Geophysical Journal International

Geophys. J. Int. (2012) **191,** 346–366 doi: 10.1111/j.1365-246X.2012.05634.x

Massively parallel structured multifrontal solver for time-harmonic elastic waves in 3-D anisotropic media

Shen Wang,[1] Maarten V. de Hoop,[1] Jianlin Xia[1] and Xiaoye S. Li[2]

[1]*Department of Mathematics, Purdue University, IN 47907, USA. E-mail: cd_VictorWang@hotmail.com*
[2]*Computational Research Division, Lawrence Berkeley National Laboratory (LBNL), CA 94720, USA*

Accepted 2012 July 27. Received 2012 July 9; in original form 2012 March 29

SUMMARY

We present a massively parallel structured multifrontal solver for the equations describing time-harmonic elastic waves in 3-D anisotropic media. We use a multicomponent second-order finite-difference method. We extend the corresponding stencil to enhance the accuracy of the discretization without increasing the order. This accuracy is aligned with the tolerance level used for the Hierarchically SemiSeparable (HSS) low rank matrix compression underlying our solver. The interplay between the finite accuracy discretization and the finite accuracy matrix solver yields the key strategy which leads to the architecture of our algorithm. We analyse the relevant matrix structures, (numerically) estimate the rank of the dense matrices prior to the HSS compression and study the effect of anisotropy, and deduce the complexity and storage requirements of our algorithm.

Key words: Numerical solutions; Seismic anisotropy; Computational seismology; Wave propagation.

GJI Seismology

1 INTRODUCTION

We consider the modelling of time-harmonic elastic waves in anisotropic media. We use a multicomponent finite-difference method. In this paper, we are concerned with solving the resulting algebraic equations for a large number of different right-hand sides, that is, surface or subsurface sources, on a large domain in the context of modelling seismic wave propagation with applications in reverse time migration (RTM) based inverse scattering and local optimization based full waveform inversion (FWI) in mind. The key result is the development of a massively parallel structured direct solver accommodating anisotropy in the low- to mid-frequency range.

In smoothly varying media, we can diagonalize the system of equations describing elastic waves using pseudodifferential operators—assuming it is of principal type—thus decoupling the polarizations (e.g. see Stolk & de Hoop 2002). We can then consider time-harmonic solutions of scalar equations. This involves the application of techniques from microlocal analysis. In general, the polarized wave equations are scalar pseudodifferential equations of second order, the discretizations of which require particular techniques; wavelet bases provide a way to carry this out in principle (Alpert *et al.* 1993). Here, we are not concerned with such discretization techniques and instead solve the original coupled system of equations, with limited smoothness conditions. More importantly, we do not impose the restriction to systems of 'real principal type'. However, if the system were of principal type, we can use the mentioned diagonalizing operators to decouple the polarizations in the solution of the system. We will show an example of the effectiveness of such a procedure.

We use a second-order finite difference scheme for the discretization of the system of equations, together with an optimization technique which involves adding more points to the basic stencil, to minimize the numerical dispersion. In principle, such an optimization procedure can be carried out at each spatial point following the heterogeneity and changing anisotropy in the medium. Our dispersion analysis and numerical examples demonstrate that with only five gridpoints per shear wavelength, we can achieve at least four digits of accuracy. This accuracy is aligned with the tolerance level used for the Hierarchically SemiSeparable (HSS) low rank matrix compression underlying our solver. The interplay between the finite accuracy discretization and the finite accuracy matrix solver yields the key strategy which leads to the architecture of our algorithm.

The direct method of choice for solving the mentioned problem is the multifrontal factorization algorithm (Liu 1992). The central idea of the multifrontal algorithm is to reorganize the sparse factorization of the discretized matrix operator into a series of dense local factorizations; this algorithm is also used in the package of MUMPS (MUltifrontal Massively Parallel Solver, Agullo *et al.* 2008). The algorithm is used together with the method of nested dissection (George 1973) to obtain a nested hierarchical structure and generate a LU factorization from the bottom up to minimize fill ins. In nested dissection, separators are introduced to recursively divide the mesh into two disjoint subdomains.

Each separator consists of a small set of mesh points. The nested partitioning leads to a sequence of separators at different levels, which forms a binary tree. This tree is used in the multifrontal method to manage the factorization from the bottom up, level by level.

The development, here, for systems of equations is a generalization of the work of Wang *et al.* (2010, 2011a, 2012) concerning scalar equations including nested dissection with separators of variable thickness. We follow the approach developed by Xia *et al.* (2009, 2010) of integrating the multifrontal method with structured matrices. The fill-in blocks of the factorization appear to be highly compressible using the framework of HSS matrices. Compression is a critical component to reduce memory requirements and enables the solution of problems defined on large subsurface domains; the accuracy of the solution is controlled and can be limited in the applications considered. We analyse the relevant matrix structures, (numerically) estimate the rank of the dense matrices prior to the HSS compression and study the effect of anisotropy, and deduce the complexity and storage requirements of our algorithm; we then compare these with the Helmholtz equation for polarized waves in isotropic media.

The solver developed, here, opens up the way to speed up significantly adjoint state computations in seismic applications for the purpose of FWI with a large number of events (Tromp *et al.* 2008). A strategy for time-harmonic FWI making use of a multifrontal solver for scalar waves was developed by Operto *et al.* (2007). Our current algorithm has been developed on a Cartesian grid, but the modifications to spherical sections is straightforward (Tromp *et al.* 2008; Fichtner *et al.* 2009).

We compare the accuracy of our algorithm with time-domain Discontinous Galerkin (DG) method. We mention related and alternative developments. For isotropic media, (Pratt 1990) developed a finite-difference method for modelling time-harmonic elastic waves. (Gauzellino *et al.* 2001) designed a non-conforming finite-element discretization emphasizing viscoelastic rheology, (see also Jr. Douglas *et al.* 1994). In fact, our solver can be adapted to a finite element method straightforwardly. (Airaksinen *et al.* 2009) developed an iterative solver based on an algebraic multigrid method and FEM in the isotropic case with a damping preconditioner. For a time-domain counterpart, we mention the work of Bansal & Sen (2008). For the high-frequency scattering of elastic waves, we refer to the work of El Kacimi & Laghrouche (2011) based on PUFEM involving wavelet based *ILU* preconditioners.

The outline of the paper is as follows. In the next section, we summarize the relevant equations, the finite-difference stencil used, the PMLs, and then introduce the system of algebraic equations. In Section 3, we discuss the modifications of our structured multifrontal solver from scalar equations to the elastic system. In particular, we give the matrix structure under nested dissection and a complexity analysis, and provide estimates on memory requirements. In Section 4, we give numerical estimates of the rank of the dense matrix prior to HSS compression. We also present the performance of our solver for a model problem. In Section 5, we present various numerical experiments: (i) multifrequency shear wave splitting and a comparison with a time-domain discontinuous Galerkin method, (ii) the formation of caustics in qSV-wave constituents , (iii) the presence of conical points and a comparison with a spectral element method, (iv) the focusing and defocusing of displacement in a strongly heterogeneous VTI medium and (v) the polarization decomposition of the solution. We end with some conclusions.

2 PROPAGATION OF TIME-HARMONIC ELASTIC WAVES

2.1 The system of partial differential equations

We write $\mathbf{x} = (x_1, x_2, x_3)$. Here 1, 2 and 3 denote the x, y and z spatial directions in 3-D, respectively. We consider the displacement formulation of the system describing time-harmonic elastic waves, with full anisotropy

$$-\frac{\partial \sigma_{ij}}{\partial x_j} - \rho\omega^2 u_i = f_i, \quad i, j = 1, 2, 3; \tag{1}$$

$\mathbf{f} = (f_1, f_2, f_3)$ is the forcing term, $\mathbf{u} = (u_1, u_2, u_3)$ is the displacement vector, ρ is the density which depends on x, and σ is the stress tensor.The constitutive relation between the stress and the strain is given by

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ * & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ * & * & C_{33} & C_{34} & C_{35} & C_{36} \\ * & * & * & C_{44} & C_{45} & C_{46} \\ * & * & * & * & C_{55} & C_{56} \\ * & * & * & * & * & C_{66} \end{pmatrix} \begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} \\ \dfrac{\partial u_2}{\partial x_2} \\ \dfrac{\partial u_3}{\partial x_3} \\ \dfrac{\partial u_2}{\partial x_3} + \dfrac{\partial u_3}{\partial x_2} \\ \dfrac{\partial u_3}{\partial x_1} + \dfrac{\partial u_1}{\partial x_3} \\ \dfrac{\partial u_1}{\partial x_2} + \dfrac{\partial u_2}{\partial x_1} \end{pmatrix}, \tag{2}$$

where $\mathbf{C} = \left[C_{ij}\right], i = 1, 2, \ldots, 6, j = 1, 2, \ldots, 6$ is the stiffness tensor flattened on a matrix; $*$ indicates the symmetry of $\mathbf{C}$. If we consider the orthorhombic anisotropy, 21 stiffness moduli reduce to nine independent ones. Then we have

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ * & C_{22} & C_{23} & 0 & 0 & 0 \\ * & * & C_{33} & 0 & 0 & 0 \\ * & * & * & C_{44} & 0 & 0 \\ * & * & * & * & C_{55} & 0 \\ * & * & * & * & * & C_{66} \end{pmatrix} \begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} \\ \dfrac{\partial u_2}{\partial x_2} \\ \dfrac{\partial u_3}{\partial x_3} \\ \dfrac{\partial u_2}{\partial x_3} + \dfrac{\partial u_3}{\partial x_2} \\ \dfrac{\partial u_3}{\partial x_1} + \dfrac{\partial u_1}{\partial x_3} \\ \dfrac{\partial u_1}{\partial x_2} + \dfrac{\partial u_2}{\partial x_1} \end{pmatrix}, \tag{3}$$

where

$$C_{33} = \rho V_p^2, \qquad C_{22} = C_{33}(1 + 2\epsilon^{(1)}), \qquad C_{11} = C_{33}(1 + 2\epsilon^{(2)}),$$

$$C_{55} = \rho V_s^2, \qquad C_{66} = C_{55}(1 + 2\gamma^{(1)}), \qquad C_{44} = \frac{C_{66}}{1 + 2\gamma^{(2)}},$$

$$C_{23} = (C_{33} - C_{44})\sqrt{1 + \frac{2\delta^{(1)}}{1 - C_{44}/C_{33}}} - C_{44},$$

$$C_{13} = (C_{33} - C_{55})\sqrt{1 + \frac{2\delta^{(2)}}{1 - C_{55}/C_{33}}} - C_{55},$$

$$C_{12} = (C_{11} - C_{66})\sqrt{1 + \frac{2\delta^{(3)}}{1 - C_{66}/C_{11}}} - C_{66}.$$

Here V_p represents the 'vertical' P-wave velocity, and V_s is the 'vertical' S-wave velocity and $\epsilon^{(1)}$, $\epsilon^{(2)}$, $\delta^{(1)}$, $\delta^{(2)}$, $\delta^{(3)}$, $\gamma^{(1)}$ and $\gamma^{(2)}$ are the extended Thomsen's parameters for orthorhombic media introduced by Tsvankin (1997). We note that if $\epsilon = \epsilon^{(1)} = \epsilon^{(2)}$, $\delta = \delta^{(1)} = \delta^{(2)} = \delta^{(3)}$ and $\gamma = \gamma^{(1)} = \gamma^{(2)}$, the orthorhombic anisotropy expressed by eq. (3) reduces to Transverse Isotropy (TI) anisotropy; ϵ, δ and γ are conventional Thomsen's parameters introduced by Thomsen (1986).

After substituting eq. (2) into eq. (1), we obtain the following coupled system of equations,

$$\left[\mathbf{A}(\mathbf{x}, \partial_{\mathbf{x}}, \omega) - \rho\omega^2\mathbf{I}\right]\mathbf{u}(\mathbf{x}, \omega) = \mathbf{f}(\mathbf{x}, \omega) \tag{4}$$

or

$$\begin{pmatrix} A_{11} - \rho\omega^2 & A_{12} & A_{13} \\ A_{21} & A_{22} - \rho\omega^2 & A_{23} \\ A_{31} & A_{32} & A_{33} - \rho\omega^2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}, \tag{5}$$

in which each element $A_{ij} (i = 1, 2, 3, j = 1, 2, 3)$ of $\mathbf{A}(\mathbf{x}, \partial_{\mathbf{x}}, \omega)$ is a second-order partial differential operator. For example,

$$\begin{aligned} A_{11} = & -\frac{\partial}{\partial x_1}\left(C_{11}\frac{\partial}{\partial x_1}\right) - \frac{\partial}{\partial x_1}\left(C_{16}\frac{\partial}{\partial x_2}\right) - \frac{\partial}{\partial x_1}\left(C_{15}\frac{\partial}{\partial x_3}\right) \\ & -\frac{\partial}{\partial x_2}\left(C_{16}\frac{\partial}{\partial x_1}\right) - \frac{\partial}{\partial x_2}\left(C_{66}\frac{\partial}{\partial x_2}\right) - \frac{\partial}{\partial x_2}\left(C_{56}\frac{\partial}{\partial x_3}\right) \\ & -\frac{\partial}{\partial x_3}\left(C_{15}\frac{\partial}{\partial x_1}\right) - \frac{\partial}{\partial x_3}\left(C_{56}\frac{\partial}{\partial x_2}\right) - \frac{\partial}{\partial x_3}\left(C_{55}\frac{\partial}{\partial x_3}\right), \end{aligned}$$

which reduces to

$$A_{11} = -\frac{\partial}{\partial x_1}\left(C_{11}\frac{\partial}{\partial x_1}\right) - \frac{\partial}{\partial x_2}\left(C_{66}\frac{\partial}{\partial x_2}\right) - \frac{\partial}{\partial x_3}\left(C_{55}\frac{\partial}{\partial x_3}\right)$$

in the orthorhombic case.

