Massively parallel structured direct solver for equations describing time-harmonic qP-polarized waves in TTI media

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

[1] *Center for Computational and Applied Mathematics,*

Department of Mathematics, Purdue University,

150 N. University St, West Lafayette, IN 47907

[2] *Computational Research Division,*

Lawrence Berkeley National Laboratory,

MS 50F-1650, One Cyclotron Rd, Berkeley, CA 94720

(October 17, 2011)

Running head: **time-harmonic qP-polarized waves**

ABSTRACT

We consider the discretization and approximate solutions of equations describing time-harmonic qP polarized waves in 3D inhomogeneous anisotropic media. The anisotropy comprises general (tilted) TI symmetries. We are concerned with solving these equations on a large domain, for a large number of different sources. We consider higher (fourth) order partial differential equations and variable order finite difference schemes, to accommodate anisotropy on the one hand and allow higher order accuracy – to control sampling rates for relatively high frequencies – on the other hand. We make use of a nested dissection based domain decomposition in a massively parallel multifrontal solver combined with Hierarchically SemiSeparable (HSS) matrix compression techniques. The higher order

partial differential operators, and the variable order finite difference schemes, require the introduction of separators with variable thickness in the nested dissection; the development of these and their integration with the multifrontal solver is the main topic of this paper. The algorithm presented here is a powerful tool for (anisotropic) full waveform inversion.

INTRODUCTION

We consider the discretization and approximate solutions of equations describing time-harmonic qP polarized waves in 3D inhomogeneous TTI media. These equations arise upon (microlocally) operator diagonalizing the system of equations describing elastic waves – assuming it is of principal type – thus decoupling the polarizations (see, for example, Stolk and De Hoop (2002)), and then considering time-harmonic solutions. We extract the 'Helmholtz' equation for qP polarized constituents. In general, the 'Helmholtz' equations are *pseudodifferential* equations of second order, but in the isotropic case they reduce to the Helmholtz equations proper. The symbols of the pseudodifferential operators can be expanded to yield (low-rank) separated representations. Instead of developing such representations, here, we consider the re-coupling of qSV and qP polarizations yielding fourth-order *partial* differential equations in TI media; we then take the 'incomplete acoustic' limit along the symmetry axis introduced by Alkhalifah (2000) to emphasize the propagation of qP-polarized waves. (The limit is naturally 'complete' in the case of elliptic media, in which the mentioned recoupling can be avoided to begin with.)

We invoke the finite difference method. We are concerned with solving the resulting algebraic equations on a large subsurface domain, for a large number of different (surface or subsurface) sources in the context of modeling seismic wave propagation with applications in so-called (local optimization based) full waveform inversion (FWI) in mind. The key result of this paper is the development of a massively parallel structured direct solver allowing the use of essentially general finite-difference stencils accommodating anisotropy and relatively high frequencies.

The direct method of choice for solving this problem is the multifrontal factorization algorithm (see Liu (1992)). The central idea of the multifrontal algorithm (see also MUMPS, Agullo et al. (2008)) is to reorganize the sparse factorization of the discretized operator into a series of dense local factorizations. The algorithm is used together with the method of nested dissection (see 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 exploited to recursively divide the mesh into two disjoint subdomains of smaller size. 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 thickness of separators is determined by the pattern of the finite difference stencil ensuring a complete partitioning of the upper level domain into two disjoint lower level subdomains. The development of nested dissection with separators of variable thickness and the integration with the multifrontal solver, generalizing the original work of Wang et al. (2010, 2011), is the main topic of this paper. We follow the approach developed by Xia et al. (2009, 2010) of integrating the multifrontal method with structured matrices. Indeed, the fill-in blocks of the factorization appear to be highly compressible using the framework of Hierarchically SemiSeparable (HSS) matrices. The key issue, here, is the memory needed for the algorithm, while the accuracy of the solution is controlled and can be limited in the applications considered.

In 3D, we specifically consider a general 125-point finite-difference stencil for discretizing the relevant operator on a regular mesh. We invoke PML boundaries. We note that the resulting matrix is non-Hermitian, indefinite, relatively poorly conditioned, but has a symmetric pattern which we use in the nested dissection. We exploit the regularity of the mesh to arrive at a complete binary assembly tree. The 125-point finite-difference stencil allows us to solve the various equations proposed to propagate qP-polarized waves.

In the case of VTI symmetry, the Alkhalifah's 'incomplete' acoustic limit yields a scalar fourth-order partial differential equation which can be discretized with a 27-point finite-difference stencil (as in the isotropic case) admitting a separator thickness of 1. The structure of this equation no longer yields the 27-point stencil in the case of general TTI symmetry. Due to the underlying re-coupling, it generates, erroneous, quasi-shear waves, unless the medium is elliptic. The natural strategy is to embed the source in a (small) ball where the medium is elliptic, and, hence, the erroneous shear waves are not excited. If the coefficients are smooth, conversions from qP to shear are relatively weak.

Time-domain strategies rely on the construction of coupled pairs of partial differential equations which are second-order in time and are equivalent to the fourth-order equation. As pointed out by Duveneck and Bakker (2011), these can be obtained directly from the original system of equations generating elastic waves, restricting the stiffness tensor to TI symmetry and then taking Alkhalifah's limit. In fact, Duveneck and Bakker (2011) started from the (first-order) constitutive equations and equations of motion to derive such a system. We adapt our solver also to this system transformed to the frequency domain and test its efficiency. Other coupled pairs of equations which are second-order in time following *ad hoc* constructions, for VTI, can be found in Grechka et al. (2004) and in Zhou et al. (2006). The analogous construction for TTI can be found in Duveneck et al. (2008); Fletcher et al. (2009). The propagator approach (in time) has been implemented by Crawley *et al.* Crawley et al. (2010). In the frequency domain, considered here, there is no need for the introduction of coupled systems of partial differential equations to lower the order in time. The complication arises in the spatial part of the relevant operator which, in our algorithm, is addressed by introducing separators of variable thickness. We derive complexity and interprocessor communication estimates of our algorithm. In particular, we compare these for separator thickness 2 with the ones for separator thickness 1, and verify the estimates by numerical experiments. The estimates aid in exploiting the trade-off between the order of finite-difference approximation and matrix size for a given frequency and subsurface domain.

