac{e^{-ik_s R}}{R} \mathbf{e}_z \right) - k_s^2 \frac{e^{-ik_s R}}{R} \mathbf{e}_z \right], \tag{41}$$

where the wavenumbers k_p and k_s become complex in the viscoelastic case. To facilitate the calculation, Gosselin-Cliche and Giroux [21] provide an expanded expression of equation (41) as follows

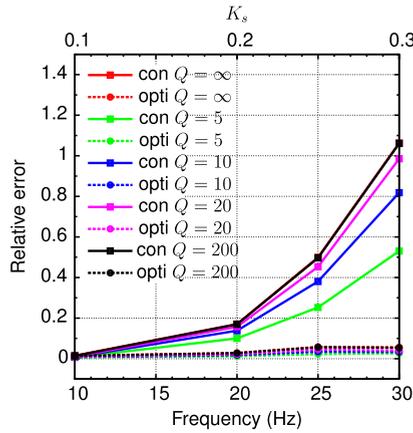


Fig. 8. Relative errors of the pressure wavefield P using the conventional and the optimal scheme, plotted as a function of the frequency with different levels of attenuation. The first x -axis (bottom) represents the frequency and the second x -axis (top) contains the corresponding value of K_s which is the inverse of the number of grid-points per minimum wavelength.

$$\begin{aligned}
 \mathbf{u} = \frac{F(\omega)}{4\pi\rho R^5\omega^2} & \left[xz \left((R^2k_p^2 - 3 - 3ik_pR)e^{-ik_pR} + (3 + 3ik_sR - R^2k_s^2)e^{-ik_sR} \right) \mathbf{e}_x \right. \\
 & + yz \left((R^2k_p^2 - 3 - 3ik_pR)e^{-ik_pR} + (3 + 3ik_sR - R^2k_s^2)e^{-ik_sR} \right) \mathbf{e}_y \\
 & + \left((x^2 + y^2 - 2z^2)(e^{-ik_pR} - e^{-ik_sR}) + (z^2R^2k_p^2 + i(x^2 + y^2 - 2z^2)Rk_p) e^{-ik_pR} \right. \\
 & \left. \left. + ((x^2 + y^2)R^2k_s^2 - i(x^2 + y^2 - 2z^2)Rk_s) e^{-ik_sR} \right) \mathbf{e}_z \right]. \tag{42}
 \end{aligned}$$

In order to be consistent with the numerically calculated particle velocities \mathbf{v} , the particle displacement field \mathbf{u} obtained from equation (42) needs to be multiplied by a factor of $i\omega$, i.e.,

$$\mathbf{v} = i\omega\mathbf{u}, \tag{43}$$

according to the principle of Fourier transform.

In the numerical experiments, we keep the geometry settings given in the previous section. The P -wave velocity is 5000 m/s, the S -wave velocity is 2500 m/s, and the density is 1000 kg/m³. The frequencies are chosen from 10 Hz to 30 Hz with a step of 1 Hz. Similarly, the corresponding number of grid-points per minimum shear wavelength varies from 10 to 3.3. Five levels of attenuation $Q = 5, 10, 20, 200$ and ∞ are taken into account. A point force along the z -axis located in the center of the domain is used as the source to conform to the wave equation (40). As for the computational requirement for the linear solver CARP-CG, the memory consumed during the iterations is 2 GB, same as the acoustic case. The elapsed time, however, is around 956.82 seconds with 8 cores for a modeling at 10 Hz without attenuation, which is also the most time-consuming test among different frequencies and different levels of attenuation. The number of iterations of CARP-CG is 4628. In spite of using twice the number of cores, the elapsed time becomes longer than that in the acoustic case. This may be caused by the generation of both P -waves and S -waves, whose complexities could delay the convergence of the CARP-CG method (see the number of iterations for the acoustic and elastic cases).

Fig. 9 shows the 3D elastic wavefields of particle velocities v_x, v_y and v_z computed from the analytic solution, the conventional and the optimal scheme with $Q = 20$ at 30 Hz (i.e., 3.3 grid-points per minimum shear wavelength). The wavefields for each particle velocity component are plotted using the same color scale. As can be seen, the optimal wavefields are almost the same as the analytic solutions both in shape and amplitude, while the conventional ones fail to fit. Similarly as in the previous section, the profiles of these wavefields are presented to check the consistency between the analytic and the numerical results. Figs. 10, 11 and 12 present the profiles of particle velocity wavefields v_x, v_y and v_z for $z = 675$ m (left) and $x = 675$ m (right) in the x - z plane sliced at $y = 1000$ m, with $Q = 20$ at 10 Hz, 20 Hz, 25 Hz and 30 Hz. As shown in these figures, the profiles of both conventional and optimal results agree very well with the analytic solutions at 10 Hz. Deviation becomes more and more severe for the conventional results as the frequency increases, whereas it is not the case for the optimal ones. Taking the profiles at 30 Hz as example, the conventional profiles for $z = 675$ m has a smaller amplitude than the analytic solution and the wave cycles are shifted. For the profiles of $x = 675$ m, the values near the center deviate badly from the analytic ones. The optimal profiles, nevertheless, remain to fit the analytic solutions very well.

Profiles of v_z at 30 Hz with different levels of attenuation are presented in Fig. 13 as an example to better illustrate the attenuation effect. The figures of v_x and v_y are omitted here due to their similar attenuation effect as shown in Fig. 13. As can be seen, the amplitude decays more as Q decreases. The better performance of the optimal scheme can also be verified by comparing the profiles in this figure. The correspondence between the optimal profiles and the analytical solutions is excellent, while the conventional profiles depart significantly in shape.