2.2 The system of algebraic equations

As in the work of Operto *et al.* (2007), we use a 27-point finite difference stencil, the structure of which is illustrated in Fig. 1, together with the mass lumping technique and convolutional PML boundary conditions, to discretize eq. (5) in orthorhombic media. The resulting

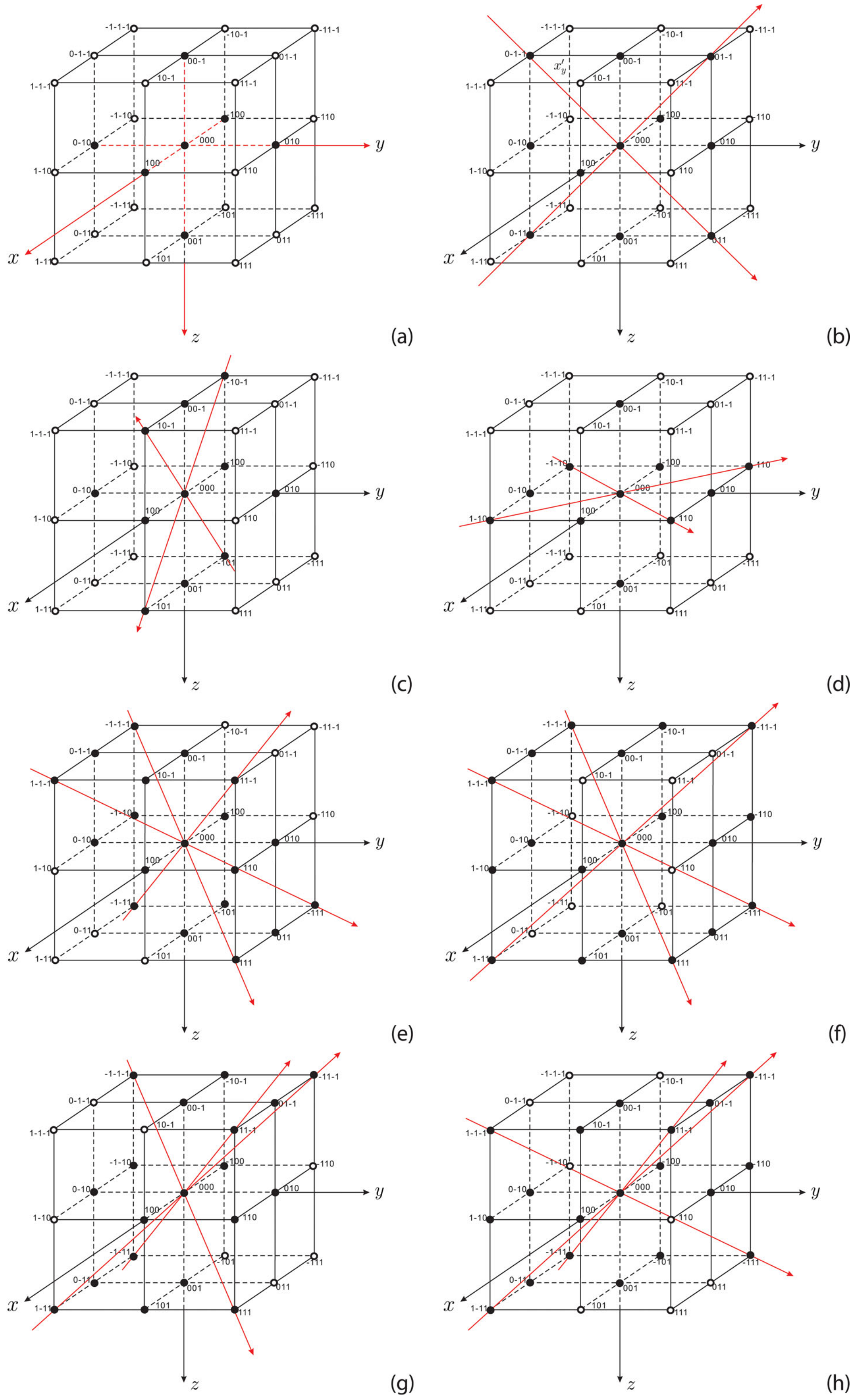


Figure 1. Groupings of points (eq. 6) in the 27-point finite difference stencil; solid balls indicate points that are included in the component stencil and circles indicate those that are not included; (a): $\mathbf{A}^{(c)}$; (b): $\mathbf{A}^{(x)}$; (c): $\mathbf{A}^{(y)}$; (d): $\mathbf{A}^{(z)}$; (e): $\mathbf{A}^{(1)}$; (f): $\mathbf{A}^{(2)}$; (g): $\mathbf{A}^{(3)}$; (h): $\mathbf{A}^{(4)}$.

discretization of $\mathbf{A}(\mathbf{x}, \partial_\mathbf{x}, \omega)$ is denoted as $\widehat{\mathbf{A}}(\mathbf{x}, \omega)$. The motivation of using a 27-point stencil is to improve the shape of the slowness surface associated with the finite difference approximation with additional degrees of freedom within the orthorhombic symmetry. This leads to the following grouping of gridpoints

$$\widehat{\mathbf{A}} = w_{s1}\mathbf{A}^{(c)} + \frac{w_{s2}}{3}\left(\mathbf{A}^{(x)} + \mathbf{A}^{(y)} + \mathbf{A}^{(z)}\right) + \frac{w_{s3}}{4}\left(\mathbf{A}^{(1)} + \mathbf{A}^{(2)} + \mathbf{A}^{(3)} + \mathbf{A}^{(4)}\right),$$

$$w_{s1} + w_{s2} + w_{s3} = 1, \tag{6}$$

where we optimize for w_{s1}, w_{s2} and w_{s3}; $\mathbf{A}^{(c)}$, $\mathbf{A}^{(x)}$, $\mathbf{A}^{(y)}$, $\mathbf{A}^{(z)}$, $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$, $\mathbf{A}^{(3)}$ and $\mathbf{A}^{(4)}$ are different groupings of gridpoints based on the orthorhombic symmetry, illustrated in Fig. 1. In Appendix A, we give an example of the complete evaluation of one entry $\widehat{A}_{11}(\mathbf{x}, \omega)$ for orthorhombic media.

To evaluate the diagonal term $\rho\omega^2\mathbf{u}$, we define the mass lumping operator $\mathbf{M}$,

$$\begin{aligned}&\mathbf{M}\mathbf{u} = w_{m1}\,[\rho\mathbf{u}]_1 + \frac{w_{m2}}{6}\,[\rho\mathbf{u}]_2 + \frac{w_{m3}}{12}\,[\rho\mathbf{u}]_3 + \frac{w_{m4}}{8}\,[\rho\mathbf{u}]_4\,,\\ &w_{m1} + w_{m2} + w_{m3} + w_{m4} = 1,\end{aligned} \tag{7}$$

where

$$\begin{aligned}
[\rho\mathbf{u}]_1 &= \rho_{000}\mathbf{u}_{000},\\
[\rho\mathbf{u}]_2 &= \rho_{-100}\mathbf{u}_{-100} + \rho_{100}\mathbf{u}_{100} + \rho_{0-10}\mathbf{u}_{0-10} + \rho_{010}\mathbf{u}_{010} + \rho_{00-1}\mathbf{u}_{00-1} + \rho_{001}\mathbf{u}_{001},\\
[\rho\mathbf{u}]_3 &= \rho_{-1-10}\mathbf{u}_{-1-10} + \rho_{1-10}\mathbf{u}_{1-10} + \rho_{-110}\mathbf{u}_{-110} + \rho_{110}\mathbf{u}_{110} + \rho_{0-1-1}\mathbf{u}_{0-1-1} + \rho_{01-1}\mathbf{u}_{01-1}\\
&\quad + \rho_{0-11}\mathbf{u}_{0-11} + \rho_{011}\mathbf{u}_{011} + \rho_{-10-1}\mathbf{u}_{-10-1} + \rho_{10-1}\mathbf{u}_{10-1} + \rho_{-101}\mathbf{u}_{-101} + \rho_{101}\mathbf{u}_{101},\\
[\rho\mathbf{u}]_4 &= \rho_{-1-1-1}\mathbf{u}_{-1-1-1} + \rho_{111}\mathbf{u}_{111} + \rho_{-111}\mathbf{u}_{-111} + \rho_{1-11}\mathbf{u}_{1-11}\\
&\quad + \rho_{11-1}\mathbf{u}_{11-1} + \rho_{-1-11}\mathbf{u}_{-1-11} + \rho_{1-1-1}\mathbf{u}_{1-1-1} + \rho_{-11-1}\mathbf{u}_{-11-1}.
\end{aligned}$$

Here, the subscripts of ρ and $\mathbf{u}$ represent different points in the 27-point stencil illustrated in Fig. 1. w_{m1}, w_{m2}, w_{m3} and w_{m4} are weights to be determined. We summarize the discrete system of eq. (5) into the following equation

$$\left[\widehat{\mathbf{A}}(\mathbf{x}, \omega) - \omega^2\mathbf{M}(\mathbf{x})\right]\mathbf{u}(., \omega) = \mathbf{f}(\mathbf{x}, \omega). \tag{8}$$

The conversion from subscripts to a linear index is chosen to be ($D = 3$)

$$\begin{aligned}&\mathbf{u}_{(k-1)N_1N_2 3+(j-1)N_1 3+(i-1)3+d} = u_d(x_{1,i}, x_{2,j}, x_{3,k}, \omega),\\ &d = 1, \ldots, 3,\ i = 1, \ldots, N_1,\ j = 1, \ldots,\ N_2,\ k = 1, \ldots, N_3;\end{aligned}$$

in which N_1, N_2 and N_3 denote the number of gridpoints in x_1, x_2 and x_3 directions, respectively. In analogy, upon sampling, the body force takes the form

$$\begin{aligned}&\mathbf{f}_{(k-1)N_1N_2 3+(j-1)N_1 3+(i-1)3+d} = f_d(x_{1,i}, x_{2,j}, x_{3,k}, \omega),\\ &d = 1, \ldots, 3,\ i = 1, \ldots, N_1,\ j = 1, \ldots, N_2,\ k = 1, \ldots, N_3.\end{aligned}$$

We end up with the a linear system of equations,

$$\mathbf{A}(\omega)\,\mathbf{u}(\omega) = \mathbf{f}(\omega), \tag{9}$$

where, by abuse of notation, the implied matrix $\mathbf{A}(\omega)$ contains the contributions both from $\widehat{\mathbf{A}}$ and $-\omega^2\mathbf{M}$. We note that $\mathbf{A}(\omega)$ is of size $(3N_1N_2N_3) \times (3N_1N_2N_3)$, and shares the same non-zero pattern for different ω. The matrix is pattern symmetric, non-Hermitian, indefinite and ill-conditioned. We set $n = 3N_1N_2N_3$. The seismic sources $\mathbf{f}(\omega)$ we consider are exploding point sources and body forces. We use a sharp Gaussian function as a regularized point source.

2.3 Numerical dispersion analysis

We present a classical numerical dispersion analysis of our finite difference discretization, and show that with only five gridpoints per shear wavelength, we can achieve reasonable accuracy. We follow the work of Holberg (1987), van Stralen *et al.* (1998), Štekl & Pratt (1998) and Operto *et al.* (2007). We consider a Fourier component, $\mathbf{u} = \mathbf{u}_0 \exp(-\mathrm{i}\mathbf{k}\cdot\mathbf{x})$, where $\mathbf{k}$ is a wave vector and $\mathbf{u}_0$ is a polarization vector. The phase velocity is given by $V_{\mathrm{ph}} = \omega/|\mathbf{k}|$. In polar coordinates, we write

$$\mathbf{k}\cdot\mathbf{x} = \frac{2\pi}{G}\left(r\cos\theta\cos\varphi + s\cos\theta\sin\varphi + t\sin\theta\right), \tag{10}$$

where (r, s, t) are the spatial coordinates and G signifies the number of gridpoints per wavelength. We cast our dispersion analysis in the framework of 3-D orthorhombic media

$$\det\left[\frac{1}{|\mathbf{k}|^2}\exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})\mathbf{M}^{-1}\widehat{\mathbf{A}}\exp(-\mathrm{i}\mathbf{k}\cdot\mathbf{x}) - V_{\mathrm{ph}}^2\mathbf{I}\right] = 0. \tag{11}$$

We write $\mathbf{B} = \frac{1}{|\mathbf{k}|^2}\exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})\mathbf{M}^{-1}\widehat{\mathbf{A}}\exp(-\mathrm{i}\mathbf{k}\cdot\mathbf{x})$, with elements

$$
\begin{aligned}
B_{11} &= \frac{C_{11}E_{xx} + C_{66}E_{yy} + C_{55}E_{zz}}{\rho J\left(\frac{2\pi}{G}\right)^2},\\
B_{22} &= \frac{C_{66}E_{xx} + C_{22}E_{yy} + C_{44}E_{zz}}{\rho J\left(\frac{2\pi}{G}\right)^2},\\
B_{33} &= \frac{C_{55}E_{xx} + C_{44}E_{yy} + C_{33}E_{zz}}{\rho J\left(\frac{2\pi}{G}\right)^2},\\
B_{12} &= B_{21} = \frac{(C_{12} + C_{66})\,E_{xy}}{\rho J\left(\frac{2\pi}{G}\right)^2},\\
B_{23} &= B_{32} = \frac{(C_{23} + C_{44})\,E_{yz}}{\rho J\left(\frac{2\pi}{G}\right)^2},\\
B_{13} &= B_{31} = \frac{(C_{13} + C_{55})\,E_{zx}}{\rho J\left(\frac{2\pi}{G}\right)^2},
\end{aligned}
$$

where

$$
\begin{aligned}
E_{xx} &= 2(1-\cos a)\left[w_{s1} + \frac{4+\cos b+\cos c}{6}w_{s2} + \frac{1+\cos b\cos c}{2}w_{s3}\right],\\
E_{yy} &= 2(1-\cos b)\left[w_{s1} + \frac{4+\cos c+\cos a}{6}w_{s2} + \frac{1+\cos c\cos a}{2}w_{s3}\right],\\
E_{zz} &= 2(1-\cos c)\left[w_{s1} + \frac{4+\cos a+\cos b}{6}w_{s2} + \frac{1+\cos a\cos b}{2}w_{s3}\right],\\
E_{xy} &= \sin a\sin b\left[w_{s1} + \frac{1+2\cos c}{3}w_{s2} + (2-\cos a\cos b\cos c)w_{s3}\right],\\
E_{yz} &= \sin b\sin c\left[w_{s1} + \frac{1+2\cos a}{3}w_{s2} + (2-\cos a\cos b\cos c)w_{s3}\right],\\
E_{zx} &= \sin c\sin a\left[w_{s1} + \frac{1+2\cos b}{3}w_{s2} + (2-\cos a\cos b\cos c)w_{s3}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
a &= \frac{2\pi}{G}\cos\theta\cos\varphi, \quad b = \frac{2\pi}{G}\cos\theta\sin\varphi, \quad c = \frac{2\pi}{G}\sin\theta,\\
\alpha &= \cos a + \cos b + \cos c,\\
\beta &= \cos a\cos b + \cos b\cos c + \cos c\cos a,\\
\eta &= \cos a\cos b\cos c,\\
J &= w_{m1} + \frac{w_{m2}}{3}\alpha + \frac{w_{m3}}{3}\beta + w_{m4}\eta.
\end{aligned}
$$