GENERAL EQUATIONS

Fourth-order partial differential equations

We consider the following scalar fourth-order partial differential equation,

$$\left[\mathit{\Gamma}(x, \partial_x, \omega) - \omega^2\right] u(x, \omega) = f(x, \omega), \quad x \in \mathbb{R}^3, \tag{1}$$

to describe the propagation of qP-polarized waves; here, $u(x, \omega)$ is a time-harmonic field excited by the forcing term $f(x, \omega)$, and

$$\begin{aligned}
-\mathit{\Gamma}(., \partial_x, .) = {} & a_{11}\frac{\partial^2}{\partial x_1^2} + a_{22}\frac{\partial^2}{\partial x_2^2} + a_{33}\frac{\partial^2}{\partial x_3^2} + a_{12}\frac{\partial^2}{\partial x_1 \partial x_2} + a_{23}\frac{\partial^2}{\partial x_2 \partial x_3} + a_{31}\frac{\partial^2}{\partial x_3 \partial x_1} \\
& + a_{1111}\frac{\partial^4}{\partial x_1^4} + a_{2222}\frac{\partial^4}{\partial x_2^4} + a_{3333}\frac{\partial^4}{\partial x_3^4} \\
& + a_{1122}\frac{\partial^4}{\partial x_1^2 \partial x_2^2} + a_{2233}\frac{\partial^4}{\partial x_2^2 \partial x_3^2} + a_{3311}\frac{\partial^4}{\partial x_3^2 \partial x_1^2} \\
& + a_{1112}\frac{\partial^4}{\partial x_1^3 \partial x_2} + a_{1222}\frac{\partial^4}{\partial x_1 \partial x_2^3} + a_{1113}\frac{\partial^4}{\partial x_1^3 \partial x_3} \\
& + a_{1333}\frac{\partial^4}{\partial x_1 \partial x_3^3} + a_{2223}\frac{\partial^4}{\partial x_2^3 \partial x_3} + a_{2333}\frac{\partial^4}{\partial x_2 \partial x_3^3} \\
& + a_{1123}\frac{\partial^4}{\partial x_1^2 \partial x_2 \partial x_3} + a_{2231}\frac{\partial^4}{\partial x_2^2 \partial x_3 \partial x_1} + a_{3312}\frac{\partial^4}{\partial x_3^2 \partial x_1 \partial x_2},
\end{aligned}$$

in which $a_{11}, \ldots, a_{3312}$ are general coefficients which can be dependent on x; the coefficients corresponding with (mixed) fourth-order derivatives are also dependent on ω^2. The principal (leading-order) symbol of this operator can reproduce the dispersion relation for qP

polarized elastic waves proposed by Alkhalifah (2000) and is discussed in the next section. Throughout this paper, the density of mass is assumed to be constant. How to incorporate variable density of mass is touched upon in the next section as well.

We introduce a Perfectly Matched Layer (PML) (Berenger (1994)) contained in the computational domain, $[0, L_1] \times [0, L_2] \times [0, L_3]$ say, where L_1, L_2 and L_3 indicate lengths. Then the damping function S_1 is defined as

$$S_1 = S_1(x_1, \omega) = \begin{cases} 1 & \text{if} \quad 0 \le x_1 \le L_{11}, \\ 1 - \mathrm{i}\dfrac{\sigma_0}{\omega}\cos^2\left(\dfrac{\pi}{2}\dfrac{x_1 - L_{11}}{L_1 - L_{11}}\right) & \text{if } L_{11} < x_1 \le L_1, \end{cases} \tag{2}$$

in which $\mathrm{i} = \sqrt{-1}$, and $0 < L_{11} < L_1$. Similar definitions hold for $S_2 = S_2(x_2, \omega)$ and $S_3 = S_3(x_3, \omega)$. Here, σ_0 is an appropriately chosen constant and has the same unit as the angular frequency ω (see, for example, Operto et al. (2007)). The PML, or complex scaling, is incorporated by adjusting the partial derivatives: $\frac{\partial}{\partial x_1}$ is replaced by $\frac{1}{S_1}\frac{\partial}{\partial x_1}$ and similarly for the partial derivatives with respect to x_2 and x_3. For example, the term

$$\frac{\partial^4}{\partial x_3^2 \partial x_1 \partial x_2} \qquad \text{becomes} \qquad \frac{1}{S_3}\frac{\partial}{\partial x_3}\left(\frac{1}{S_3}\frac{\partial}{\partial x_3}\left(\frac{1}{S_1}\frac{\partial}{\partial x_1}\left(\frac{1}{S_2}\frac{\partial}{\partial x_2}\right)\right)\right).$$

General algebraic equations

We introduce a regular mesh and lattice,

$$x_{1,i} = (i-1)\,h_1,\ x_{2,j} = (j-1)\,h_2,\ x_{3,k} = (k-1)\,h_3,$$
$$i = 1, \ldots, N_1,\ j = 1, \ldots, N_2,\ k = 1, \ldots, N_3,$$

with $h_1, h_2, h_3 \approx h$. We implement a basic centered finite-difference approach leading to a 125-point stencil. More sophisticated designs are possible which yield such a stencil. Each term in the Helmholtz operator needs to be treated separately.