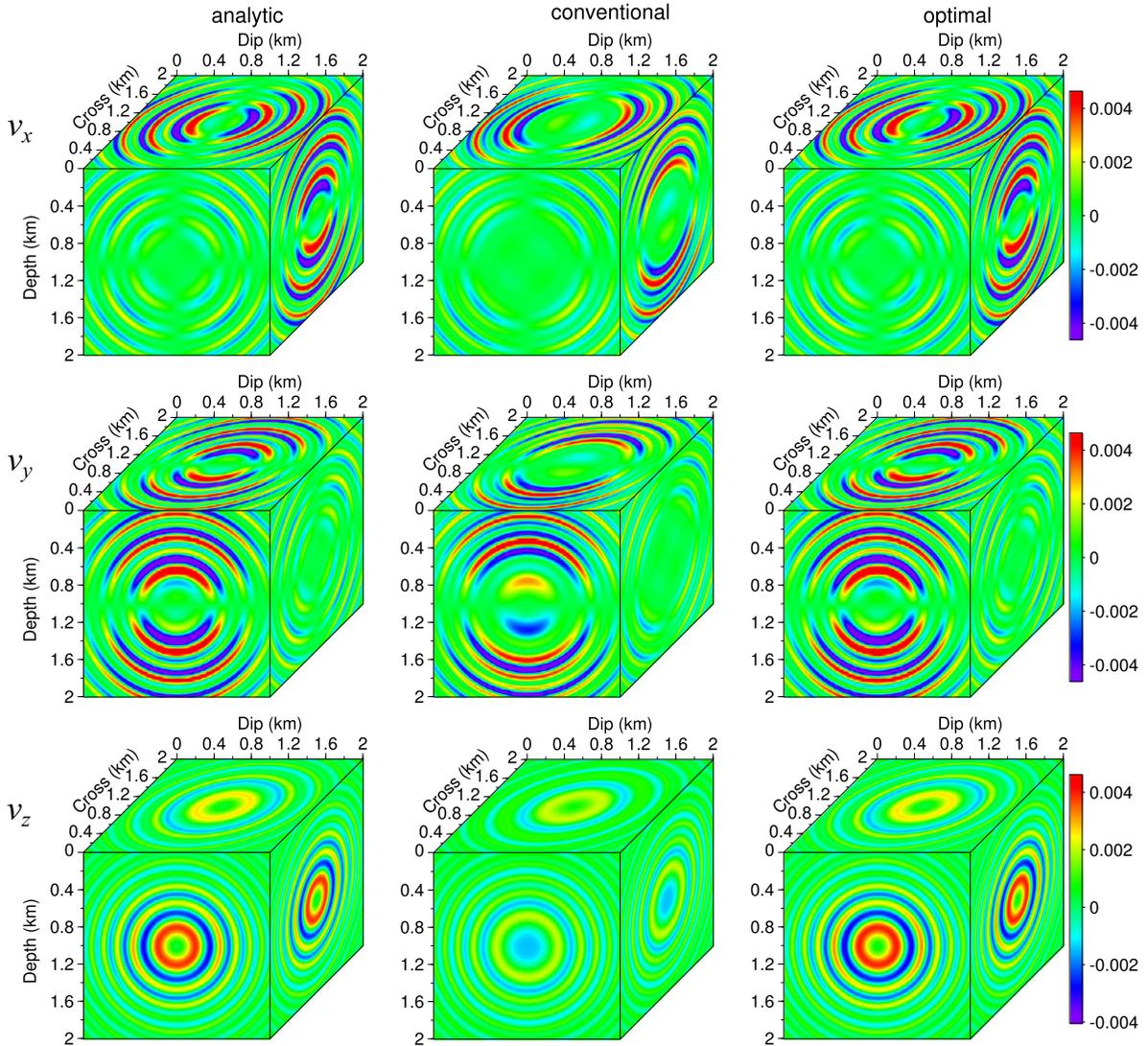


Fig. 9. 3D elastic wavefields of the particle velocity v_x (first row), v_y (second row) and v_z (third row) with $Q = 20$ at 30 Hz, computed from the analytic solution (first column), the conventional scheme (second column) and the optimal scheme (third column). The wavefields in each row are plotted using the same color scale shown on the right-most side. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For the quantitative comparison, the definition given in equation (39) are used for calculating the relative errors. Fig. 14 shows the relative errors of the particle velocity wavefields v_x (left), v_y (middle) and v_z (right) using the conventional and the optimal scheme. The errors are plotted as a function of the frequency with different levels of attenuation. The first x -axis (bottom) represents the frequency and the second x -axis (top) contains the corresponding value of K_s which is the inverse of the number of grid-points per minimum wavelength. In general, the same phenomenon can be observed as in the acoustic case: the error produced by the optimal results are smaller than the conventional ones and error of attenuated wavefields are smaller than those without attenuation. The error of the optimal results between 10 Hz and 30 Hz are very small which can be considered a natural consequence of the optimization on K_s in the range of [0, 0.33].

4. Conclusion

We have investigated an optimal four-order staggered-grid FD scheme for 3D frequency-domain viscoelastic wave modeling. An ALM strategy is implemented to minimize the grid dispersion and grid anisotropy. The optimal FD coefficients and the mass weighting coefficients are determined by optimizing the misfit between the normalized phase velocities and the unity. The Levenberg–Marquardt method and the simulated annealing algorithm are used to tackle the optimization problem. Inheriting from the properties of the staggered-grid scheme, the optimal scheme is stable for different values of Poisson’s ratio, and the grid dispersion and grid anisotropy are small and relatively insensitive to Poisson’s ratio. Dispersion