We compare the solutions with the solutions of

$$
\det\left[\frac{\mathbf{A}(\mathbf{x},\mathbf{k},\omega)}{\rho|\mathbf{k}|^2} - \widetilde{V}_{ph}^2\mathbf{I}\right] = 0, \tag{12}
$$

where $\widetilde{V}_{\mathrm{ph}}$ denotes exact phase velocities, and

$$
\begin{aligned}
\left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{11} &= \frac{C_{11}}{\rho}(\cos\theta\cos\varphi)^2 + \frac{C_{66}}{\rho}(\cos\theta\sin\varphi)^2 + \frac{C_{55}}{\rho}(\sin\theta)^2,\\
\left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{22} &= \frac{C_{66}}{\rho}(\cos\theta\cos\varphi)^2 + \frac{C_{22}}{\rho}(\cos\theta\sin\varphi)^2 + \frac{C_{44}}{\rho}(\sin\theta)^2,\\
\left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{33} &= \frac{C_{55}}{\rho}(\cos\theta\cos\varphi)^2 + \frac{C_{44}}{\rho}(\cos\theta\sin\varphi)^2 + \frac{C_{33}}{\rho}(\sin\theta)^2,\\
\left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{12} &= \left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{21} = \frac{C_{12}+C_{66}}{\rho}\cos^2\theta\sin\varphi\cos\varphi,\\
\left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{23} &= \left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{32} = \frac{C_{23}+C_{44}}{\rho}\sin\theta\cos\theta\sin\varphi,\\
\left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{13} &= \left(\rho^{-1}|\mathbf{k}|^{-2}\mathbf{A}(\mathbf{x},\mathbf{k},\omega)\right)_{31} = \frac{C_{13}+C_{55}}{\rho}\sin\theta\cos\theta\cos\varphi.
\end{aligned}
$$

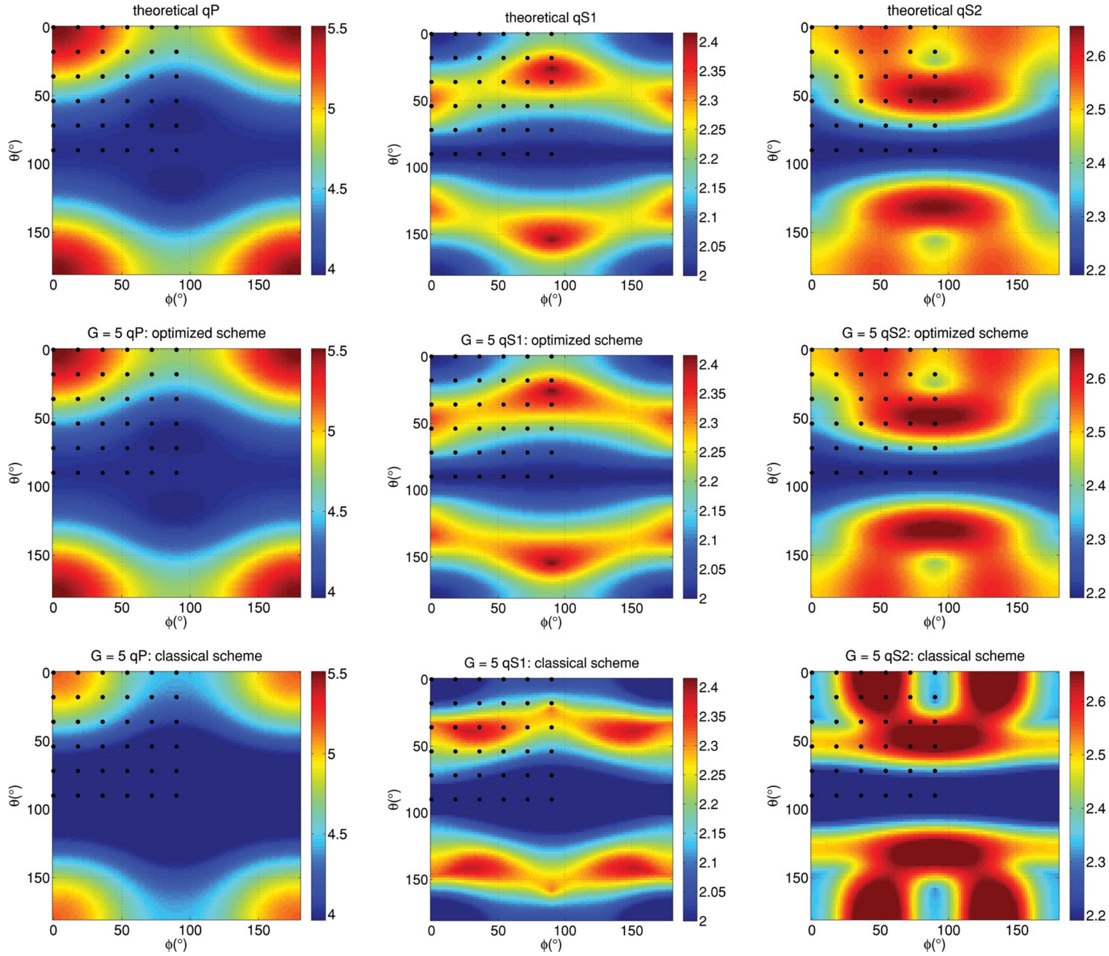


Figure 2. Phase velocities for the orthorhombic model: $V_p = 4000\,\mathrm{m\,s^{-1}}$, $V_s = 2000\,\mathrm{m\,s^{-1}}$, $\rho = 1.0$, $\epsilon^{(1)} = 0.2$, $\epsilon^{(2)} = 0.45$, $\delta^{(1)} = -0.1$, $\delta^{(2)} = 0.2$, $\delta^{(3)} = -0.15$, $\gamma^{(1)} = 0.28$, $\gamma^{(2)} = 0.15$; row one: theoretical phase velocities; row two: numerical phase velocities for $G = 5$ using the mixed finite difference scheme with mass lumping; row three: numerical phase velocities for $G = 5$ using the classical second order centred finite difference scheme without mass lumping; column one: phase velocities for qP; column two: phase velocities for $qS1$; column three: phase velocities for $qS2$.

As an example, we compute phase velocities in an orthorhombic model described in Fig. 2, with respect to the full range of $1/G = [0.01, 0.3]$, $\theta = [0, \pi]$ and $\phi = [0, \pi]$. We cast the evaluation of weights into a global optimization framework proposed by Sen & Stoffa (1995), based on a fast simulation annealing algorithm. The weights are found to be $w_{s1} = 0.2$, $w_{s2} = 0.6$, $w_{s3} = 0.2$; $w_{m1} = 0.5$, $w_{m2} = 0.45$, $w_{m3} = 0.05$, $w_{m4} = 0.0$. In row one, we solve eq. (12) and obtain three theoretical phase velocities for qP, $qS1$ and $qS2$ modes. Then we solve the eq. (11) with $G = 5$. Row two and three display results for classical second-order centred finite difference without mass lumping and the mixed grid finite difference with mass lumping, respectively. The three columns display phase velocities for qP, $qS1$ and $qS2$, respectively. We present normalized phase velocities $V_{\mathrm{ph}}/\widetilde{V}_{\mathrm{ph}}$ with respect to $1/G$ in Fig. 3, for angles indicated by solid dots in Fig. 2. Row one and two display results for the second-order centred finite difference and the optimized finite difference approximation, respectively. Column one to three display the phase velocities for qP, $qS1$ and $qS2$, respectively. We observe that five gridpoint per shear wavelength allow us to conduct modelling with at least three-digit accuracy.

3 NESTED DISSECTION AND THE MULTIFRONTAL METHOD FOR THE ELASTIC SYSTEM OF EQUATIONS

In this section, we give a brief overview of the structured multifrontal solver together with the nested dissection based domain decomposition techniques introduced by Wang *et al.* (2011a) and Wang *et al.* (2012), for modelling time-harmonic waves in (anisotropic) acoustic media. Then we study in detail similarities and differences between the acoustic and the elastic modelling, via the same structured multifrontal factorization approach. Eventually we present comparisons of both computational complexity and storage between the acoustic and elastic modelling, for the same problem size.

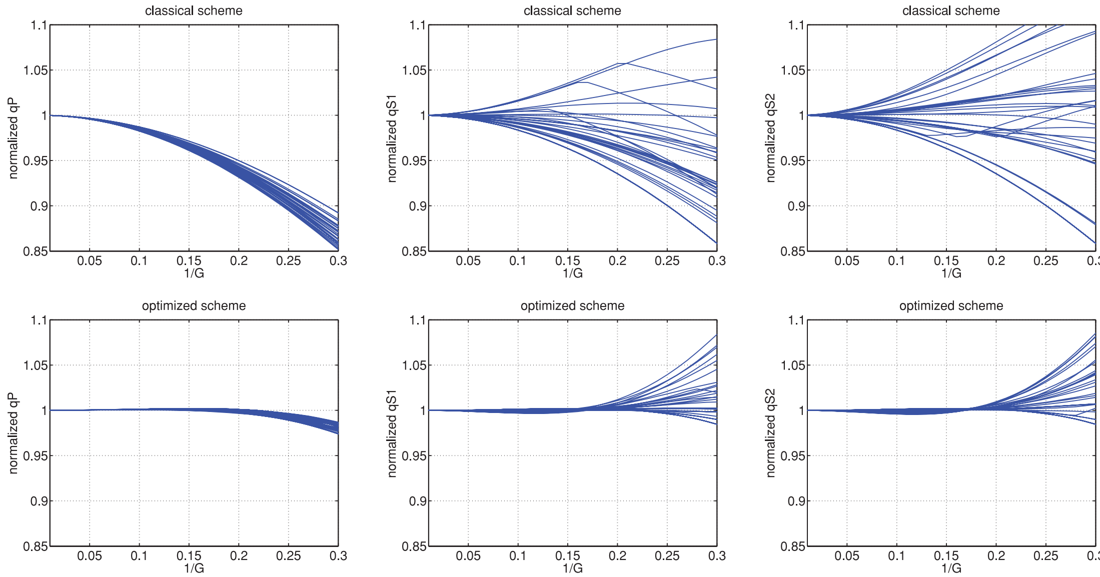


Figure 3. Dispersion curves for the orthorhombic model and angles indicated by dots in Fig. 2; row one: the classical second-order centred finite difference scheme without mass lumping; row two: the mixed finite difference scheme with mass lumping; column one: qP dispersion curves; column two: $qS1$ dispersion curves; column three: $qS2$ dispersion curves.

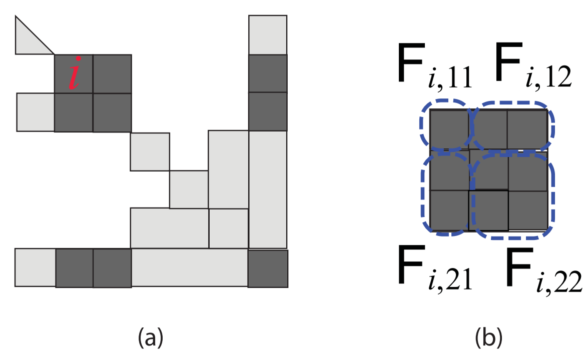


Figure 4. The illustration of the multifrontal method summarized in eq. (13): (a) the multifrontal factorization stage associated with the node i; (b) the formation of the frontal matrix $\mathbf{F}_i$.

3.1 Overview of the structured multifrontal solver

To solve the acoustic analogue matrix system of equation (*cf.* 9), (Wang *et al.* 2011a) introduced a massively parallel structured multifrontal solver together with the nested dissection based domain decomposition, imbedding a scalable HSS matrix solver. They showed that, in 3-D, the computational complexity associated with the factorization is between $\mathcal{O}(n \log n)$ and $\mathcal{O}(n^{4/3} \log n)$, and storage is between $\mathcal{O}(n)$ and $\mathcal{O}(n \log n)$, reminding that n denotes the size of the matrix $\mathbf{A}(\omega)$.

They first conduct the nested dissection reordering (see George 1973) of $\mathbf{A}(\omega)$ by dividing upper level domains into lower level subdomains and separators recursively, imposing that the mesh points associated with subdomains are reordered prior to ones associated with separators, and lower level domains are reordered prior to the upper level ones. This yields a post-ordering tree structure named assembly tree. The nested dissection reordering, which essentially can be viewed as hierarchical domain decompositions, has been proven to be the optimal reordering strategy that minimizes the fill-in of the factorization. Furthermore, to account for the anisotropy and variable order of accuracy, (Wang *et al.* 2012) extended the nested dissection to incorporate separators of variable thickness.