We write $\mathit{\Gamma}(x, \partial_x, \omega) = \sum_{\mu=1}^{21} \mathit{\Gamma}^\mu(x, \partial_x, \omega)$ and standardly approximate

$$
\begin{aligned}
&[\mathit{\Gamma}^\mu(x, \partial_x, \omega) - \omega^2]\, u(x_{1,i}, x_{2,j}, x_{3,k}, \omega) \\
&\qquad\approx \sum_{m_1,m_2,m_3=-2}^{2} \mathcal{D}^\mu_{m_1 m_2 m_3}(i, j, k)\; u(x_{1,i} + m_1 h_1, x_{2,j} + m_2 h_2, x_{3,k} + m_3 h_3, \omega), \quad (3)
\end{aligned}
$$

yielding second-order accuracy. The $\mathcal{D}^\mu_{m_1 m_2 m_3}(i, j, k)$ are constructed from discretizing the first to fourth order derivatives; taking the x_1 direction for example:

$$
\left(\frac{1}{S_1}\frac{\partial u}{\partial x_1}\right)(x_{1,i}, x_{2,j}, x_{3,k}, \omega) \approx \frac{u(x_{1,i} + h_1, x_{2,j}, x_{3,k}, \omega) - u(x_{1,i} - h_1, x_{2,j}, x_{3,k}, \omega)}{2\; h_1\; S_1(x_{1,i}, \omega)},
$$

while

$$
\begin{aligned}
\left(\frac{1}{S_1}\frac{\partial}{\partial x_1}\left(\frac{1}{S_1}\frac{\partial u}{\partial x_1}\right)\right)(x_{1,i}, x_{2,j}, x_{3,k}, \omega) &\approx \frac{1}{h_1\; S_1(x_{1,i}, \omega)} \\
&\times \left(\frac{u(x_{1,i} + h_1, x_{2,j}, x_{3,k}, \omega) - u(x_{1,i}, x_{2,j}, x_{3,k}, \omega)}{h_1\; S_1(x_{1,i} + \frac{1}{2}h_1, \omega)}\right. \\
&\qquad \left. - \frac{u(x_{1,i}, x_{2,j}, x_{3,k}, \omega) - u(x_{1,i} - h_1, x_{2,j}, x_{3,k}, \omega)}{h_1\; S_1(x_{1,i} - \frac{1}{2}h_1, \omega)}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
&\left(\frac{1}{S_1}\frac{\partial}{\partial x_1}\left(\frac{1}{S_1}\frac{\partial}{\partial x_1}\left(\frac{1}{S_1}\frac{\partial u}{\partial x_1}\right)\right)\right)(x_{1,i}, x_{2,j}, x_{3,k}, \omega) \\
&\qquad\approx \frac{\mathcal{U}(x_{1,i} + h_1, x_{2,j}, x_{3,k}, \omega) - \mathcal{U}(x_{1,i} - h_1, x_{2,j}, x_{3,k}, \omega)}{2 h_1\; S_1(x_{1,i}, \omega)},
\end{aligned}
$$

and

$$
\begin{aligned}
\left(\frac{1}{S_1}\frac{\partial}{\partial x_1}\left(\frac{1}{S_1}\frac{\partial}{\partial x_1}\left(\frac{1}{S_1}\frac{\partial}{\partial x_1}\left(\frac{1}{S_1}\frac{\partial u}{\partial x_1}\right)\right)\right)\right)(x_{1,i}, x_{2,j}, x_{3,k}, \omega) &\approx \frac{1}{h_1\; S_1(x_{1,i}, \omega)} \\
&\times \left(\frac{\mathcal{U}(x_{1,i} + h_1, x_{2,j}, x_{3,k}, \omega) - \mathcal{U}(x_{1,i}, x_{2,j}, x_{3,k}, \omega)}{h_1\; S_1(x_{1,i} + \frac{1}{2}h_1, \omega)}\right. \\
&\qquad \left. - \frac{\mathcal{U}(x_{1,i}, x_{2,j}, x_{3,k}, \omega) - \mathcal{U}(x_{1,i} - h_1, x_{2,j}, x_{3,k}, \omega)}{h_1\; S_1(x_{1,i} - \frac{1}{2}h_1, \omega)}\right)
\end{aligned}
$$

where

$$\mathcal{U}(x_{1,i}+h_1,x_{2,j},x_{3,k},\omega) \approx \frac{1}{h_1\ S_1(x_{1,i}+h_1,\omega)} \times \left(\frac{u(x_{1,i}+2h_1,x_{2,j},x_{3,k},\omega)-u(x_{1,i}+h_1,x_{2,j},x_{3,k},\omega)}{h_1\ S_1(x_{1,i}+\frac{3}{2}h_1,\omega)} - \frac{u(x_{1,i}+h_1,x_{2,j},x_{3,k},\omega)-u(x_{1,i},x_{2,j},x_{3,k},\omega)}{h_1\ S_1(x_{1,i}+\frac{1}{2}h_1,\omega)} \right),$$

and

$$\mathcal{U}(x_{1,i},x_{2,j},x_{3,k},\omega) \approx \frac{1}{h_1\ S_1(x_{1,i},\omega)} \times \left(\frac{u(x_{1,i}+h_1,x_{2,j},x_{3,k},\omega)-u(x_{1,i},x_{2,j},x_{3,k},\omega)}{h_1\ S_1(x_{1,i}+\frac{1}{2}h_1,\omega)} - \frac{u(x_{1,i},x_{2,j},x_{3,k},\omega)-u(x_{1,i}-h_1,x_{2,j},x_{3,k},\omega)}{h_1\ S_1(x_{1,i}-\frac{1}{2}h_1,\omega)} \right),$$

and

$$\mathcal{U}(x_{1,i}-h_1,x_{2,j},x_{3,k},\omega) \approx \frac{1}{h_1\ S_1(x_{1,i}-h_1,\omega)} \times \left(\frac{u(x_{1,i},x_{2,j},x_{3,k},\omega)-u(x_{1,i}-h_1,x_{2,j},x_{3,k},\omega)}{h_1\ S_1(x_{1,i}-\frac{1}{2}h_1,\omega)} - \frac{u(x_{1,i}-h_1,x_{2,j},x_{3,k},\omega)-u(x_{1,i}-2h_1,x_{2,j},x_{3,k},\omega)}{h_1\ S_1(x_{1,i}-\frac{3}{2}h_1,\omega)} \right).$$