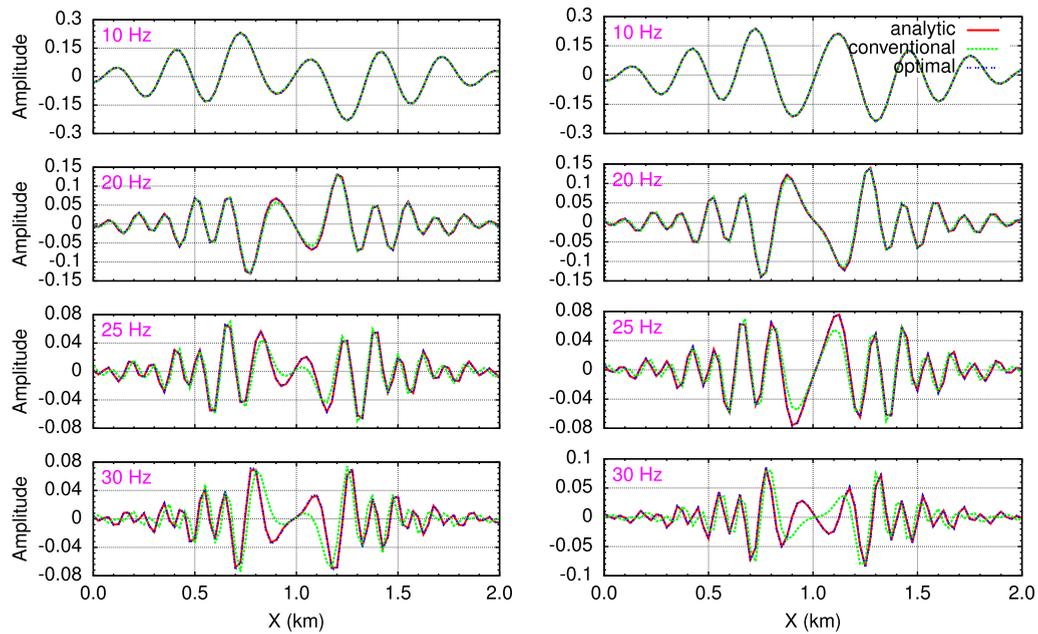


Fig. 10. Profiles of particle velocity wavefield v_x for $z = 675$ m (left) and $x = 675$ m (right) in the x - z plane sliced at $y = 1000$ m with $Q = 20$ at different frequencies.

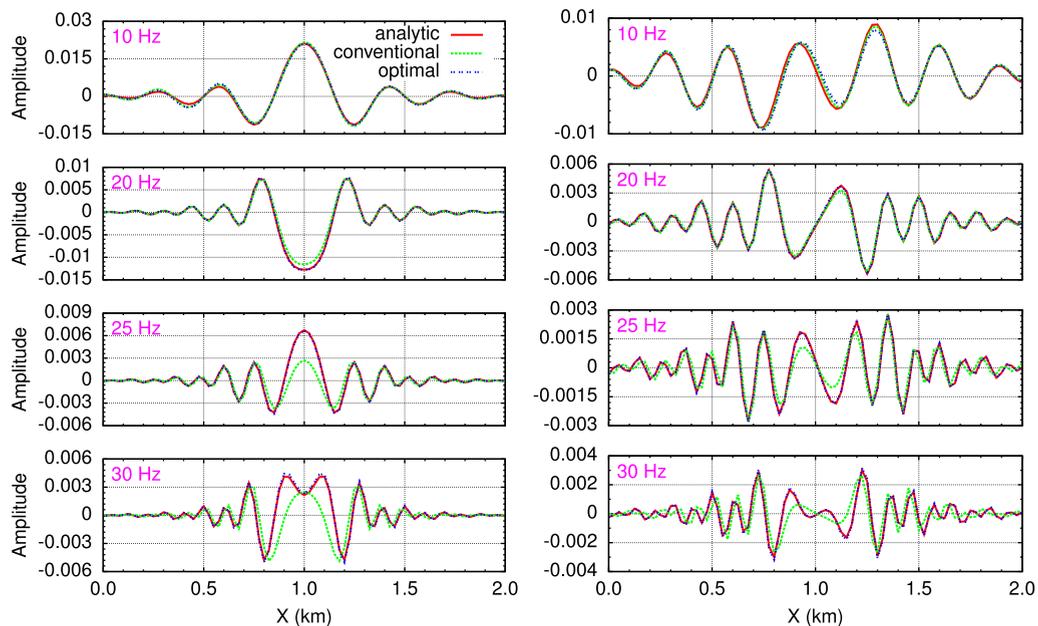


Fig. 11. Profiles of particle velocity wavefield v_y for $z = 675$ m (left) and $x = 675$ m (right) in the x - z plane sliced at $y = 1000$ m with $Q = 20$ at different frequencies.

analysis also shows that using the optimal coefficients allows a coarse sampling that requires only 3.7 grid-points per minimum shear wavelength under the error limit of 1%. Moreover, compared with using the conventional fourth-order scheme, the memory requirement reduces to a percentage of 5.37% for a direct solver and 11.2% for an iterative solver under the error limit of 1%, and these numbers change to 10.3% and 18.3% respectively if the error limit is 3%. Such a reduction would significantly relieve the computational burden for 3D viscoelastic wave modeling problems.

Validations of the numerical solutions are conducted using both the viscoacoustic and the viscoelastic analytic solutions with different levels of attenuation. The acoustic pressure wavefields are calculated within the elastic discretization framework by setting $V_s = 0$ and using an explosive source. For the elastic case, a point force along z -axis serves as the source to generate the particle velocity wavefields. Several 3D volume visualizations and 1D profiles of these wavefields show that the

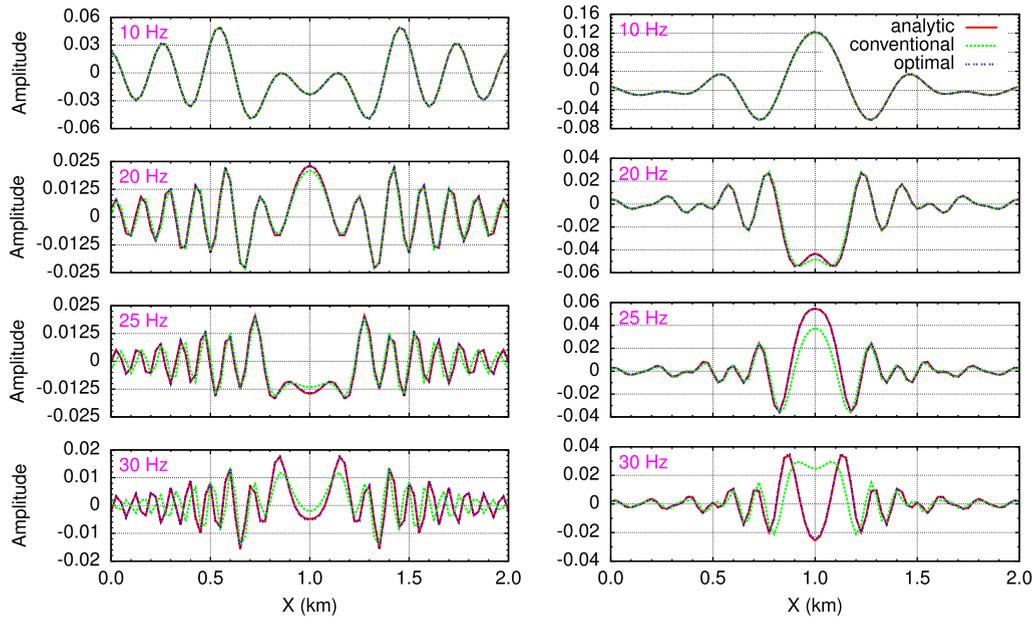


Fig. 12. Profiles of particle velocity wavefield v_z for $z = 675$ m (left) and $x = 675$ m (right) in the x - z plane sliced at $y = 1000$ m with $Q = 20$ at different frequencies.