Secondly, after the nested dissection reordering, (Wang *et al.* 2011a) carry out local partial LU factorizations upon the reordered matrix, via forming frontal matrices $\mathbf{F}_i$ and computing update matrices $\mathbf{U}_i$ locally on each node i of the assembly tree, by taking advantage of the multifrontal method introduced by Liu (1992). We summarize the mathematics of the multifrontal method in the concise way below

$$
\mathbf{F}_i = \begin{pmatrix} \mathbf{F}_{i,11} & \mathbf{F}_{i,12} \\ \mathbf{F}_{i,21} & \mathbf{F}_{i,22} \end{pmatrix} = \begin{cases} \begin{pmatrix} \mathbf{A}_{i,11} & \mathbf{A}_{i,12} \\ \mathbf{A}_{i,21} & 0 \end{pmatrix}, & \text{if } i \text{ is a leaf node,} \\ \begin{pmatrix} \mathbf{A}_{i,11} & \mathbf{A}_{i,12} \\ \mathbf{A}_{i,21} & 0 \end{pmatrix} + \left(\mathbf{U}_{c_1} \updownarrow\!\!\leftrightarrow \mathbf{U}_{c_2}\right), & \text{if } i \text{ is a non-leaf node,} \end{cases} \quad (13)
$$
$$
\mathbf{U}_i = \mathbf{F}_{i,22} - \mathbf{F}_{i,21}\, \mathbf{F}_{i,11}^{-1}\, \mathbf{F}_{i,12},
$$

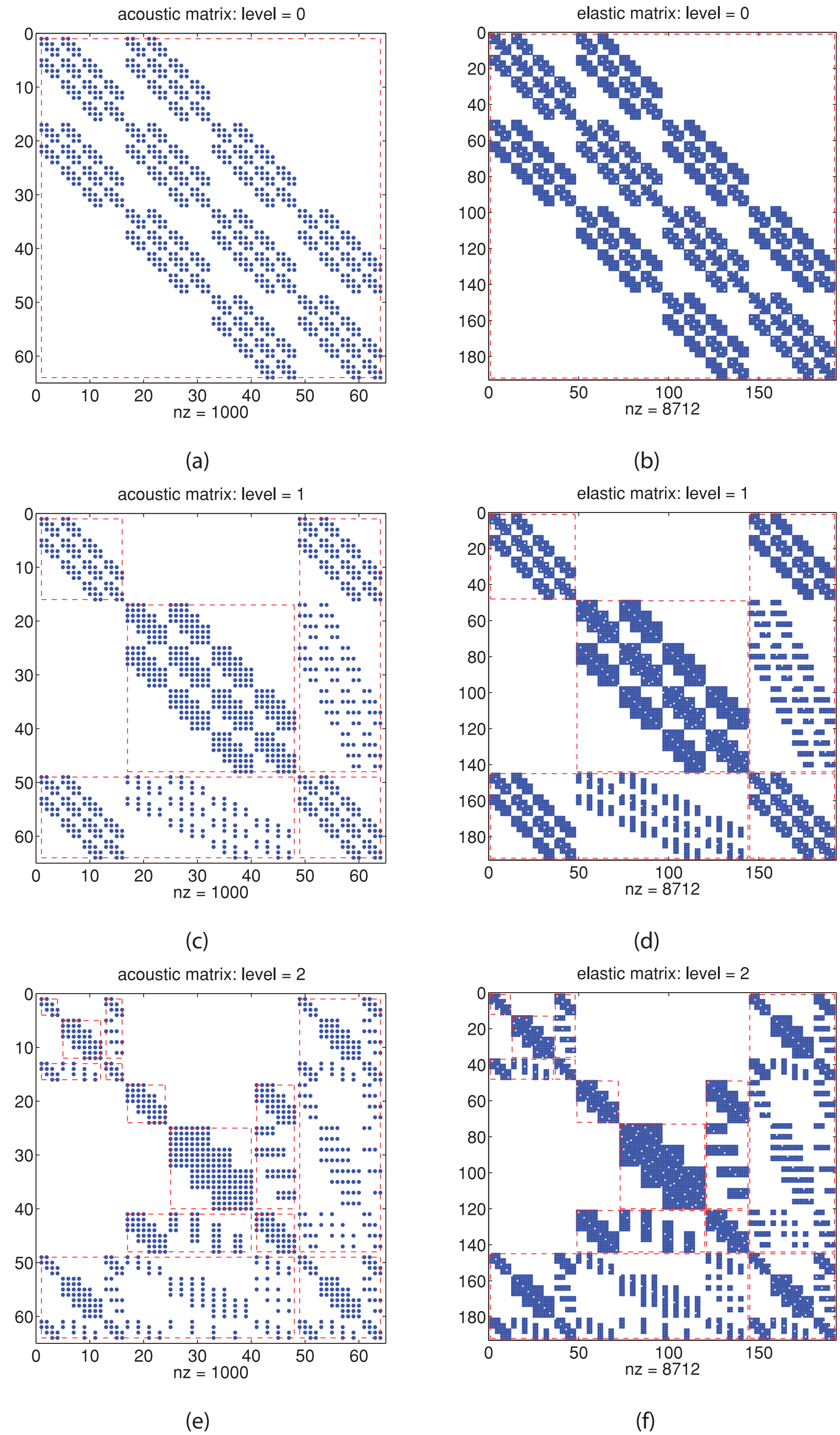
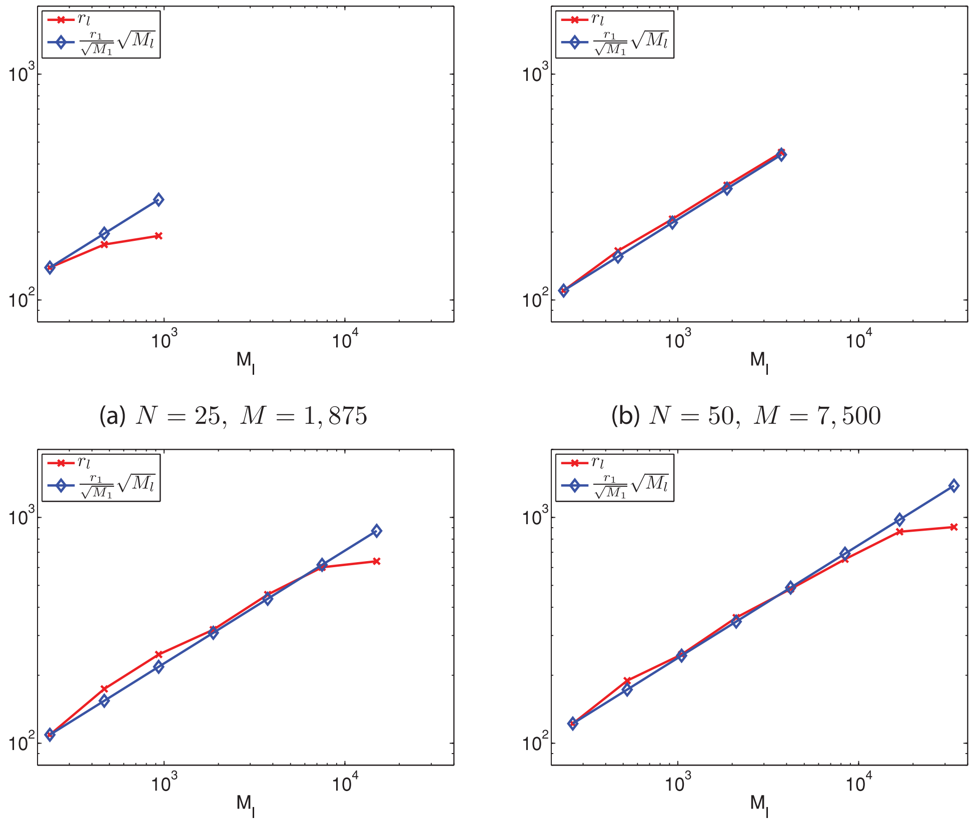


Figure 5. The pattern of the matrix $\mathbf{A}(\omega)$ discretized on a 3-D $4 \times 4 \times 4$ mesh. (a) The acoustic matrix without nested dissection; (b) the elastic matrix without nested dissection; (c) the acoustic matrix with one-level nested dissection; (d) the elastic matrix with one-level nested dissection; (e) the acoustic matrix with two-level nested dissection and (f) the elastic matrix with two-level nested dissection.

where $\mathbf{A}_i$ denotes the portion of the global matrix $\mathbf{A}(\omega)$ associated with the node i on the assembly tree; c_1 and c_2 are two children of the node i if i is a non-leaf, satisfying $c_1 < c_2 < i$. The extend-add operation is denoted by $\updownarrow\!\!\!\!\leftrightarrow$. Fig. 4 illustrates the multifrontal process associated with the node i on the assembly tree.

Thirdly, by exploiting low rank properties of off-diagonal blocks of frontal matrices $\mathbf{F}_i$, (Wang *et al.* 2011b) proposed a series of parallel HSS compression, *ULV* factorization and solution techniques. They show that with the aid of HSS implementations, the cost of factorization



(a) $N = 25,\ M = 1,875$ (b) $N = 50,\ M = 7,500$

(c) $N = 100,\ M = 30,000$ (d) $N = 150,\ M = 67,500$

Figure 6. Rank patterns of frontal matrices $\mathbf{F}_i$ of order M that arises from the factorization of the sparse matrices discretized on $N \times N \times N$ meshes.

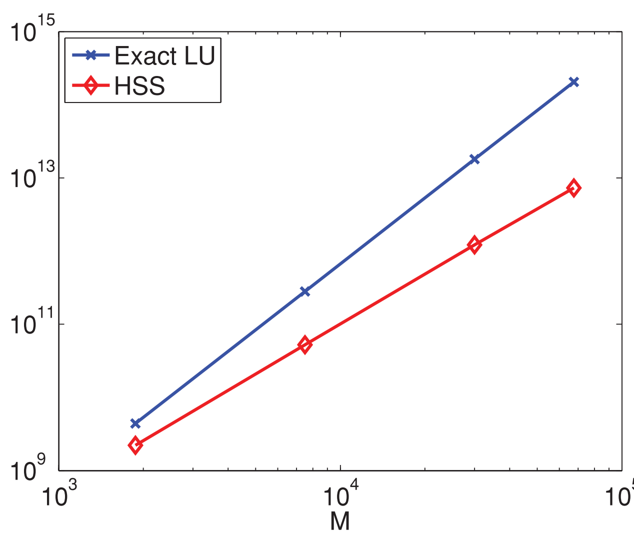


Figure 7. HSS construction costs (flops) for the largest frontal matrix $\mathbf{F}_i$ of order M that arises from the factorization of the sparse matrices discretized on $N \times N \times N$ meshes, where $N = 25$, 50, 100, 150, with the corresponding $M = 1875, 7500, 30000, 67500$.

of each frontal matrix is reduced from $\mathcal{O}(n_i^3)$ to $\mathcal{O}(r_i n_i^2)$, as well as the storage is reduced from $\mathcal{O}(n_i^2)$ to $\mathcal{O}(r_i n_i)$, where n_i denotes the size of each frontal matrix $\mathbf{F}_i$, and r_i denotes the maximum numerical rank of all off-diagonal blocks associated with each $\mathbf{F}_i$.

3.2 Similarities and differences: acoustic versus elastic

We point out that a big difference between the acoustic modelling and the elastic modelling for the same problem size $N_1 \times N_2 \times N_3$ lies in that the number of unknowns for the elastic system is $3N_1N_2N_3$, which is exactly three times as large as the number of unknowns for the

Table 1. The benchmark comparison between MUMPS and the hybrid MF-HSS solver, solving for a 10 Hz time-harmonic wavefield on a 201 × 201 × 151 mesh. Nested dissection wall time T_{ND}, matrix factorization wall time T_{fact}, one shot solution wall time T_{sol}, and the memory consumption are recorded.

	T_{ND}(s)	T_{fact}(s)	T_{sol}(s)	Memory (GB)
MUMPS	48.63	31108	133.45	3947
hybrid MF-HSS	0.12	2947	3.67	2316

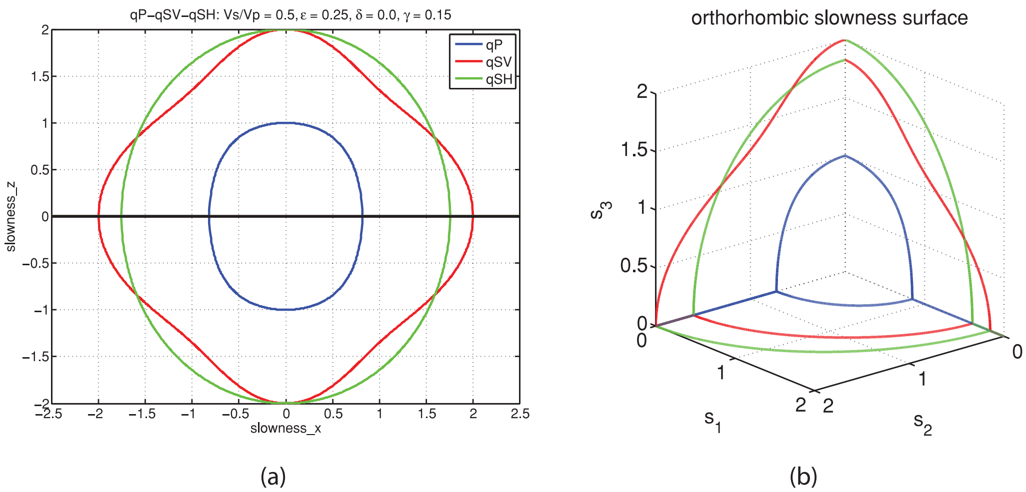


Figure 8. (a) The slowness surfaces of the qP, qSV and qSH polarization for the homogeneous elastic VTI model. The parameters of the model are: 'vertical' P-wave velocity $V_p = 4000\,\mathrm{m\,s^{-1}}$, 'vertical' S-wave velocity $V_s = 2000\,\mathrm{m\,s^{-1}}$, density $\rho = 1.0$, Thomsen's parameters $\epsilon = 0.25$, $\delta = 0.0$, $\gamma = 0.15$; (b) The slowness surfaces of the qP, $qS1$ and $qS2$ polarizations for the homogeneous elastic orthorhombic model. The parameters of the model are: 'vertical' P-wave velocity $V_p = 4000\,\mathrm{m\,s^{-1}}$, 'vertical' S-wave velocity $V_s = 2000\,\mathrm{m\,s^{-1}}$, density $\rho = 1.0$, Thomsen's parameters $\epsilon^{(1)} = 0.2$, $\epsilon^{(2)} = 0.45$, $\delta^{(1)} = -0.1$, $\delta^{(2)} = 0.2$, $\delta^{(3)} = -0.15$, $\gamma^{(1)} = 0.28$, $\gamma^{(2)} = 0.15$.

acoustic system. This is obviously due to the fact that there are three components associated with each mesh point for the elastic system rather than only one component that is for the acoustic system. This implies that the size of the elastic matrix $\mathbf{A}(\omega)$ is three times as large as the size of the acoustic matrix.