For example, in the term $\Gamma^{14}(x,\partial_x,\omega) = a_{1222}(x,\omega)\frac{\partial^4}{\partial x_1 \partial x_2^3}$ the coefficient $a_{1222}(x,\omega)$ is evaluated on the following ten grid points: $u(x_{1,i}+m_1h_1, x_{2,j}+m_2h_2, x_{3,k},\omega), m_1=\pm1, m_2=\pm2,\pm1,0$, based on the formula above.

Matrix equation. We apply the usual conversion from subscripts to a linear index,

$$\mathbf{u}_{(k-1)N_1N_2+(j-1)N_1+i}(\omega) = u(x_{1,i},x_{2,j},x_{3,k},\omega),$$

$$i=1,\ldots,N_1,\ j=1,\ldots,N_2,\ k=1,\ldots,N_3,$$

with $(N_1-1)\,h_1 = L_1$, $(N_2-1)\,h_2 = L_2$, $(N_3-1)\,h_3 = L_3$; in a similar fashion we obtain $\mathbf{f}(\omega)$ from $f(.,.,.,\omega)$. The vectors $\mathbf{u}(\omega)$ and $\mathbf{f}(\omega)$ are of size $N_1N_2N_3$. We cast the discretized equation in corresponding matrix form:

$$\mathbf{A}(\omega)\ \mathbf{u}(\omega) = \mathbf{f}(\omega). \tag{4}$$

Naturally, the sparse matrix $\mathbf{A}(\omega)$ is of size $(N_1N_2N_3) \times (N_1N_2N_3)$, and shares the same nonzero pattern for different values of ω. The matrix is indefinite, relatively poorly conditioned, non-Hermitian but has a symmetric pattern. Our approach addresses the complications associated with these properties. For a prescribed accuracy and given computational domain, N_1, N_2, N_3 grow linearly with increasing frequency, to keep the number of grid points per wavelength constant.

TRANVERSE ISOTROPY; MULTI-FREQUENCY – PROPAGATING WAVES

The propagation of singularities by the solution operator of equation (1) is governed by the solutions to the dispersion relation,

$$\mathit{\Gamma}(x, \mathrm{i}\xi, \omega) = \omega^2; \tag{5}$$

if we write $\xi = \omega p$, in which ξ denotes the wave vector and p the slowness vector, we find the equation defining the slowness surface, $\mathit{\Gamma}(x, \mathrm{i}p, 1) = 1$. This equation is quartic in the components of p, and typically generates two sheets. Now, we discuss the choice of coefficients in $\mathit{\Gamma}(x, \partial_x, \omega)$.

The polarized 'Helmholtz' equations

We revisit the original general polarized wave equations. The polarized (pseudodifferential) 'Helmholtz' equations are of the form (Stolk and De Hoop (2002))

$$[\omega^2 - A(x, -\mathrm{i}\partial_x)]\, u(x,\omega) = -f(x,\omega). \tag{6}$$

We will restrict the discussion in this section to the principal (leading-order) parts, $A^{\mathrm{prin}}(x,\xi)$, of the symbols of $A(x, -\mathrm{i}\partial_x)$. It is straightforward to extend the calculations to the sub-principal symbols involving derivatives of the stiffness tensor components; density variations can be incorporated then as well.

In the case of VTI symmetry, in the $(1,3)$-plane ($n = 2$), we have

$$A^{\mathrm{prin}}_{qP}(x,\xi) = \tfrac{1}{2}[(c_{11}(x) + c_{55}(x))\xi_1^2 + (c_{33}(x) + c_{55}(x))\xi_3^2] + \tfrac{1}{2}\sqrt{[(c_{11}(x) - c_{55}(x))\xi_1^2 + (c_{33}(x) - c_{55}(x))\xi_3^2]^2 - 4E(x)^2\xi_1^2\xi_3^2},$$

and

$$A^{\mathrm{prin}}_{qSV}(x,\xi) = \tfrac{1}{2}[(c_{11}(x) + c_{55}(x))\xi_1^2 + (c_{33}(x) + c_{55}(x))\xi_3^2] - \tfrac{1}{2}\sqrt{[(c_{11}(x) - c_{55}(x))\xi_1^2 + (c_{33}(x) - c_{55}(x))\xi_3^2]^2 - 4E(x)^2\xi_1^2\xi_3^2},$$

where

$$E^2 = (c_{11} - c_{55})(c_{33} - c_{55}) - (c_{13} + c_{55})^2.$$

Here, the c_{ij} are the stiffness moduli divided by density. (E^2 is directly related to the anisotropic parameter η introduced by Alkhalifah and Tsvankin (1995).) Both symbols lead to the symbol of a standard Helmholtz operator if $E^2 = 0$, the elliptic case.