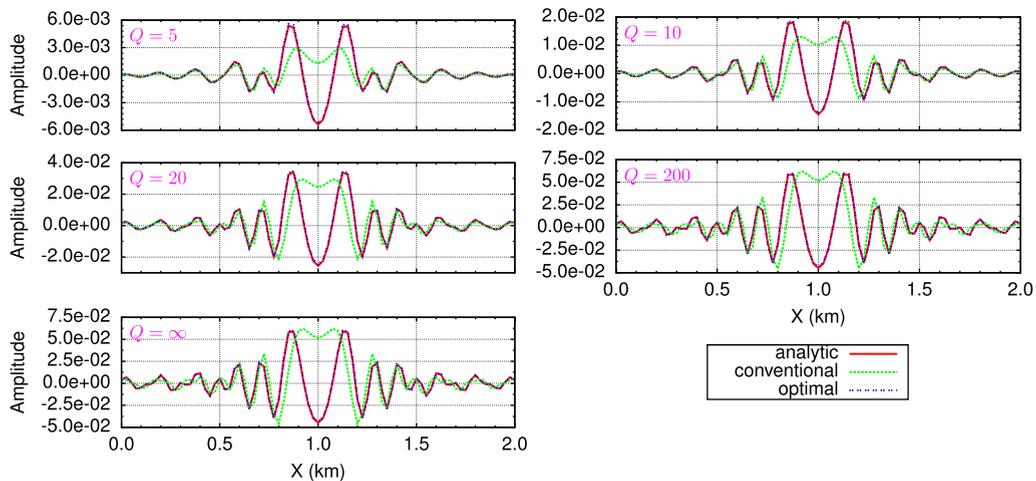


Fig. 13. Profiles of particle velocity wavefield v_z for $x = 675$ m in the x - z plane sliced at $y = 1000$ m at 30 Hz with different levels of attenuation.

optimal results conform better to the analytic solutions, even by a sampling with 2.5 grid-points per minimum wavelength. Quantitative comparison verifies the smaller error of the optimal results, especially for samplings finer than 3 grid-points per minimum wavelength, which is consistent with the optimization over K_s in the range of $[0, 0.33]$. We have observed that the errors of wavefields with strong attenuation are smaller than those with weak attenuation. This could be explained by the fact that with strong attenuation, the amplitude of the wavefields decays faster as the distance from the source point increases. The error hence diminishes due to the attenuation. The CARP-CG method used as the linear solver has shown satisfactory memory efficiency compared to the direct solvers and good converging ability for viscoelastic modeling.

Note that the optimal coefficients could also be applied in the time-domain, where the time step should be chosen according to the Courant–Friedrichs–Lewy stability condition. On the other hand, given the significant reduction in memory cost, this optimal scheme could be potentially incorporated into large scale memory-consuming seismic imaging applications, such as reverse time migration or full waveform inversion, and further study is thus lead.

Acknowledgements

This study was funded by the National Nature Science Foundation of China under Grant No. 91230119 and No. 41474102, and Chinese Scholarship Council. This study was also partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>),

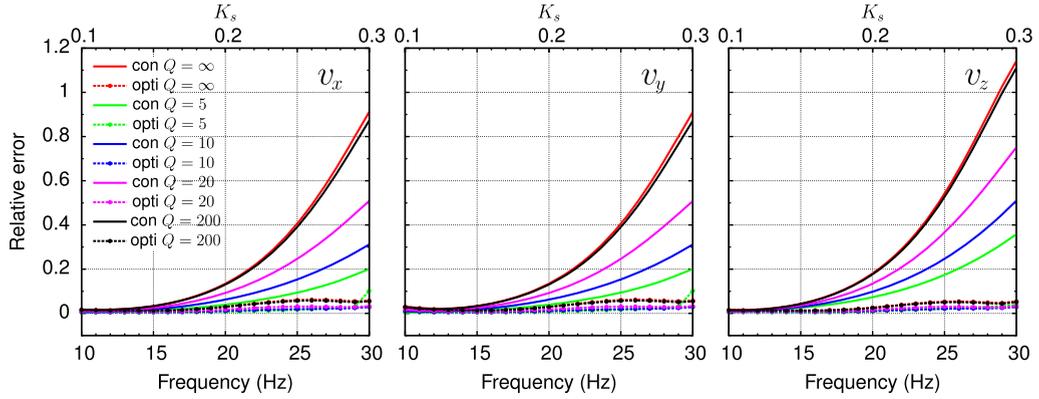


Fig. 14. Relative errors of the particle velocity wavefields v_x (left), v_y (middle) and v_z (right) using the conventional and the optimal scheme, plotted as a function of the frequency with different levels of attenuation. The first x-axis (bottom) represents the frequency and the second x-axis (top) contains the corresponding value of K_s , which is the inverse of the number of grid-points per minimum wavelength.

sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SCHLUMBERGER, SHELL, SINOPEC, STA-TOIL, TOTAL and WOODSIDE. The authors would like to acknowledge the associate editor and the reviewers for their instructive and detailed comments. The authors appreciate the fruitful discussions with Wei Zhou (ISTerre, Université Grenoble Alpes), Qinglong He and Yong Chen (Department of Mathematics, Harbin Institute of Technology).