On the other hand, we assume that the thickness of the separators for both the acoustic and elastic systems is t, which yields that the number of mesh points for the finite difference stencil associated with the acoustic system is $(2t + 1)^3$, while the one for the elastic system finite difference stencil is $3(2t + 1)^3$. For example, for the acoustic system (Wang *et al.* 2011a) utilize a 27-point stencil ($t = 1$), which corresponds with an 81-point stencil in the elastic case.

Because both the acoustic system and the elastic system share the same problem size $N_1 \times N_2 \times N_3$ and the same thickness of separator t, thus they share exactly the same nested dissection strategy, which, in other words, means that positions and sizes associated with subdomains and separators are exactly the same for both acoustic and elastic system. The only difference between two systems is that after nested dissection reordering, the size n_i of each frontal matrix $\mathbf{F}_i$ associated with the elastic system is exactly three times as large as each n_i associated with the acoustic system. Fig. 5 illustrates matrix patterns for both acoustic and elastic systems discretized on the same 4 × 4 × 4 mesh, for various levels of nested dissection reordering. We note the similarity of matrix patterns, and the size difference of each matrix block by a factor of three.

Without resorting to the HSS compression and factorization techniques, we can straightforwardly conclude that the computational complexity for the elastic matrix factorization is 3^3 times larger than the one associated with the acoustic system, and the storage for the elastic system is 3^2 times larger than the one for the acoustic system. This is due to the well known fact that the cost of exact LU factorization of each dense frontal matrix $\mathbf{F}_i$ is $\mathcal{O}(n_i^3)$, and the storage is $\mathcal{O}(n_i^2)$. By virtue of HSS low rank compression techniques, the complexity and storage are of the order $\mathcal{O}(r_i n_i^2)$ and $\mathcal{O}(r_i n_i)$, respectively. This brings the complexity ratio from 3^3 to 3^2, and the storage ratio from 3^2 to 3, which comprises the main result of this paper.

4 PERFORMANCE

Here, we briefly discuss the complexity of the solver. The solver performs well if the off-diagonal blocks of the frontal matrices $\mathbf{F}_i$ in (13) have small numerical ranks. In general, such a requirement is not satisfied for 3-D (elastic) problems. However, in Xia (2012a,b), it is shown

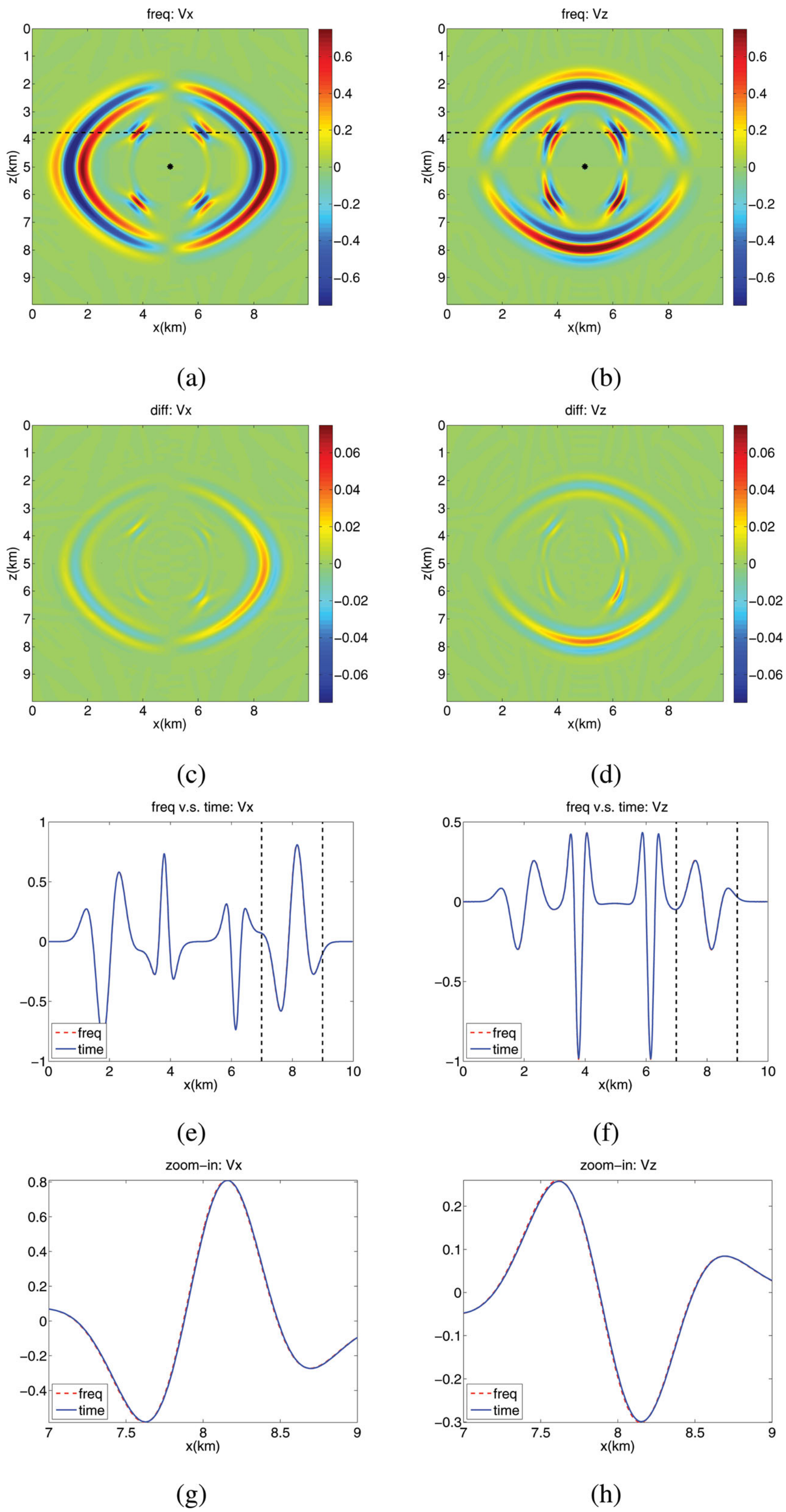


Figure 9. A comparison between our frequency domain finite difference multicomponent modelling and a multicomponent time domain Discontinuous Galerkin method based modelling. (a) $v_1 = \partial_t u_1$ component computed by the frequency domain code; (b) $v_3 = \partial_t u_3$ component computed by the frequency domain code; (c) v_1 difference from the time domain results displayed with a $10\times$ clipping; (d) v_3 difference from the time domain results displayed with a $10\times$ clipping; (e) a trace comparison of v_1 on the indicated dashed line of (a); (f) a trace comparison of v_3 on the indicated dashed line of (b); (g) the zoom-in comparison of the window indicated by the dashed line in (e); (h) the zoom-in comparison of the window indicated by the dashed line in (f).

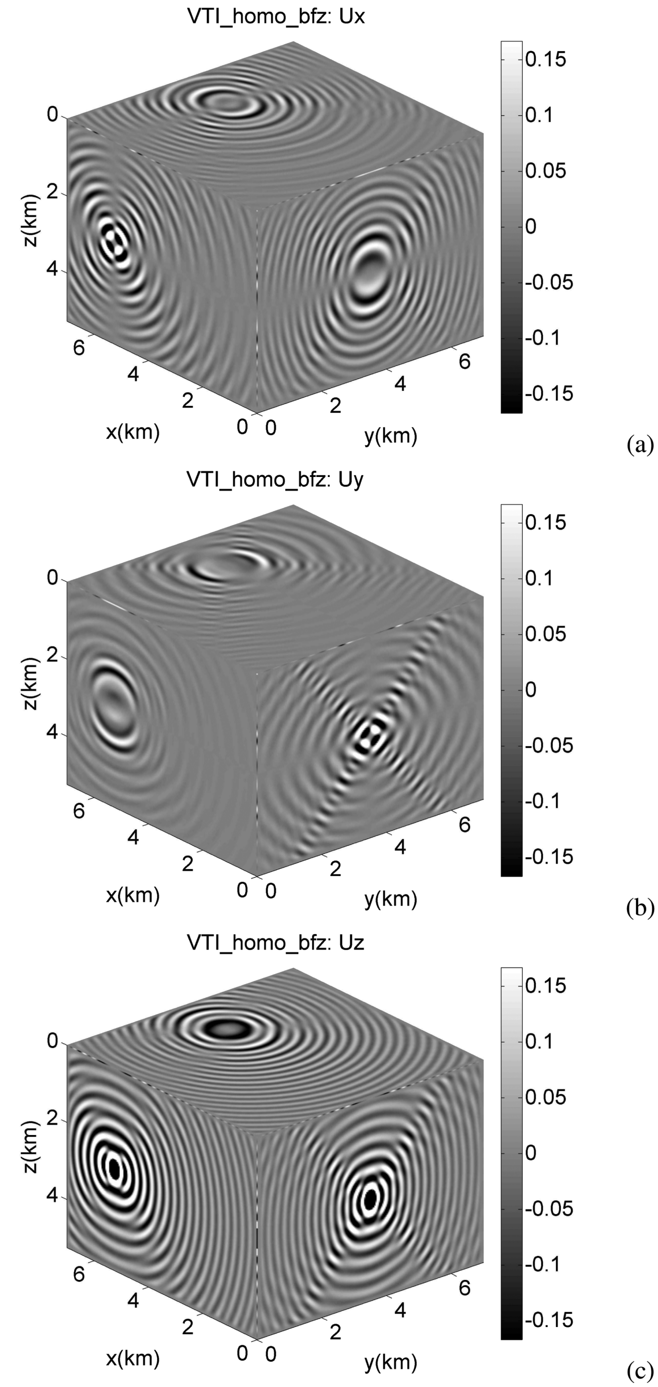


Figure 10. Multicomponent time harmonic wavefield generated by a vertical point body force in the homogeneous elastic VTI model depicted in Fig. 8. The point body force is located at (5.25, 3.5, 2.625) km. (a) u_1 displayed on the planes of $x_1 = 6.125$ km, $x_2 = 3.5$ km and $x_3 = 3.5$ km; (b) u_2 displayed on the planes of $x_1 = 5.25$ km, $x_2 = 4.375$ km and $x_3 = 3.5$ km; (c) u_3 displayed on the planes of $x_1 = 5.25$ km, $x_2 = 3.5$ km and $x_3 = 3.5$ km.

that if the off-diagonal numerical ranks satisfy certain patterns, then these structured solvers can still work well, and the complexity is similar to the case where the ranks are bounded.

For mesh dimensions $N = 25, 50, 100, 150$, we demonstrate the off-diagonal numerical ranks of the largest frontal matrix $\mathbf{F}_i$ that is of order, say M. $\mathbf{F}_i$ is hierarchically partitioned into multiple levels following the definition of HSS matrices (Xia *et al.* 2010). The largest level is where there are most subblocks. The maximum numerical rank r_l at each level l of the partition is recorded, and plotted in Fig. 6. Although a precise justification is not yet available, these ranks are observed to closely follow the following pattern