Extension to $n = 3$, TTI parametrization by rotation. We replace ξ_1 in the expression presented in the previous subsection by $\xi_1^2 + \xi_2^2$. We then introduce the rotation matrix

$$R_{\theta,\varphi} = \begin{pmatrix} \cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\varphi & \cos\varphi & 0 \\ \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \end{pmatrix},$$

where $\theta = \theta(x)$ is the local tilt angle and $\varphi = \varphi(x)$ is the local azimuth. The direction of the symmetry axis is given by the third column of $R_{\theta,\varphi}$. We then generate

$$A^{\text{prin}}_{qP;\text{TTI}}(x,\xi) = A^{\text{prin}}_{qP}(x, R^{-1}_{\theta(x),\varphi(x)}\xi). \tag{7}$$

In view of the changing multiplicity of the qS waves, we cannot globally extract a qSV equation in 3D (see **?**).

TTI scalar partial differential equation

The TTI scalar *partial* differential equation is obtained by the re-coupling of qSV and qP polarizations, upon forming

$$-\frac{1}{\omega^2}\,[\omega^2 - A^{\text{prin}}_{qSV;\text{TTI}}(x,-\mathrm{i}\partial_x)][\omega^2 - A^{\text{prin}}_{qP;\text{TTI}}(x,-\mathrm{i}\partial_x)]u(.,\omega) = 0,$$

or (up to principal parts)

$$\begin{aligned} \Bigg[-\left(1 + 2(\epsilon-\delta)\frac{v_{pz}^2}{\omega^2}(H_1+H_2) \right) H_3 - (1+2\epsilon)(H_1+H_2) \\ - \frac{v_{sz}^2}{v_{pz}^2}(H_3+H_1+H_2) - \frac{v_{sz}^2}{\omega^2}(H_3^2 + (1+2\epsilon)(H_1+H_2)^2 + (2+2\delta)(H_1+H_2)H_3) \\ - \frac{\omega^2}{v_{pz}^2} \Bigg] u(.,\omega) = 0, \end{aligned} \tag{8}$$

and taking the limit of vanishing SV velocity along the symmetry axis, $v_{sz} \downarrow 0$; see also Fletcher et al. (2008, 2009). The forcing term can be incorporated by introducing $\omega^{-2}[\omega^2 -$

$A^{\mathrm{prin}}_{qSV}(x, -\mathrm{i}\partial_x)]f_{qP}$ as the right-hand side. Here,

$$H_1(.,\partial_x) + H_2(.,\partial_x) = (\cos^2\theta\,\cos^2\varphi + \sin^2\varphi)\,\frac{\partial^2}{\partial x_1^2} + (\cos^2\theta\,\sin^2\varphi + \cos^2\varphi)\,\frac{\partial^2}{\partial x_2^2}$$
$$+ \sin^2\theta\,\frac{\partial^2}{\partial x_3^2} - \sin^2\theta\,\sin 2\varphi\,\frac{\partial^2}{\partial x_1\partial x_2} - \sin 2\theta\,\sin\varphi\,\frac{\partial^2}{\partial x_2\partial x_3} - \sin 2\theta\,\cos\varphi\,\frac{\partial^2}{\partial x_3\partial x_1},$$
$$H_3(.,\partial_x) = \sin^2\theta\,\cos^2\varphi\,\frac{\partial^2}{\partial x_1^2} + \sin^2\theta\,\sin^2\varphi\,\frac{\partial^2}{\partial x_2^2} + \cos^2\theta\,\frac{\partial^2}{\partial x_3^2}$$
$$+ \sin^2\theta\,\sin 2\varphi\,\frac{\partial^2}{\partial x_1\partial x_2} + \sin 2\theta\,\sin\varphi\,\frac{\partial^2}{\partial x_2\partial x_3} + \sin 2\theta\,\cos\varphi\,\frac{\partial^2}{\partial x_3\partial x_1}.$$

We have made the substitutions

$$\begin{aligned} c_{11} &= v_{pz}^2(1+2\epsilon), \\ c_{33} &= v_{pz}^2, \\ c_{55} &= v_{sz}^2, \\ (c_{13}+c_{55})^2 &= (v_{pz}^2 - v_{sz}^2)^2 + 2\delta v_{pz}^2(v_{pz}^2 - v_{sz}^2), \end{aligned} \tag{9}$$

where $\epsilon = \epsilon(x)$ and $\delta = \delta(x)$ denote Thomsen's (1986) parameters, and $v_{pz} = v_{pz}(x)$ denotes the P-wave velocity along the symmetry axis. In the mentioned limit, $c_{55} = 0$ and $c_{13} = v_{pz}^2\sqrt{1+2\delta}$; one obtains the so-called acoustic equation associated with TTI media, which is of the form (1) with

$$\Gamma = -v_{pz}^2\left[\left(1 + 2(\epsilon-\delta)\frac{v_{pz}^2}{\omega^2}(H_1+H_2)\right)H_3 + (1+2\epsilon)(H_1+H_2)\right]. \tag{10}$$

The limit $v_{sz} \downarrow 0$ mimics the proper limit from an elastic to an acoustic equation, which holds up to the elliptic case. It does remove the qSH-polarized constituents in the general (TI) case ($c_{66} = v_{sz}^2(1+2\gamma) \downarrow 0$, γ denoting another Thomsen's parameter).

The slowness surface associated with eq.(8) consists of a qP and a qSV sheet, and is illustrated in Figure (**??**) left; in Figure (**??**) right, we illustrate the limit $v_{sz} \downarrow 0$ and identify the asymptotes appearing in the rational approximation. The slowness sheet for qSV is significantly deformed in view of the mentioned limit (reflecting that it is incomplete), however, the qP sheet remains intact; this observation was already made by Alkhalifah (1998) and Grechka et al. (2004).

It is straightforward to circumvent the excitation of the erroneous qSV waves: One embeds the source in a ball within which $\epsilon(x) = \delta(x)$ in which case the equation becomes a second-order equation for qP polarized waves. If the coefficients, ϵ, δ, v_{pz} are smooth, the conversion will be weak. (An anaologous argument and construction can be applied to equation (8) using a pseudodifferential projection operator associated with the qP polarization without taking the mentioned limit.) To mitigate numerical instability, we smoothly adapt the coefficients near the boundary, such that the medium is elliptic in the PML.