Appendix A. Derivation of P_0 , P_1 and P_2

In order to elaborate on the derivation of the normalized phase velocities, the intermediate formulas for computing P_0 , P_1 and P_2 are presented, which is originally provided in [21]. Calculating the determinant of matrix (4) generates a polynomial of the third degree on the variable ω^2 which is

$$A + B(\omega^2) + C(\omega^2)^2 + (\omega^2)^3 = 0, \quad (\text{A.1})$$

where A , B , C are defined as

$$\begin{aligned} A &= (\Delta^2 \rho D_m)^{-3} (a_{3,0} \mu^3 + a_{2,1} \mu^2 \nu + a_{1,2} \mu \nu^2 + a_{0,3} \nu^3), \\ B &= (\Delta^2 \rho D_m)^{-2} (b_{2,0} \mu^2 + b_{1,1} \mu \nu + b_{0,2} \nu^2), \\ C &= (\Delta^2 \rho D_m)^{-1} (c_{1,0} \mu + c_{0,1} \nu). \end{aligned} \quad (\text{A.2})$$

The lowercase coefficients a , b , c are equations of the discrete differential operators

$$\begin{aligned} a_{3,0} &= D_{xx}(D_{yy}^2 + D_{zz}^2 - D_{zx}^2 - D_{xy}^2) \\ &\quad + D_{yy}(D_{xx}^2 + D_{zz}^2 - D_{yz}^2 - D_{xy}^2) \\ &\quad + D_{zz}(D_{xx}^2 + D_{yy}^2 - D_{yz}^2 - D_{zx}^2) \\ &\quad - 2D_{xy}D_{yz}D_{zx} + 2D_{xx}D_{yy}D_{zz}, \\ a_{2,1} &= D_{xx}^3 + D_{yy}^3 + D_{zz}^3 \\ &\quad + D_{xx}(D_{yy}^2 + D_{zz}^2 - D_{yz}^2 + 2D_{zx}^2 + 2D_{xy}^2) \\ &\quad + D_{yy}(D_{xx}^2 + D_{zz}^2 + 2D_{yz}^2 - D_{zx}^2 + 2D_{xy}^2) \\ &\quad + D_{zz}(D_{xx}^2 + D_{yy}^2 + 2D_{yz}^2 + 2D_{zx}^2 - D_{xy}^2) \\ &\quad + 6D_{xy}D_{yz}D_{zx} + 3D_{xx}D_{yy}D_{zz}, \\ a_{1,2} &= D_{xx}(D_{yy}^2 + D_{zz}^2 + 2D_{yz}^2 - D_{zx}^2 - D_{xy}^2) \\ &\quad + D_{yy}(D_{xx}^2 + D_{zz}^2 - D_{yz}^2 + 2D_{zx}^2 - D_{xy}^2) \\ &\quad + D_{zz}(D_{xx}^2 + D_{yy}^2 - D_{yz}^2 - D_{zx}^2 + 2D_{xy}^2) \\ &\quad - 6D_{xy}D_{yz}D_{zx}, \\ a_{0,3} &= -(D_{xx}D_{yz}^2 + D_{yy}D_{zx}^2 + D_{zz}D_{xy}^2) \\ &\quad + 2D_{xy}D_{yz}D_{zx} + D_{xx}D_{yy}D_{zz}, \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned}
 b_{2,0} &= (D_{xx}^2 + D_{yy}^2 + D_{zz}^2) - (D_{yz}^2 + D_{zx}^2 + D_{xy}^2) \\
 &\quad + 3(D_{xx}D_{yy} + D_{xx}D_{zz} + D_{yy}D_{zz}), \\
 b_{1,1} &= 2(D_{xx}^2 + D_{yy}^2 + D_{zz}^2) + 2(D_{yz}^2 + D_{zx}^2 + D_{xy}^2) \\
 &\quad + 2(D_{xx}D_{yy} + D_{xx}D_{zz} + D_{yy}D_{zz}), \\
 b_{0,2} &= - (D_{yz}^2 + D_{zx}^2 + D_{xy}^2) \\
 &\quad + (D_{xx}D_{yy} + D_{xx}D_{zz} + D_{yy}D_{zz}), \\
 c_{1,0} &= 2(D_{xx} + D_{yy} + D_{zz}), \\
 c_{0,1} &= D_{xx} + D_{yy} + D_{zz}.
 \end{aligned}
 \tag{A.4}$$

The cube roots of polynomial (A.1) are

$$\begin{aligned}
 \omega_p^2 &= -\frac{C}{3} + (S + T), \\
 \omega_{s1}^2 &= -\frac{C}{3} - \frac{1}{2}(S + T) + \frac{i\sqrt{3}}{2}(S - T), \\
 \omega_{s2}^2 &= -\frac{C}{3} - \frac{1}{2}(S + T) - \frac{i\sqrt{3}}{2}(S - T),
 \end{aligned}
 \tag{A.5}$$

where

$$\begin{aligned}
 S &= (R + \sqrt{R^2 + Q^3})^{\frac{1}{3}}, \\
 T &= (R - \sqrt{R^2 + Q^3})^{\frac{1}{3}}, \\
 R &= \frac{9BC - 2C^3 - 27A}{54}, \\
 Q &= \frac{3B - C^2}{9}.
 \end{aligned}
 \tag{A.6}$$

With these developments, we could recall equations (5) and (6) and give the expressions of P_0 , P_1 and P_2 in the following

$$\begin{aligned}
 P_0 &= c_{1,0}\mu + c_{0,1}\nu, \\
 P_1 &= q_{2,0}\mu^2 + q_{1,1}\mu\nu + q_{0,2}\nu^2, \\
 P_2 &= r_{3,0}\mu^3 + r_{2,1}\mu^2\nu + r_{1,2}\mu\nu^2 + r_{0,3}\nu^3,
 \end{aligned}
 \tag{A.7}$$

and

$$\begin{aligned}
 q_{2,0} &= -\frac{1}{9}c_{1,0}^2 + \frac{1}{3}b_{2,0}, \\
 q_{1,1} &= -\frac{2}{9}c_{0,1}c_{1,0} + \frac{1}{3}b_{1,1}, \\
 q_{0,2} &= -\frac{1}{9}c_{0,1}^2 + \frac{1}{3}b_{0,2}, \\
 r_{3,0} &= -\frac{1}{27}c_{1,0}^3 + \frac{1}{6}c_{1,0}b_{2,0} - \frac{1}{2}a_{3,0}, \\
 r_{2,1} &= -\frac{1}{9}c_{0,1}c_{1,0}^2 + \frac{1}{6}c_{0,1}b_{2,0} + \frac{1}{6}c_{1,0}b_{1,1} - \frac{1}{2}a_{2,1}, \\
 r_{1,2} &= -\frac{1}{9}c_{1,0}c_{0,1}^2 + \frac{1}{6}c_{1,0}b_{0,2} + \frac{1}{6}c_{0,1}b_{1,1} - \frac{1}{2}a_{1,2}, \\
 r_{0,3} &= -\frac{1}{27}c_{0,1}^3 + \frac{1}{6}c_{0,1}b_{0,2} - \frac{1}{2}a_{0,3}.
 \end{aligned}
 \tag{A.8}$$