$$r_l = \mathcal{O}\left(\sqrt{M_l}\right),$$

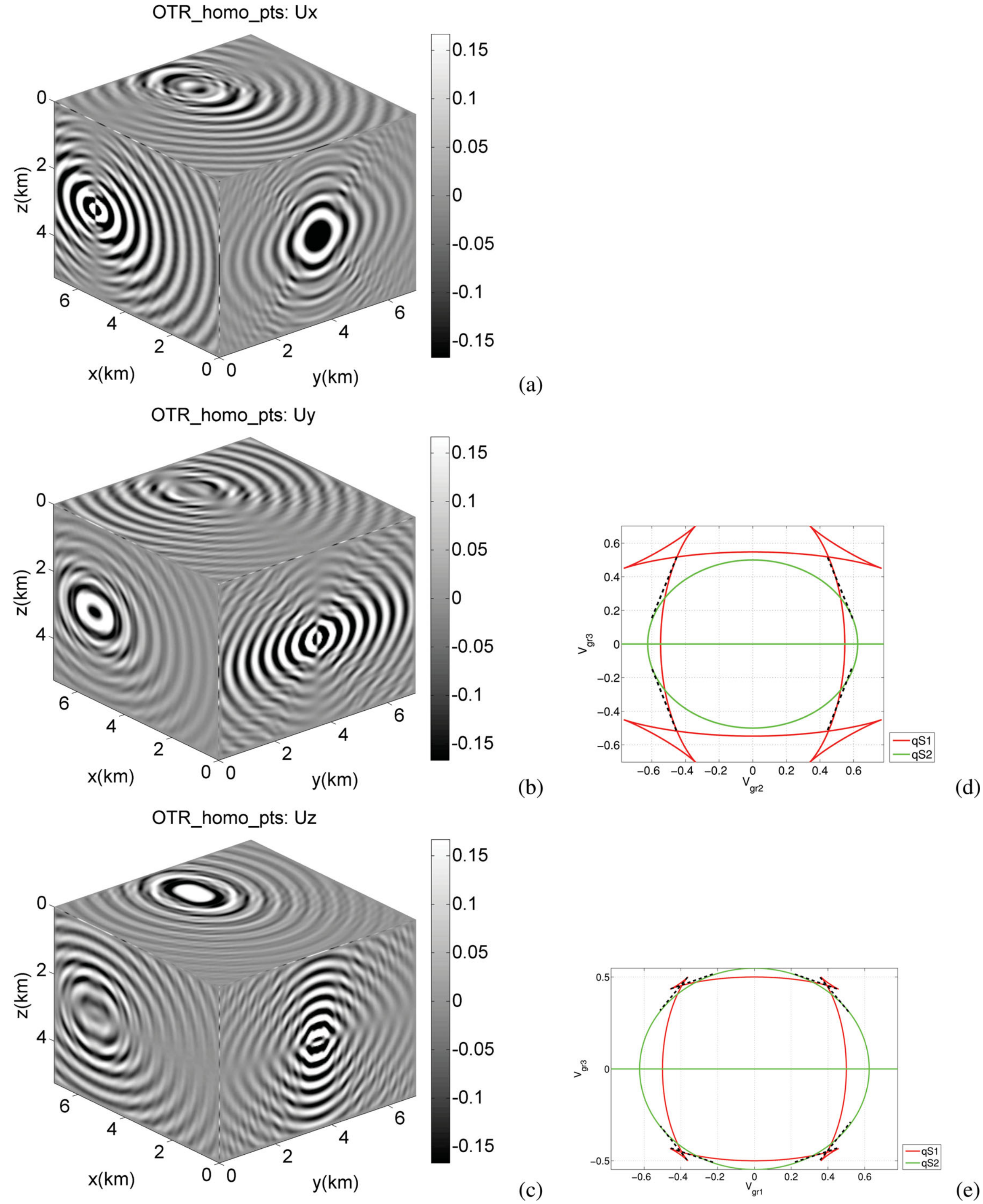


Figure 11. Multicomponent time harmonic wavefield generated by an explosive point source in the homogeneous orthorhombic model with parameters depicted in Fig. 8 and additional parameters: $\epsilon^{(1)} = 0.2$, $\epsilon^{(2)} = 0.45$, $\delta^{(1)} = -0.1$, $\delta^{(2)} = 0.2$, $\delta^{(3)} = -0.15$, $\gamma^{(1)} = 0.28$, $\gamma^{(2)} = 0.15$. The explosive point source is located at (5.25, 3.5, 2.625) km. (a) u_1 displayed on the planes of $x_1 = 6.125$ km, $x_2 = 3.5$ km and $x_3 = 3.5$ km; (b) u_2 displayed on the planes of $x_1 = 5.25$ km, $x_2 = 4.375$ km and $x_3 = 3.5$ km; (c) u_3 displayed on the planes of $x_1 = 5.25$ km, $x_2 = 3.5$ km and $x_3 = 3.5$ km; (d) The wavefront in the symmetry plane $x_2 = 3.5$ km; (e) The wavefront in the symmetry plane $x_1 = 5.25$ km.

where M_l is the maximum row size of the blocks at level l of the partition. With this pattern, it is shown in Xia (2012a,b) that an HSS approximation to $\mathbf{F}_i$ can be constructed in ξ_0 flops and factorized in ξ_1 flops, where

$$\xi_0 = \mathcal{O}\left(M^2 \log M\right), \ \xi_1 = \mathcal{O}\left(M^{3/2}\right).$$

Moreover, the solution cost is

$$\xi_2 = \mathcal{O}\left(M \log M\right).$$

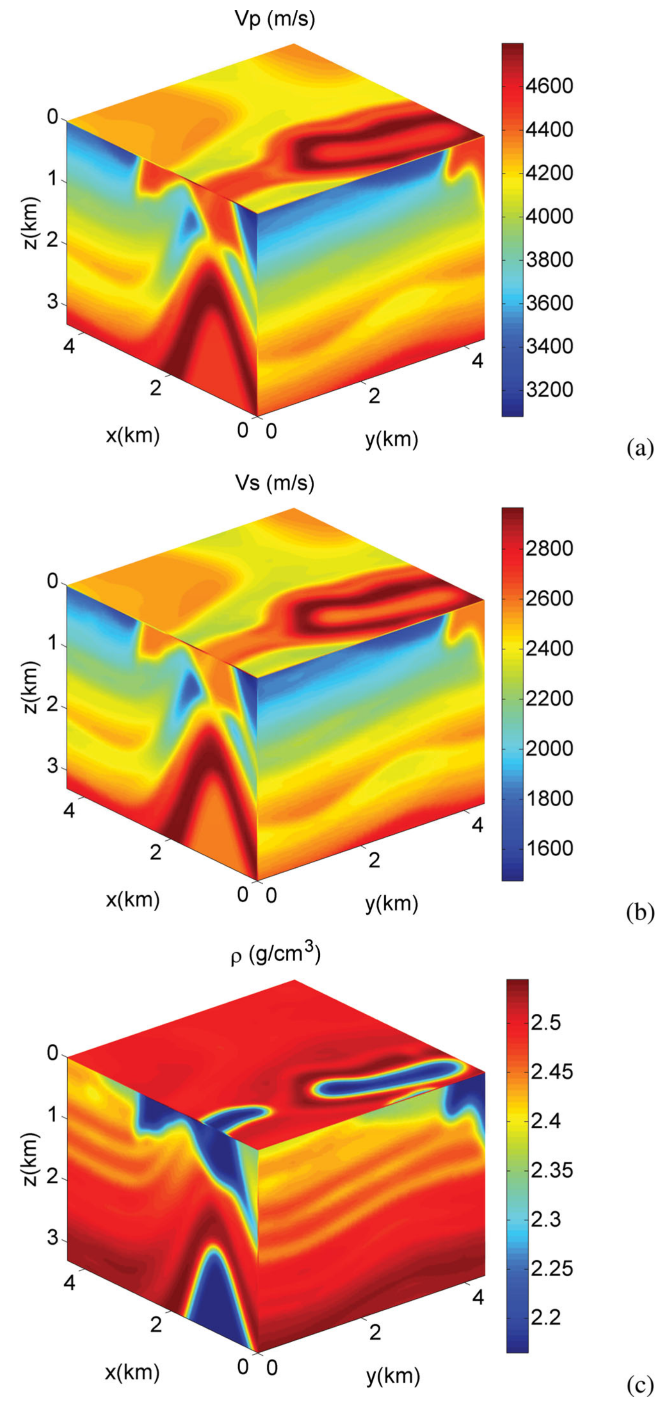


Figure 12. The inhomogeneous elastic VTI model discretized on a 201 × 201 × 151 mesh with stepsizes $h_1 = h_2 = h_3 = 0.022$ km. All models are displayed on the planes of $x_1 = 3.3$ km, $x_2 = 2.2$ km and $x_3 = 1.65$ km. (a) 'vertical' P-wave velocity V_p; (b) 'vertical' S-wave velocity V_s; (c) density ρ.

In contract, the exact factorization and solution costs are $\tilde{\xi}_1 = \mathcal{O}\left(M^3\right)$ and $\tilde{\xi}_2 = \mathcal{O}\left(M^2\right)$, respectively. Notice that when N doubles, M becomes four times larger. Then ξ_0, ξ_1 and ξ_2 increase by factors of 16, 8 and 4, respectively, while $\tilde{\xi}_1$ and $\tilde{\xi}_2$ increase by factors of 64 and 16, respectively. This is illustrated in Fig. 7. The overall sparse structured factorization cost is then $O(n^{4/3}\log n)$, and the solution cost is $O(n\log n)$, which is nearly linear in n (Xia 2012b).

5 NUMERICAL EXPERIMENTS

We present various numerical experiments illustrating the behaviour of time-harmonic elastic waves in 3-D anisotropic media. Both homogeneous and inhomogeneous 3-D models are discretized on a 201 × 201 × 151 mesh, with different step sizes. We set the HSS compression

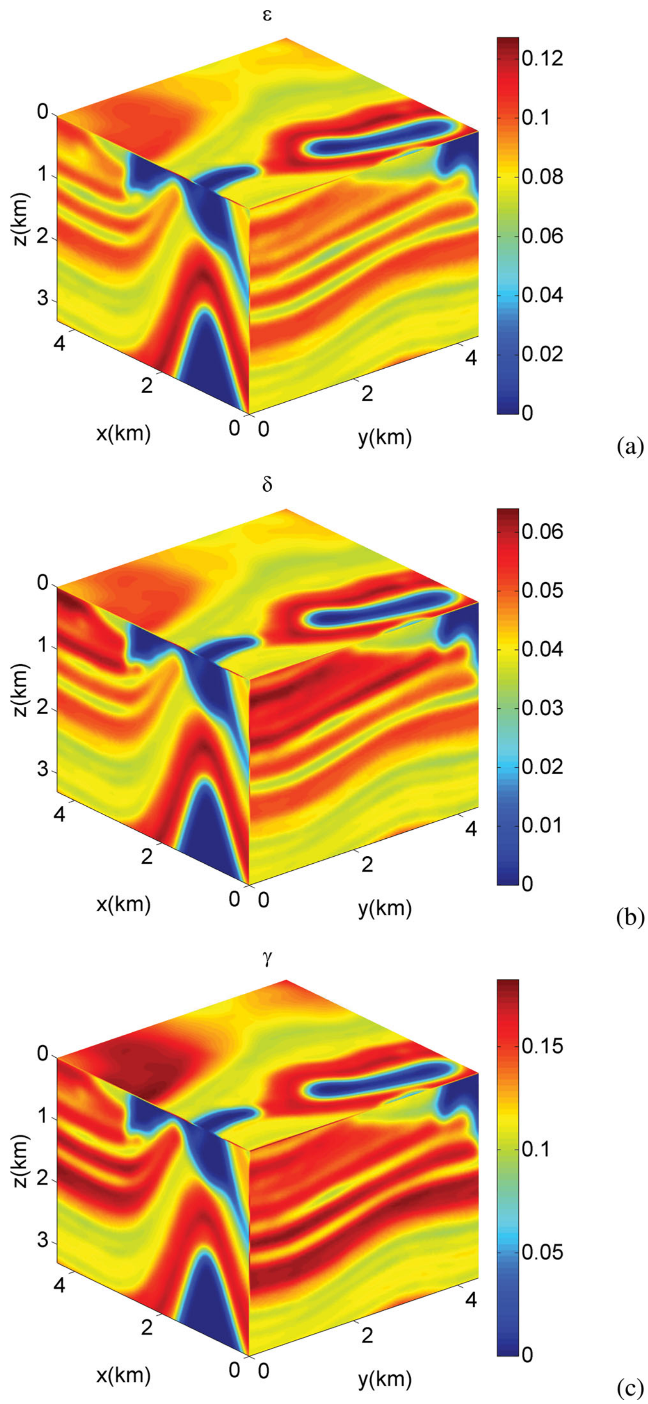


Figure 13. The inhomogeneous elastic VTI model discretized on a 201 × 201 × 151 mesh with stepsizes $h_1 = h_2 = h_3 = 0.022$ km. All models are displayed on the planes of $x_1 = 3.3$ km, $x_2 = 2.2$ km and $x_3 = 1.65$ km. (a) Thomsen's parameter ϵ; (b) Thomsen's parameter δ; (c) Thomsen's parameter γ.

threshold to be $1.0e - 4$, which results in a four digit accuracy of our computed time-harmonic wavefields. All 3-D computations are 10Hz time harmonic wavefields, and are conducted on a National Energy Research Scientific Computing Center (NERSC) supercomputer named Hopper.nersc.gov, utilizing 128 nodes with 16 cores and 32 GB of memory per node. The CPU wall time is 2947 s, and the total memory consumption is 2316 GB.

5.1 Benchmark comparison with MUMPS

Prior to showing numerical results of multicomponent time harmonic wavefields, we conduct benchmark comparison tests on both computational complexity and memory consumption, between our hybrid MF-HSS solver and a general standard matrix solver MUMPS (Agullo

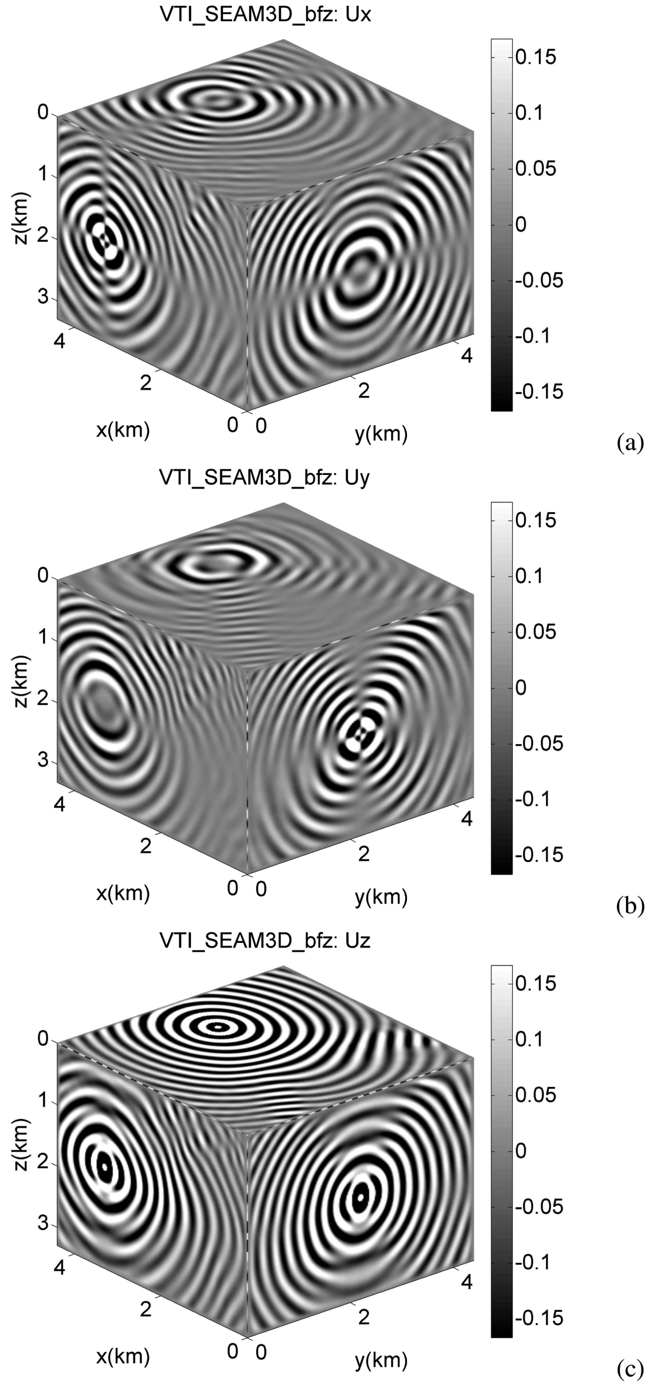


Figure 14. Multicomponent time harmonic wavefield generated by a vertical point body force in the inhomogeneous VTI model depicted in Figs 12 and 13. The body force is located at (3.3, 2.2, 1.65) km. (a) u_1 displayed on the planes of $x_1 = 3.85$ km, $x_2 = 2.2$ km and $x_3 = 2.2$ km; (b) u_2 displayed on the planes of $x_1 = 3.3$ km, $x_2 = 2.75$ km and $x_3 = 2.2$ km; (c) u_3 displayed on the planes of $x_1 = 3.3$ km, $x_2 = 2.2$ km and $x_3 = 1.65$ km.