TTI coupled system of second-order partial differential equations

As discussed in the previous subsection, the incomplete acoustic limit does remove the qSH-polarized constituents. In the complete acoustic limit, the shear stress vanishes and the stress tensor becomes isotropic; one considers the pressure obtained from the diagonal of the stress matrix to obtain a scalar equation. In the incomplete acoustic limit, again, the shear stress vanishes; now, in the remaining normal stress one distinguishes the 'horizontal' components, σ_h, in the symmetry plane, and the 'vertical' component, σ_v, along the symmetry axis. These components satisfy a coupled system of two (second-order) partial differential equations, describing coupled qP-qSV polarized waves. The analysis was carried out by Duveneck and Bakker (2011). In the TTI case, up to principal parts,

$$\left[\omega^2 I - v_{pz}^2 \begin{pmatrix} -(1+2\epsilon)(H_1+H_2) & -\sqrt{1+2\delta}H_3 \\ -\sqrt{1+2\delta}(H_1+H_2) & -H_3 \end{pmatrix}\right] \begin{pmatrix} \sigma_h \\ \sigma_v \end{pmatrix} = -\begin{pmatrix} f_h \\ f_v \end{pmatrix}. \tag{11}$$

The result is naturally consistent with the previous approach: one can (microlocally) operator diagonalize the system and apply the re-coupling procedure to recover (10).

NESTED DISSECTION AND THE MULTIFRONTAL METHOD

In this section, we discuss a structured multifrontal solver together with a 3D nested dissection ordering. The basic ideas are similar to those in Wang et al. (2010, 2011), but we use separators of variable thickness. Thus, the massively parallel multifrontal factorization and solution methods in Wang et al. (2010, 2011) are used here also, and we only need to focus on the situation of separators with variable thickness.

To reduce the fill-in of a direct solver, the matrix $\mathbf{A}(\omega)$ in eq. (4) generally needs to be reordered prior to the factorization stage. The nested dissection reordering (see George (1973); Liu (1992)), which in principle can be regarded as hierarchical domain decompositions, has been proven the optimal reordering strategy which minimizes the fill-in under certain circumstances. In the process of nested dissection, one divides the mesh into subdomains and separators. A separator can be precisely defined as a set of grid points the removal of which divides the mesh into two disjoint subdomains. Since we use finite difference discretizations and regular meshes, each separator consists of t straight lines or t planes in the mesh, where t is called the thickness of the separator. In Wang et al. (2010, 2011), $t = 1$, because the FD stencil was 27-point compact. In the case of a 125-point FD stencil, $t = 2$ is used, because two subdomains can be fully disconnected provided that there are at least two layers of grid points on the separator. We can generalize our discussion into the following formula:

$$\text{The number of grid points in the 3D compact stencil is } (2t+1)^3. \tag{12}$$

We note that t is determined by the order of the partial differential operator and by the order of the accuracy of the FD stencil. The variable t plays a vital role in the parallel multifrontal solver, especially in the data communication stage. The larger t is, the larger the intermediate dense matrices are, and the more expensive the method is. See the detailed count in the next section.

We show the pattern of matrix $\mathbf{A}(\omega)$, for a 27-point stencil and a 125-point stencil, in Figure 3, for a $20 \times 20 \times 20$ mesh. We note that $t = 1$ (27-point stencil) yields a block-tri-diagonal system, while $t = 2$ (125-point stencil) yields a block-penta-diagonal system.

At the nested dissection level one, a z direction separator of thickness t divides the entire 3D mesh into two subdomains and the separator itself. The grid points associated with the subdomains are reordered prior to the ones associated with the separator. Figure **??** top illustrates the first level nested dissection. Figures **??** row two left and row two right display the reordered matrix patterns for $t = 1$ and $t = 2$, respectively. We note that the size of the submatrix associated with the separator (lower right corner) in the $t = 2$ case is twice as large as the size of the submatrix in the $t = 1$ case.

Then, each subdomain is recursively partitioned following the same rule. At the second level of the nested dissection, figure **??** middle illustrates that two y direction separators are introduced. The further reordered matrix patterns are displayed in figure **??** row three. Figure **??** bottom together with figure **??** row four display the nested dissection at level three, when four x direction separators are introduced.

After the nested dissection with a preset total number of levels $\mathbf{l}_{\max}$, which is chosen to allow sufficient levels of domain decomposition and is determined by the mesh size N_1, N_2 and N_3, the matrix $\mathbf{A}(\omega)$ in equation (4) is reordered into the pattern as illustrated by Figure **??**(left). At the same time, an assembly tree which is a postordered binary tree defining the order of the Gaussian elimination is also formed. Figure **??**(right) displays this assembly tree.

The parallel multifrontal solver together with HSS approximation introduced in Wang et al. (2010, 2011) is based on the traversal of the assembly tree, provided that the neighboring information is determined before carrying out any factorization. Such neighboring information is used to form the frontal matrices. Deciding the precise neighboring information can help minimize the factorization and solution cost.