Replacing the elastic moduli (μ, ν) by the Poisson's ratio σ in equation (8) yields new expressions of P_0 , P_1 and P_2 :

$$\begin{aligned}
 P_0 &= p_1^{(0)}\sigma + p_0^{(0)}, \\
 P_1 &= p_2^{(1)}\sigma^2 + p_1^{(1)}\sigma + p_0^{(1)}, \\
 P_2 &= p_3^{(2)}\sigma^3 + p_2^{(2)}\sigma^2 + p_1^{(2)}\sigma + p_0^{(2)},
 \end{aligned}
 \tag{A.9}$$

with

$$\begin{aligned}
 p_1^{(0)} &= -c_{0,1} - c_{1,0}, \\
 p_0^{(0)} &= c_{0,1} + \frac{1}{2}c_{1,0}, \\
 p_2^{(1)} &= q_{0,2} + q_{1,1} + q_{2,0}, \\
 p_1^{(1)} &= -2q_{0,2} - \frac{3}{2}q_{1,1} - q_{2,0}, \\
 p_0^{(1)} &= q_{0,2} + \frac{1}{2}q_{1,1} + \frac{1}{4}q_{2,0}, \\
 p_3^{(2)} &= -r_{0,3} - r_{1,2} - r_{2,1} - r_{3,0}, \\
 p_2^{(2)} &= 3r_{0,3} + \frac{5}{2}r_{1,2} + 2r_{2,1} + \frac{3}{2}r_{3,0}, \\
 p_1^{(2)} &= -3r_{0,3} - 2r_{1,2} + \frac{5}{4}r_{2,1} - \frac{3}{4}r_{3,0}, \\
 p_0^{(2)} &= r_{0,3} + \frac{1}{2}r_{1,2} + \frac{1}{4}r_{2,1} + \frac{1}{8}r_{3,0}.
 \end{aligned} \tag{A.10}$$

These equations enable us to formulate the discrete phase velocities (10).

Appendix B. Derivation of normalized group velocities

The discrete group velocities of P - and S -waves are defined as

$$V_p^{\text{gr}} = \frac{\partial \omega_p}{\partial k}, \quad V_{s1}^{\text{gr}} = \frac{\partial \omega_{s1}}{\partial k}, \quad V_{s2}^{\text{gr}} = \frac{\partial \omega_{s2}}{\partial k}. \tag{B.1}$$

Calculating the derivative of both sides in equations (5) with respect to k gives

$$2\omega_p \frac{\partial \omega_p}{\partial k} = \frac{1}{\rho \Delta^2} \left(-D_m^{-2} \frac{\partial D_m}{\partial k} \Phi + \frac{1}{D_m} \frac{\partial \Phi}{\partial k} \right), \tag{B.2}$$

$$2\omega_{s1} \frac{\partial \omega_{s1}}{\partial k} = \frac{1}{\rho \Delta^2} \left(-D_m^{-2} \frac{\partial D_m}{\partial k} \Psi + \frac{1}{D_m} \frac{\partial \Psi}{\partial k} \right), \tag{B.3}$$

$$2\omega_{s2} \frac{\partial \omega_{s2}}{\partial k} = \frac{1}{\rho \Delta^2} \left(-D_m^{-2} \frac{\partial D_m}{\partial k} \Theta + \frac{1}{D_m} \frac{\partial \Theta}{\partial k} \right), \tag{B.4}$$

where

$$\Phi = -\frac{P_0}{3} + P_4 + P_5, \tag{B.5}$$

$$\Psi = -\frac{P_0}{3} - \frac{P_4 + P_5}{2} + i\sqrt{3} \left(\frac{P_4 - P_5}{2} \right), \tag{B.6}$$

$$\Theta = -\frac{P_0}{3} - \frac{P_4 + P_5}{2} - i\sqrt{3} \left(\frac{P_4 - P_5}{2} \right). \tag{B.7}$$

Then we obtain the expressions of the normalized group velocities

$$\frac{V_p^{\text{gr}}}{V_p} = \frac{1}{2\Delta\sqrt{\nu\Phi/D_m}} \left(-D_m^{-2} \frac{\partial D_m}{\partial k} \Phi + \frac{1}{D_m} \frac{\partial \Phi}{\partial k} \right), \tag{B.8}$$

$$\frac{V_{s1}^{\text{gr}}}{V_{s1}} = \frac{1}{2\Delta\sqrt{\mu\Psi/D_m}} \left(-D_m^{-2} \frac{\partial D_m}{\partial k} \Psi + \frac{1}{D_m} \frac{\partial \Psi}{\partial k} \right), \tag{B.9}$$

$$\frac{V_{s2}^{\text{gr}}}{V_{s2}} = \frac{1}{2\Delta\sqrt{\mu\Theta/D_m}} \left(-D_m^{-2} \frac{\partial D_m}{\partial k} \Theta + \frac{1}{D_m} \frac{\partial \Theta}{\partial k} \right). \tag{B.10}$$

To be consistent with previous development, a variable change from (μ, ν) to σ should also be conducted. We omit the expressions of $\partial D_m/\partial k$, $\partial P_i/\partial k$, $i = 0, 4, 5$, which can be obtained from Appendix A, section 2.2 and equation (21) with proper calculations.