et al. 2008), on exactly the same platform of 128 nodes of Hopper.nersc.gov. We compute 10 Hz time harmonic wavefields on the same 201 × 201 × 151 mesh, using both solvers. In Table 1, we record the wall time for the nested dissection T_{ND}, the wall time for the matrix factorization T_{fact}, the wall time for the resolution to one right hand side associated with one seismic shot T_{sol}, and the total memory consumption. We note that the computational time associated with our hybrid MF-HSS solver is at least one order of magnitude faster than the one associated with the general MUMPS solver, for all three stages of nested dissection, matrix factorization and solution. In particular, bearing in mind that T_{sol} associates with one seismic shot resolution, it is straightforward to notice the efficiency of the MF-HSS solver, given a large number of seismic shots in practice. We also point out that memory consumed by the MF-HSS solver is around one half of that consumed by the MUMPS solver.

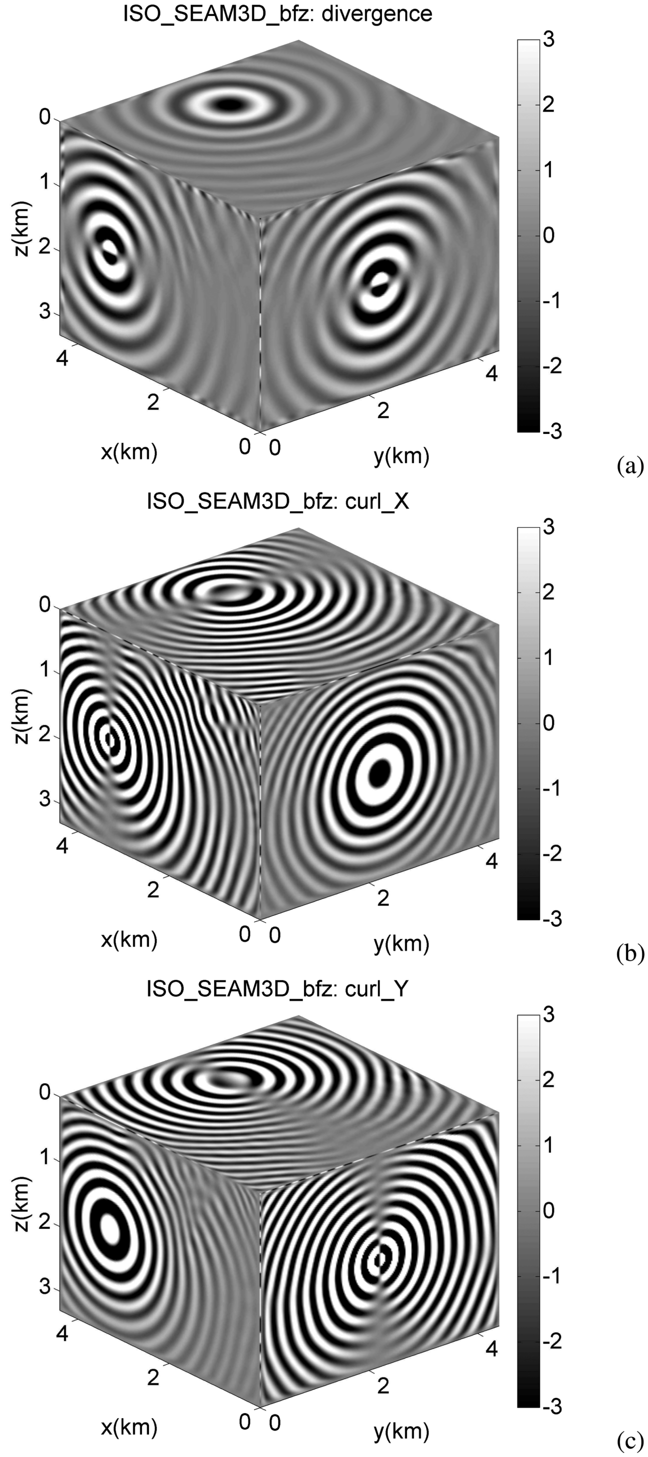


Figure 15. The divergence $\nabla \cdot \mathbf{u}$ and the curl $\nabla \times \mathbf{u}$ of the time-harmonic wavefield generated by a vertical point body force in the inhomogeneous elastic isotropic model depicted in Fig. 12. The point body force is located at (3.3, 2.2, 1.65) km. (a) $\nabla \cdot \mathbf{u}$ displayed on the planes of $x_1 = 3.3$ km, $x_2 = 2.2$ km and $x_3 = 2.2$ km; (b) $(\nabla \times \mathbf{u})_1$ displayed on the planes of $x_1 = 3.85$ km, $x_2 = 2.2$ km and $x_3 = 1.65$ km; (c) $(\nabla \times \mathbf{u})_2$ displayed on the planes of $x_1 = 3.3$ km, $x_2 = 2.75$ km and $x_3 = 1.65$ km.

5.2 Homogenous media

Our reference model is depicted by the left slowness surface in Fig. 8. Here, we consider a homogeneous medium and different point sources. The system of equations is discretized on a $201 \times 201 \times 151$ mesh with step sizes $h_1 = h_2 = h_3 = 0.035$ km, which implies that the model size is $[0, 7]\,\text{km} \times [0, 7]\,\text{km} \times [0, 5.25]\,\text{km}$.

First, we show in Fig. 9(a) multifrequency computation and a comparison with a time-domain Discontinuous Galerkin (DG) method in 2-D. We made use of a 10Hz Ricker wavelet as our source signature. We computed and stored time harmonic wavefields for altogether 300 frequencies ranging from 0 to 30 Hz, followed by an inverse fast Fourier transform (IFFT) back to the time domain to draw comparisons. We observe that the four digit accuracy is sufficient to capture the change of phase in the reverse branch of the triplication. Next, we illustrate the time-harmonic displacement in 3-D, in Fig. 10, for a vertical point body force. We note the interference between the different polarizations. Our computation has been carried out with a sampling rate of five points per shear wavelength. In the u_3 component, in the x_2x_3 plane, we observe the excitation of qP waves in a cone aligned with the x_3 direction and of qSV waves in a cone aligned with the x_2 direction, while the presence of a caustic is clear. In the u_2 component, in the x_2x_3 plane, we observe the qP waves in all directions. Essentially, we observe the imprint of the radiation pattern of the source.

We then consider an orthorhombic model depicted by the right slowness surface in Fig. 8. Our computation has been carried out with a sampling rate of five points per shear wavelength. We illustrate the time-harmonic displacement generated by an explosive point source in Figs 11 (a)–(c). We note the presence of conical points, for example, in the u_1 component in the x_2x_3 plane. In Figs 11 (d)–(e), we show the wavefronts in the $x_2 = 3.5$ km and $x_1 = 5.25$ km (symmetry) planes for comparison. The wavefronts both show the formation of caustics and lids associated with the conical points; the locations of such points is illustrated in the $qS1 - qS2$ slowness surfaces presented in Fig. 8. The wavefronts aid in clarifying the interference of wave constituents in the Figs 11 (a)–(c).

5.3 Inhomogeneous medium

We consider a heterogeneous VTI model, derived from the SEAM3D model, and discretized on a $201 \times 201 \times 151$ mesh with step sizes $h_1 = h_2 = h_3 = 0.022$ km that yields the model size [0, 4.4] km × [0, 4.4] km × [0, 3.3] km. It is illustrated in Figs 12–13. We show the time-harmonic displacement generated by a vertical point body force in Fig. 14. This example confirms and illustrates the performance of our algorithm in a salt tectonic geological environment with strong heterogeneities.

Finally, we generate the time-harmonic displacement generated by a vertical point body force in an isotropic heterogeneous model by using the parameters in Fig. 12. We decompose the solution into P and S polarizations; the results are shown in Fig. 15. The model is smooth and, indeed, the separation is clean.

6 CONCLUSION

We presented a finite-difference modelling algorithm for time-harmonic seismic waves in anisotropic media using locally optimized finite difference stencils. We developed a massively parallel direct structured solver for the relevant system of equations. The system of equations need not be of real principal type. We carried out computational experiments both for TI and orthorhombic media. For fixed frequency, in 3-D, the complexity associated with the elastic system is nine times larger than the one for the acoustic system, and the storage requirement for the elastic modelling is three times larger than the one associated with the acoustic modelling. The solver will play a key role in wave-equation tomography and FWI, and is very well suited for adjoint state computations with a large number of sources (events) potentially on planetary scale.

ACKNOWLEDGMENTS

The authors thank the members, ConocoPhillips, ExxonMobil, PGS, BGP, Statoil and Total, of the Geo-Mathematical Imaging Group (GMIG) for partial financial support. J. Xia was supported in part by NSF grants DMS-1115572 and CHE-0957024, M.V. de Hoop was supported in part by NSF CMG grant DMS-1025318, and X.S. Li was supported in part by the Director, Office of Science, Office of Advanced Scientific Computing Research of the U.S. Department of Energy under the contract DE-AC02-05CH11231. The authors thank the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory (LBNL) for providing the computing resources, and Ruichao Ye for generating time domain multi-component modeling results using the Discontinuous Galerkin method.

REFERENCES

Agullo, E., Guermouche, A. & L'Excellent, J., 2008. A parallel out-of-core multifrontal method: storage of factors on disk and analysis of models for an out-of-core active memory, *Parallel Comput.,* **34,** 296–317.

Airaksinen, T., Pennanen, A. & Toivannen, J., 2009. A damping preconditioner for time-harmonic wave equations in fluid and elastic material, *J. Comput. Phys.,* **228,** 1466–1479.

Alpert, B., Beylkin, G., Coifman, R. & Rokhlin, V., 1993. Wavelet-like bases for the fast solution of second-kind integral equations, *SIAM J. Sci. Comput.,* **14,** 159–184.

Bansal, R. & Sen, M., 2008. Finite-difference modeling of S-wave splitting in anisotropic media, *Geophys. Prospect.,* **56,** 293–312.

El Kacimi, A. & Laghrouche, O., 2011. Wavelet based *ILU* preconditioners for the numerical solution by PUFEM of high frequency elastic wave scattering, *J. Comput. Phys.,* **230,** 3119–3134.

Fichtner, A., Igel, H., Bunge, H. & Kennett, B., 2009. Simulation and inversion of seismic wave propagation on continental scales based on a spectral-element method, *J. Numer. Anal. Indust. appl. Math.,* **4,** 11–22.

Gauzellino, P., Santos, J. & Sheen, D., 2001. Frequency domain wave propagation modeling in exploration seismology, *J. Comput. Acoust.,* **9,** 941–955.

George, J., 1973. Nested dissection of a regular finite element mesh, *SIAM J. Numer. Anal.,* **10,** 345–363.

Holberg, O., 1987. Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena, *Geophys. Prospect.,* **35,** 629–655.

Jr. Douglas, J., Sheen, D. & Santos, J., 1994. Approximation of scalar waves in the space-frequency domain, *Math. Models Methods appl. Sci.,* **4,** 509–531.

Liu, J., 1992. The multifrontal method for sparse matrix solution: theory and practice, *SIAM Rev.,* **34,** 82–109.

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

Pratt, R., 1990. Frequency-domain elastic wave modeling by finite differences: a tool for crosshole seismic imaging, *Geophysics,* **55,** 626–632.

Sen, M. & Stoffa, P., 1995. *Global Optimization Methods in Geophysical Inversion,* Elsevier, Amsterdam.

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

Stolk, C. & de Hoop, M., 2002. Microlocal analysis of seismic inverse scattering in anisotropic elastic media, *Commun. Pure appl. Math.,* **55,** 261–301.

van Stralen, M., de Hoop, M. & Blok, H., 1998. Generalized bremmer series with rational approximation for the scattering of waves in inhomogeneous media, *J. acoust. Soc. Am.,* **104,** 1943–1963.

Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics,* **51,** 1954–1966.

Tromp, J., Komatitsch, D. & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.,* **3,** 1–32.

Tsvankin, I., 1997. Anisotropic parameters and P-wave velocity for orthorhombic media, *Geophysics,* **62,** 1292–1309.

Wang, S., De Hoop, M. & Xia, J., 2010. Seismic inverse scattering via Helmholtz operator factorization and optimization, *J. Comput. Phys.,* **229,** 8445–8462.

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

Wang, S., Li, X., Xia, J., Situ, Y. & de Hoop, M., 2011. Efficient scalable algorithms for hierarchically semiseparable matrices, *SIAM J. Sci. Comput.,* in press.

Wang, S., Xia, J., de Hoop, M. & Li, X., 2012. Massively parallel structured direct solver for equations describing time-harmonic *qp*-polarized waves in TTI media, *Geophysics,* **77,** 69–82.