We note that after the nested dissection, each separator is associated with one node on the assembly tree depicted in Figure **??**(right). Additionally, each separator i is uniquely determined by the three coordinates of its two end points on the diagonal. We denote them as X^i_{head}, X^i_{tail}, Y^i_{head}, Y^i_{tail}, Z^i_{head}, Z^i_{tail}. Figure **??** uses dots to illustrate the *head* and *tail* points representation of that piece of the mesh. We will use the x direction as an example to explain how to determine neighbors at each direction. We denote i as the current node whose neighbors are to be determined, and p as its possible neighbors. We point out that all the possible neighbors of i should be among its ancestors in the assembly tree. The condition for p to be a neighbor of i is

$$\left|X^p_{head} - X^i_{tail}\right| = 1 \text{ or } \left|X^i_{head} - X^p_{tail}\right| = 1,$$

If $t = 1$, then $X^p_{head} = X^p_{tail}$, which implies that there will be no x direction points reordering if we calculate the overlapping range between i and p. However, if $t > 1$, then $X^p_{head} < X^p_{tail}$, which implies that the separator is no longer a perfect 2D plane, but a 3D block with thickness t in the x direction. Hence, there will be an additional x direction reordering if we calculate the overlapping range between i and p. This makes the calculation more expensive when we have larger t. Furthermore, when we calculate the higher level neighbors, the separator i together with lower level calculated neighboring information should be projected onto higher levels (see figure **??**). If $t = 1$, then the corner region should be either a line or a point. However when $t > 1$, the corner regions are all 3D blocks. These regions may be shared by several separators, and the positions in each separator should be carefully identified.

PERFORMANCE

With the nested dissection ordering, the exact factorization of the matrix $\mathbf{A}(\omega)$ in 3D requires $\mathcal{O}(n^2)$ flops and $\mathcal{O}(n^{4/3})$ storage (see Xia (2010); Wang et al. (2011)), where $n = N_1 \times N_2 \times N_3$. The counts are improved with the incorporation of HSS matrix structures. In the following, we will simplify the expressions by assuming that $N_1 \approx N_2 \approx N_3 \approx N$.

Costs: VTI versus TTI

Here, we address the complexities for solving different problems, namely VTI and TTI, using variable t and with the same order of accuracy. We note that the size of each frontal matrix for the TTI case (in the case of the 125-point stencil, $t = 2$) is t times larger than the one for the VTI or isotropic case (in the case of a 27-point stencil, $t = 1$), which is illustrated by Figure **??**. Thus the storage for TTI is t^2 times larger than the storage for VTI. With the exact LU factorization of each frontal matrix during the process of the multifrontal solver, the complexity for TTI is t^3 times larger than the complexity for VTI, due to the fact that the complexity for the exact LU factorization is $\mathcal{O}(m^3)$ in which m is the size of each dense frontal matrix. Xia et al. (2010) conclude that the complexities associated with the HSS construction and factorization are $\mathcal{O}(rm^2)$ and $\mathcal{O}(rm)$, respectively, in which r is the largest off-diagonal rank of the HSS matrix. Therefore, with the HSS construction and factorization of each frontal matrix, the complexity for TTI is t^2 rather than t^3 times larger than the complexity for VTI. Because our structured multifrontal solver is a hybrid of both exact LU and HSS factorizations (see Wang et al. (2011)), the entire complexity associated with TTI is hence between t^2 and t^3 times larger than the complexity associated with VTI.

We present the strong scaling (variable number of processors for a fixed problem) for a fixed 3D $128 \times 128 \times 128$ mesh, for both the isotropic/VTI ($t = 1$) and TTI ($t = 2$) cases. The CPU wall time is collected in table **??** with the number of processors varying from 32 to 512. We note that the wall time associated with the TTI case is around five times larger than the wall time associated with the VTI case for a fixed 3D problem using the same number of processors, which is consistent with the analysis we carried out. Figure **??**(a) shows the strong scaling curve and figure **??**(b) displays the parallel efficiency. We note that both $t = 1$ and $t = 2$ scale much the same, and that we achieve high parallel efficiency, above 70%, while the parallel efficiency for $t = 2$ is slightly higher than for $t = 1$.

Coupled system of second-order equations. We point out that two formulations of the TTI system in the frequency domain (equations (1), (10) and (11)) are computationally equivalent. Although (11) is of second order rather than of fourth order, as (1), (10), the number of unknowns for equation (11) is $2N_1N_2N_3$ rather than $N_1N_2N_3$ as for (1), (10), which balances out the computational cost. We further observe that both equations yield dense frontal matrices of exactly the same size, and hence they have the same computational cost using the structured multifrontal solver. However, we gain accuracy with formulation (11) for the same complexity; for the stencils considered, (11) yields fourth-order accuracy while (1), (10) is accurate only to second order.

Cost: order versus sampling rate

Here, we address the complexities for solving a problem while increasing accuracy, by increasing order or sampling rate. We first revisit the cost of the multifrontal method with HSS structures, which can be roughly analyzed as follows (see Xia et al. (2009)). First, we count the costs for $t = 1$, and then generalize the results to a variable t. The costs associated with the separators are listed in Table , where the counts in Xia (2010) are used.

	traditional factorizations	structured factorizations
$\mathbf{l}_{\max} = \mathcal{O}(\log_2 N)$ levels	$\mathbf{l} - \mathbf{l}_s$ bottom levels	$\mathbf{l}_s$ upper levels
Each level $\mathbf{l} = 0, 1, \ldots, \mathbf{l}_{\max}$	$2^{\mathbf{l}}$ separators, each of size $\mathcal{O}(N/2^{\lfloor \mathbf{l}/2 \rfloor})$	
Cost (each separator)	$\mathcal{O}((N/2^{\lfloor \mathbf{l}/2 \rfloor})^3)$	$\mathcal{O}(r(N/2^{\lfloor \mathbf{l}/2 \rfloor})^2)$
Cost (subtotal)	$\sum_{\mathbf{l}=\mathbf{l}_s+1}^{\mathbf{l}_{\max}} 2^{\mathbf{l}} \mathcal{O}((N/2^{\lfloor \mathbf{l}/2 \rfloor})^3)$	$\sum_{\mathbf{l}=0}^{\mathbf{l}_s} 2^{\mathbf{l}} \mathcal{O}(r(N/2^{\lfloor \mathbf{l}/2 \rfloor})^2)$

Table 1: Flop count for the multifrontal method with intermediate HSS operations.