References

- [1] P. Amestoy, C. Ashcraft, O. Boiteau, A. Buttari, J.-Y. L'Excellent, C. Weisbecker, Improving multifrontal methods by means of block low-rank representations, *SIAM J. Sci. Comput.* 37 (3) (2015) A1451–A1474.
- [2] J.E. Anderson, L. Tan, D. Wang, Time-reversal checkpointing methods for RTM and FWI, *Geophysics* 77 (2012) S93–S103.
- [3] J.-P. Béranger, A perfectly matched layer for absorption of electromagnetic waves, *J. Comput. Phys.* 114 (1994) 185–200.
- [4] Å. Björck, T. Elfving, Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations, *BIT Numer. Math.* 19 (2) (1979) 145–163, <http://dx.doi.org/10.1007/BF01930845>.
- [5] N. Bleistein, *Mathematical Methods for Wave Phenomena*, Academic Press Inc., London, 1984.
- [6] T. Bohlen, Parallel 3-D viscoelastic finite-difference seismic modeling, *Comput. Geosci.* 28 (2002) 887–899.
- [7] R. Brossier, V. Etienne, S. Operto, J. Virieux, Frequency-domain numerical modelling of visco-acoustic waves based on finite-difference and finite-element discontinuous Galerkin methods, in: D.W. Dissanayake (Ed.), *Acoustic Waves, SCYO*, 2010, pp. 125–158.
- [8] R. Brossier, B. Pajot, L. Combe, S. Operto, L. Métivier, J. Virieux, Time and frequency-domain FWI implementations based on time solver: analysis of computational complexities, in: *Expanded Abstracts, 76th Annual EAGE Meeting*, 2014, We G102 05.
- [9] R. Brossier, J. Virieux, S. Operto, Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling, *Geophys. J. Int.* 175 (2) (2008) 541–559.
- [10] J.M. Carcione, Seismic modeling in viscoelastic media, *Geophysics* 58 (1) (1993) 110–120.
- [11] T.F. Chan, M.K. Ng, Galerkin projection methods for solving multiple linear systems, *SIAM J. Sci. Comput.* 21 (3) (1999) 836–850.
- [12] R. Clapp, Reverse time migration: saving the boundaries, Tech. Rep. SEP-136, Stanford Exploration Project, 2008.
- [13] R. Courant, K. Friedrichs, H. Lewy, Über die partiellen Differenzengleichungen der mathematischen Physik, *Math. Ann.* 100 (1928) 32–74.
- [14] A.C. Eringen, E.S. Suhubi, *Elastodynamics. Vol. 2, Linear Theory*, Academic Press, 1975.
- [15] Y.A. Erlangga, R. Nabben, On a multilevel Krylov method for the Helmholtz equation preconditioned by shifted Laplacian, *Electron. Trans. Numer. Anal.* 31 (2008) 403–424.
- [16] W. Futterman, Dispersive body waves, *J. Geophys. Res.* 67 (1962) 5279–5291.
- [17] L. Gao, R. Brossier, B. Pajot, J. Tago, J. Virieux, An immersed free-surface boundary treatment for seismic wave simulation, *Geophysics* 80 (5) (2015) T193–T209, <http://dx.doi.org/10.1190/geo2014-0609.1>.
- [18] D. Gordon, R. Gordon, Component-averaged row projections: a robust block-parallel scheme for sparse linear systems, *SIAM J. Sci. Comput.* 27 (3) (2005) 1092–1117.
- [19] D. Gordon, R. Gordon, CGMN revisited: robust and efficient solution of stiff linear systems derived from elliptic partial differential equations, *ACM Trans. Math. Softw.* 35 (3) (2008) 18:1–18:27.
- [20] D. Gordon, R. Gordon, CARP-CG: a robust and efficient parallel solver for linear systems, applied to strongly convection dominated PDEs, *Parallel Comput.* 36 (2010) 495–515.
- [21] B. Gosselin-Cliche, B. Giroux, 3D frequency-domain finite-difference viscoelastic-wave modeling using weighted average 27-point operators with optimal coefficients, *Geophysics* 79 (3) (2014) T169–T188, <http://dx.doi.org/10.1190/geo2013-0368.1>.
- [22] R. Graves, Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bull. Seismol. Soc. Am.* 86 (1996) 1091–1106.
- [23] A. Griewank, A. Walther, Algorithm 799: Revolve: an implementation of checkpointing for the reverse or adjoint mode of computational differentiation, *ACM Trans. Math. Softw.* 26 (2000) 19–45.
- [24] B. Hustedt, S. Operto, J. Virieux, Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling, *Geophys. J. Int.* 157 (2004) 1269–1296.
- [25] C.H. Jo, C. Shin, J.H. Suh, An optimal 9-point, finite-difference, frequency-space 2D scalar extrapolator, *Geophysics* 61 (1996) 529–537.
- [26] S. Kaczmarz, Angenäherte Auflösung von Systemen linearer Gleichungen, *Bull. Int. Acad. Pol. Sci. Lett.* 35 (1937) 355–357 (English translation by Jason Stockmann).
- [27] H. Kolsky, The propagation of stress pulses in viscoelastic solids, *Philos. Mag.* 1 (1956) 693–710.
- [28] D. Komatitsch, C. Barnes, J. Tromp, Wave propagation near a fluid–solid interface: a spectral element approach, *Geophysics* 65 (2) (2000) 623–631.
- [29] J. Kristek, P. Moczo, Seismic-wave propagation in viscoelastic media with material discontinuities: a 3D fourth-order staggered-grid finite-difference modeling, *Bull. Seismol. Soc. Am.* 93 (5) (2003) 2273–2280, <http://www.bssaonline.org/content/93/5/2273.abstract>.
- [30] A.R. Levander, Fourth-order finite-difference P-SV seismograms, *Geophysics* 53 (11) (1988) 1425–1436.
- [31] Y. Li, L. Métivier, R. Brossier, B. Han, J. Virieux, 2D and 3D frequency-domain elastic wave modeling in complex media with a parallel iterative solver, *Geophysics* 80 (3) (2015) T101–T118, <http://dx.doi.org/10.1190/geo2014-0480.1>.
- [32] V.V. Lisitsa, V. Tcheverda, A. Bakulin, V. Etienne, Combining discontinuous Galerkin with finite differences to simulate seismic waves in presence of free-surface topography, in: *Expanded Abstracts, EAGE*, 2014.
- [33] Y. Luo, G.T. Schuster, Parsimonious staggered grid finite-differencing of the wave equation, *Geophys. Res. Lett.* 17 (2) (1990) 155–158.
- [34] K. Marfurt, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations, *Geophysics* 49 (1984) 533–549.
- [35] D.J. Min, C. Sin, B.-D. Kwon, S. Chung, Improved frequency-domain elastic wave modeling using weighted-averaging difference operators, *Geophysics* 65 (2000) 884–895.
- [36] R. Mittet, Free-surface boundary conditions for elastic staggered-grid modeling schemes, *Geophysics* 67 (5) (2002) 1616–1623, <http://library.seg.org/doi/abs/10.1190/1.1512752>.
- [37] J. Moré, The Levenberg–Marquardt algorithm: implementation and theory, in: G. Watson (Ed.), *Numerical Analysis*, in: *Lecture Notes in Mathematics*, vol. 630, Springer, Berlin/Heidelberg, 1978, pp. 105–116.
- [38] MUMPS-team, MUMPS – MULTifrontal Massively Parallel Solver users' guide – version 4.10.0 (May 10, 2011). ENSEEIHT-ENS Lyon, <http://mumps.enseeiht.fr>, or <http://graal.ens-lyon.fr/MUMPS>, 2011.
- [39] K.T. Nihei, X. Li, Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD), *Geophys. J. Int.* 169 (2007) 1069–1078.
- [40] D.P. O'Leary, The block conjugate gradient algorithm and related methods, *Linear Algebra Appl.* 29 (1980) 293–322.
- [41] S. Operto, R. Brossier, L. Combe, L. Métivier, A. Ribodetti, J. Virieux, Computationally-efficient three-dimensional visco-acoustic finite-difference frequency-domain seismic modeling in vertical transversely isotropic media with sparse direct solver, *Geophysics* 79 (5) (2014) T257–T275.
- [42] S. Operto, A. Miniussi, R. Brossier, L. Combe, L. Métivier, A. Ribodetti, J. Virieux, 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, *Geophys. J. Int.* 202 (2) (2015) 1362–1391, <http://gji.oxfordjournals.org/content/202/2/1362.abstract>.
- [43] S. Operto, J. Virieux, P. Amestoy, J.-Y. L'Excellent, L. Giraud, H. Ben Hadj Ali, 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: a feasibility study, *Geophysics* 72 (5) (2007) SM195–SM211.
- [44] W.L. Piliant, *Elastic Waves in the Earth. Developments in Solid Earth Geophysics*, vol. 11, Elsevier, Amsterdam, 1979.
- [45] R.E. Plessix, A Helmholtz iterative solver for 3D seismic-imaging problems, *Geophysics* 72 (5) (2007) SM185–SM194.
- [46] R.E. Plessix, W.A. Mulder, Frequency domain finite difference amplitude preserving migration, *Geophys. J. Int.* 157 (2004) 975–987.