Xia, J., 2012a. On the complexity of some hierarchical structured matrices, *SIAM J. Matrix Anal. Appl.,* **33,** 388–410.

Xia, J., 2012b. Efficient structured multifrontal factorization for general large sparse matrices, *SIAM J. Sci. Comput.* Available at: http://www.math.purdue.edu/~xiaj/work/mfhss.pdf.

Xia, J., Chandrasekaran, S., Gu, M. & Li, X., 2009. Superfast multifrontal method for large structured linear systems of equations, *SIAM J. Matrix Anal. Appl.,* **31,** 1382–1411.

Xia, J., Chandrasekaran, S., Gu, M. & Li, X., 2010. Fast algorithms for hierarchically semiseparable matrices, *Numer. Linear Algebra Appl.,* **17,** 953–976.

APPENDIX: THE EVALUATION OF A_{11} IN ORTHORHOMBIC MEDIA

In this section, we present the evaluation of one entry $\widehat{A}_{11}(\mathbf{x}, \omega)$ out of $\widehat{\mathbf{A}}(\mathbf{x}, \omega)$ in eq. (6), for orthorhombic media. For the sake of brevity, we denote the u_1 component as v, and normalize A_{11} by the multiplication of h^2 where h is the mesh step size. We recall that

$$A_{11}(\mathbf{x}, \partial_{\mathbf{x}}, \omega) = -\frac{\partial}{\partial x_1}\left(C_{11}\frac{\partial}{\partial x_1}\right) - \frac{\partial}{\partial x_2}\left(C_{66}\frac{\partial}{\partial x_2}\right) - \frac{\partial}{\partial x_3}\left(C_{55}\frac{\partial}{\partial x_3}\right).$$

Similar to eq. (6), we have

$$A_{11} = w_{s1}A_{11}^{(c)} + \frac{w_{s2}}{3}\left(A_{11}^{(x)} + A_{11}^{(y)} + A_{11}^{(z)}\right) + \frac{w_{s3}}{4}\left(A_{11}^{(1)} + A_{11}^{(2)} + A_{11}^{(3)} + A_{11}^{(4)}\right);$$

here,

$$\begin{aligned} A_{11}^{(c)} = {} & \left[(C_{11})_{-\frac{1}{2}00}(v_{000} - v_{-100}) + (C_{11})_{\frac{1}{2}00}(v_{000} - v_{100})\right] + \left[(C_{66})_{0-\frac{1}{2}0}(v_{000} - v_{0-10}) + (C_{66})_{0\frac{1}{2}0}(v_{000} - v_{010})\right] \\ & + \left[(C_{55})_{00-\frac{1}{2}}(v_{000} - v_{00-1}) + (C_{55})_{00\frac{1}{2}}(v_{000} - v_{001})\right]; \end{aligned}$$

and

$$\begin{aligned} A_{11}^{(x)} = {} & \left[(C_{11})_{-\frac{1}{2}00}(v_{000} - v_{-100}) + (C_{11})_{\frac{1}{2}00}(v_{000} - v_{100})\right] + \frac{1}{4}\left[(C_{66})_{0-\frac{1}{2}\frac{1}{2}}(v_{000} - v_{0-11}) + (C_{66})_{0\frac{1}{2}-\frac{1}{2}}(v_{000} - v_{01-1})\right] \\ & + \frac{1}{4}\left[(C_{66})_{0-\frac{1}{2}-\frac{1}{2}}(v_{000} - v_{0-1-1}) + (C_{66})_{0\frac{1}{2}\frac{1}{2}}(v_{000} - v_{011})\right] + \frac{1}{4}\left[(C_{66})_{0\frac{1}{2}-\frac{1}{2}}(v_{010} - v_{00-1}) - (C_{66})_{0-\frac{1}{2}\frac{1}{2}}(v_{001} - v_{0-10})\right] \\ & + \frac{1}{4}\left[(C_{66})_{0\frac{1}{2}\frac{1}{2}}(v_{010} - v_{001}) - (C_{66})_{0-\frac{1}{2}-\frac{1}{2}}(v_{00-1} - v_{0-10})\right] + \frac{1}{4}\left[(C_{55})_{0-\frac{1}{2}\frac{1}{2}}(v_{000} - v_{0-11}) + (C_{55})_{0\frac{1}{2}-\frac{1}{2}}(v_{000} - v_{01-1})\right] \\ & + \frac{1}{4}\left[(C_{55})_{0-\frac{1}{2}-\frac{1}{2}}(v_{000} - v_{0-1-1}) + (C_{55})_{0\frac{1}{2}\frac{1}{2}}(v_{000} - v_{011})\right] - \frac{1}{4}\left[(C_{55})_{0\frac{1}{2}-\frac{1}{2}}(v_{010} - v_{00-1}) - (C_{55})_{0-\frac{1}{2}\frac{1}{2}}(v_{001} - v_{0-10})\right] \\ & - \frac{1}{4}\left[(C_{55})_{0\frac{1}{2}\frac{1}{2}}(v_{010} - v_{001}) - (C_{55})_{0-\frac{1}{2}-\frac{1}{2}}(v_{00-1} - v_{0-10})\right]; \end{aligned}$$

similar expressions are obtained for $A_{11}^{(y)}$ and $A_{11}^{(z)}$. Moreover,

$$
\begin{aligned}
A_{11}^{(1)} = {} & \frac{1}{4}\left[(C_{11})_{-\frac12\frac12\frac12}(v_{000}-v_{-111})+(C_{11})_{\frac12-\frac12-\frac12}(v_{000}-v_{1-1-1})\right]+\frac{1}{4}\left[(C_{11})_{\frac12\frac12-\frac12}(v_{000}-v_{11-1})+(C_{11})_{-\frac12-\frac12\frac12}(v_{000}-v_{-1-11})\right]\\
& -\frac{1}{4}\left[(C_{11})_{-\frac12\frac12\frac12}(v_{010}-v_{-101})-(C_{11})_{\frac12-\frac12-\frac12}(v_{10-1}-v_{0-10})\right]-\frac{1}{4}\left[(C_{11})_{\frac12\frac12-\frac12}(v_{010}-v_{10-1})-(C_{11})_{-\frac12-\frac12\frac12}(v_{-101}-v_{0-10})\right]\\
& +\frac{1}{4}\left[(C_{66})_{\frac12\frac12\frac12}(v_{000}-v_{111})+(C_{66})_{-\frac12-\frac12-\frac12}(v_{000}-v_{-1-1-1})\right]+\frac{1}{4}\left[(C_{66})_{\frac12\frac12-\frac12}(v_{000}-v_{11-1})+(C_{66})_{-\frac12-\frac12\frac12}(v_{000}-v_{-1-11})\right]\\
& +\frac{1}{4}\left[(C_{66})_{\frac12\frac12-\frac12}(v_{110}-v_{00-1})-(C_{66})_{-\frac12-\frac12\frac12}(v_{001}-v_{-1-10})\right]+\frac{1}{4}\left[(C_{66})_{\frac12\frac12\frac12}(v_{110}-v_{001})-(C_{66})_{-\frac12-\frac12-\frac12}(v_{00-1}-v_{-1-10})\right]\\
& +\frac{1}{4}\left[(C_{55})_{\frac12\frac12\frac12}(v_{000}-v_{111})+(C_{55})_{-\frac12-\frac12-\frac12}(v_{000}-v_{-1-1-1})\right]+\frac{1}{4}\left[(C_{55})_{-\frac12\frac12\frac12}(v_{000}-v_{-111})+(C_{55})_{\frac12-\frac12-\frac12}(v_{000}-v_{1-1-1})\right]\\
& +\frac{1}{4}\left[(C_{55})_{\frac12\frac12\frac12}(v_{011}-v_{100})-(C_{55})_{-\frac12-\frac12-\frac12}(v_{-100}-v_{0-1-1})\right]+\frac{1}{4}\left[(C_{55})_{-\frac12\frac12\frac12}(v_{011}-v_{-100})-(C_{55})_{\frac12-\frac12-\frac12}(v_{100}-v_{0-1-1})\right],
\end{aligned}
$$

and similar expressions are obtained for $A_{11}^{(2)}$, $A_{11}^{(3)}$ and $A_{11}^{(4)}$. Here the subscript $\frac{1}{2}$ denotes the interpolation of the stiffness tensor, following the notation in Operto *et al.* (2007).

Conversely, we find the weighting for each gridpoint,

$$
\begin{aligned}
v_{000}: {} & \left(w_{s1}+\frac{w_{s2}}{3}\right)\left[(C_{11})_{-\frac12 00}+(C_{11})_{\frac12 00}+(C_{66})_{0-\frac12 0}+(C_{66})_{0\frac12 0}+(C_{55})_{00-\frac12}+(C_{55})_{00\frac12}\right]\\
& +\frac{w_{s2}}{12}\left[(C_{66})_{0-\frac12-\frac12}+(C_{66})_{0\frac12-\frac12}+(C_{66})_{0-\frac12\frac12}+(C_{66})_{0\frac12\frac12}(C_{55})_{0-\frac12-\frac12}+(C_{55})_{0\frac12-\frac12}+(C_{55})_{0-\frac12\frac12}+(C_{55})_{0\frac12\frac12}\right.\\
& +(C_{11})_{-\frac12 0-\frac12}+(C_{11})_{\frac12 0-\frac12}+(C_{11})_{-\frac12 0\frac12}+(C_{11})_{\frac12 0\frac12}+(C_{55})_{-\frac12 0-\frac12}+(C_{55})_{\frac12 0-\frac12}+(C_{55})_{-\frac12 0\frac12}+(C_{55})_{\frac12 0\frac12}\\
& \left.+(C_{11})_{-\frac12-\frac12 0}+(C_{11})_{\frac12-\frac12 0}+(C_{11})_{-\frac12\frac12 0}+(C_{11})_{\frac12\frac12 0}+(C_{66})_{-\frac12-\frac12 0}+(C_{66})_{\frac12-\frac12 0}+(C_{66})_{-\frac12\frac12 0}+(C_{66})_{\frac12\frac12 0}\right]\\
& +\frac{w_{s3}}{16}\left[(C_{11})_{-\frac12-\frac12-\frac12}+(C_{11})_{\frac12-\frac12-\frac12}+(C_{11})_{-\frac12\frac12-\frac12}+(C_{11})_{\frac12\frac12-\frac12}+(C_{11})_{-\frac12-\frac12\frac12}+(C_{11})_{\frac12-\frac12\frac12}+(C_{11})_{-\frac12\frac12\frac12}+(C_{11})_{\frac12\frac12\frac12}\right.\\
& +(C_{66})_{-\frac12-\frac12-\frac12}+(C_{66})_{\frac12-\frac12-\frac12}+(C_{66})_{-\frac12\frac12-\frac12}+(C_{66})_{\frac12\frac12-\frac12}+(C_{66})_{-\frac12-\frac12\frac12}+(C_{66})_{\frac12-\frac12\frac12}+(C_{66})_{-\frac12\frac12\frac12}+(C_{66})_{\frac12\frac12\frac12}\\
& \left.+(C_{55})_{-\frac12-\frac12-\frac12}+(C_{55})_{\frac12-\frac12-\frac12}+(C_{55})_{-\frac12\frac12-\frac12}+(C_{55})_{\frac12\frac12-\frac12}+(C_{55})_{-\frac12-\frac12\frac12}+(C_{55})_{\frac12-\frac12\frac12}+(C_{55})_{-\frac12\frac12\frac12}+(C_{55})_{\frac12\frac12\frac12}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
v_{100}: {} & -\left(w_{s1}+\frac{w_{s2}}{3}\right)(C_{11})_{\frac12 00}\\
& -\frac{w_{s2}}{12}\left[2(C_{11})_{\frac12-\frac12 0}+2(C_{11})_{\frac12\frac12 0}-(C_{66})_{\frac12-\frac12 0}-(C_{66})_{\frac12\frac12 0}-(C_{55})_{\frac12-\frac12 0}-(C_{55})_{\frac12\frac12 0}\right]\\
& -\frac{w_{s3}}{16}\left[(C_{55})_{\frac12-\frac12-\frac12}+(C_{55})_{\frac12\frac12-\frac12}+(C_{55})_{\frac12-\frac12\frac12}+(C_{55})_{\frac12\frac12\frac12}\right.\\
& \left.+(C_{66})_{\frac12-\frac12-\frac12}+(C_{66})_{\frac12\frac12-\frac12}+(C_{66})_{\frac12-\frac12\frac12}+(C_{66})_{\frac12\frac12\frac12}\right],
\end{aligned}
$$

while similar results are obtained for v_{-100}, $v_{0\,-10}$, v_{010}, $v_{00\,-1}$ and v_{001}. Moreover,

$$
\begin{aligned}
v_{110}: {} & -\frac{w_{s2}}{12}\left[(C_{11})_{\frac12\frac12 0}+(C_{66})_{\frac12\frac12 0}\right]\\
& -\frac{w_{s3}}{16}\left[(C_{11})_{\frac12\frac12-\frac12}+(C_{11})_{\frac12\frac12\frac12}+(C_{66})_{\frac12\frac12-\frac12}+(C_{66})_{\frac12\frac12\frac12}\right],
\end{aligned}
$$

while similar results are obtained for $v_{-1\,-10}$, $v_{1\,-10}$, v_{-110}, $v_{-10\,-1}$, $v_{10\,-1}$, v_{-101}, v_{101}, $v_{0\,-1\,-1}$, $v_{01\,-1}$, $v_{0\,-11}$ and v_{011}. Finally,

$$
v_{111}: -\frac{w_{s3}}{8}\left[(C_{11})_{\frac12\frac12\frac12}+(C_{66})_{\frac12\frac12\frac12}+(C_{55})_{\frac12\frac12\frac12}\right],
$$

while similar results are obtained for $v_{-1\,-1\,-1}$, $v_{1\,-1\,-1}$, $v_{-11\,-1}$, $v_{11\,-1}$, $v_{-1\,-11}$, $v_{1\,-11}$ and v_{-111}.