- [47] R.G. Pratt, Inverse theory applied to multi-source cross-hole tomography. Part II: elastic wave-equation method, *Geophys. Prospect.* 38 (1990) 311–330.
- [48] R.G. Pratt, Seismic waveform inversion in the frequency domain, part I: theory and verification in a physic scale model, *Geophysics* 64 (1999) 888–901.
- [49] R.G. Pratt, M.H. Worthington, Inverse theory applied to multi-source cross-hole tomography. Part I: acoustic wave-equation method, *Geophys. Prospect.* 38 (1990) 287–310.
- [50] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes*, in: *The Art of Scientific Computing*, 3rd edition, Cambridge University Press, New York, NY, USA, September 2007.
- [51] J. Pujol, The solution of nonlinear inverse problems and the Levenberg–Marquardt method, *Geophysics* 72 (4) (2007) W1–W16, <http://dx.doi.org/10.1190/1.2732552>.
- [52] C.D. Riyanti, A. Kononov, Y.A. Erlangga, C. Vuik, C. Oosterlee, R.E. Plessix, W.A. Mulder, A parallel multigrid-based preconditioner for the 3D heterogeneous high-frequency Helmholtz equation, *J. Comput. Phys.* 224 (2007) 431–448.
- [53] J. Robertsson, J. Blanch, W. Symes, Viscoelastic finite-difference modeling, *Geophysics* 59 (1994) 1444–1456.
- [54] J.O.A. Robertsson, A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography, *Geophysics* 61 (1996) 1921–1934.
- [55] V. Simoncini, S. Gallopoulos, A hybrid block GMRES method for nonsymmetric systems with multiple right-hand sides, *J. Comput. Appl. Math.* 36 (1–2) (1996) 457–469.
- [56] W.W. Symes, Reverse time migration with optimal checkpointing, *Geophysics* 72 (5) (2007) SM213–SM221, <http://link.aip.org/link/?GPY/72/SM213/1>.
- [57] E. Turkel, D. Gordon, R. Gordon, S. Tsynkov, Compact 2D and 3D sixth order schemes for the Helmholtz equation with variable wave number, *J. Comput. Phys.* 232 (2013) 272–287.
- [58] P. van Laarhoven, E. Aarts, *Simulated Annealing: Theory and Applications*, D. Reidel Publishing Co., 1987.
- [59] T. van Leeuwen, D. Gordon, R. Gordon, F. Herrmann, Preconditioning the Helmholtz equations via row projections, in: *Expanded Abstracts, EAGE, 2012*, p. A002.
- [60] J. Virieux, P-SV wave propagation in heterogeneous media, velocity-stress finite difference method, *Geophysics* 51 (1986) 889–901.
- [61] J. Virieux, V. Etienne, V. Cruz-Atienza, R. Brossier, E. Chaljub, O. Coutant, S. Garambois, D. Mercerat, V. Prieux, S. Operto, A. Ribodetti, J. Tago, Modelling seismic wave propagation for geophysical imaging, in: M. Kanao (Ed.), *Seismic Waves – Research and Analysis*, InTech, 2012, pp. 253–304.
- [62] J. Virieux, S. Operto, An overview of full waveform inversion in exploration geophysics, *Geophysics* 74 (6) (2009) WCC1–WCC26.
- [63] I. Štekl, R.G. Pratt, Accurate viscoelastic modeling by frequency-domain finite difference using rotated operators, *Geophysics* 63 (1998) 1779–1794.
- [64] D. Whitley, A genetic algorithm tutorial, *Stat. Comput.* 4 (2) (1994) 65–85, <http://dx.doi.org/10.1007/BF00175354>.