Thus, we have the total cost for our algorithm

$$
\begin{aligned}
\mathcal{C}_{\text{fact}} &= \sum_{\mathbf{l}=\mathbf{l}_s+1}^{\mathbf{l}_{\max}} 8^{\mathbf{l}} \mathcal{O}((N/2^{\mathbf{l}})^6) + \sum_{\mathbf{l}=0}^{\mathbf{l}_s} 8^{\mathbf{l}} \mathcal{O}((N/2^{\mathbf{l}})^4) \\
&\approx \mathcal{O}(N^6(2^{-3\mathbf{l}_s} - 2^{-3\mathbf{l}_{\max}})) + \mathcal{O}((2 - 2^{-\mathbf{l}_s}) N^4 \log_2 N)
\end{aligned}
$$

We choose $\mathbf{l}_s$ so that

$$
\mathcal{O}(N^6(2^{-3\mathbf{l}_s} - 2^{-3\mathbf{l}_{\max}})) = \mathcal{O}((2 - 2^{-\mathbf{l}_s}) N^4 \log_2 N).
$$

That is,

$$
\mathbf{l}_s \approx \mathcal{O}(\log_2 N) - \mathcal{O}(\log_2 \log_2 N).
$$

In this situation, we have

$$
\mathcal{C}_{\text{fact}}^1 = \mathcal{O}(N^4 \log_2 N) = \mathcal{O}(n^{4/3} \log_2 n).
$$

Similarly, we can show that the solution cost is

$$\mathcal{C}^1_{\text{sol}} = \mathcal{O}(n \log_2 n)$$

and that the storage requirement is

$$\mathcal{S}^1_{\text{mem}} = \mathcal{O}(n \log_2 n).$$

For a separator with variable thickness, or with t layers of single planes, the number of mesh points in a separator increases by a factor of t. We can simply replace N in the above by tN, and have

$$\mathcal{C}^t_{\text{fact}} \approx t^4 \mathcal{C}^1_{\text{fact}}, \ \mathcal{S}^t_{\text{mem}} \approx t^3 \mathcal{S}^1_{\text{mem}}. \tag{13}$$

For example, if $t = 2$, then $\mathcal{C}^t_{\text{fact}} \approx 16\mathcal{C}^1_{\text{fact}}, \ \mathcal{S}^t_{\text{mem}} \approx 8\mathcal{S}^1_{\text{mem}}$.

In our method, the error is proportional to N^{-r}, if r is the order of the scheme. Since $h = O(N^{-1})$ is small when N is large, it is preferrable to increase r so as to increase the accuracy, which leads to the increase of t. This is because, if otherwise we increase the sampling rate so that the mesh dimension becomes tN, then the number of mesh points in a separator increases by a factor of roughly t^2. Thus, we can similarly show that the factorization cost and the storage become $t^8\mathcal{C}^1_{\text{fact}}$ and $t^6\mathcal{S}^1_{\text{mem}}$. These are much higher than the counts in (13), especially when t is large.

NUMERICAL EXPERIMENTS

In the first example, we show the 3D time-harmonic wavefields in VTI and TTI homogeneous media using both the coupled system (11) and the fourth-order scalar equation (1), (10), computed on a $151 \times 151 \times 151$ mesh with the step size $h_1 = h_2 = h_3 = 30$m. The P-wave velocity along the symmetry axis is 3000 m/s, $\epsilon = 0.25, \delta = 0.1, \theta = 45^\circ$ and $\omega/2\pi = 10$

Hz. That is, we use a sampling rate of about 10 points per wave length. The explosive point source is located at the center of the domain. The homogeneous medium is smoothly tapered to be elliptic ($\epsilon = \delta$) in the PML region (see figure **??**) to avoid the PML instability for the TTI case (see also Operto et al. (2009)). We set the total level of the nested dissection to $\mathbf{l}_{\max} = 15$, and the HSS compression is of four-digit accuracy which we preserve for all the following numerical examples. Figures (**??**a) and (**??**b) show the fields computed in the VTI and TTI medium respectively, using (11). Figure (**??**c) and (**??**d) show the fields computed in the VTI and TTI medium respectively, using (1), (10). Figure (**??**e) displays the true amplitude difference between Figures (**??**a) and (**??**c) in the VTI medium, while Figure (**??**f) displays the true amplitude difference between Figures (**??**b) and (**??**d) in the TTI medium. The results are in agreement with one another. Furthermore, the wall time for solving the coupled TTI system is 1249s using 512 cores, while the wall time for solving the scalar TTI equation is 1235s also using 512 cores. However, we could have reduced the sampling rate in the case of the coupled TTI system to about 5 points per wave length and thus have reduced the matrix size by about a factor 2^3 and the wall time accordingly.

We compare the relative accuracy of the TTI and VTI simulations in Figure **??**, by rotating the TTI result $-45°$ to the orientation of the symmetry axis of the VTI medium. Figure **??**(a) displays the 2D VTI slice extracted from figure (**??**c). The 2D TTI slice extracted from figure (**??**d) after the rotation is shown in figure **??**(b). We show the amplitude difference between these extracted 2D slices in figure **??**(c). The amplitude of the green dashed line in figure **??** (a) and (b) is plotted in figure **??**(d). We note that our simulation is at least three digits accurate for both VTI and TTI simulations in the frequency domain.

We test the relative efficiency between separator thicknesses using the BP2007 TTI model. We generate an isotropic counterpart, BP2007 ISO, of this model by using the P-wave velocity along the symmetry axis. Figure **??** (top) displays the P-wave velocity along the symmetry axis; $\epsilon - \delta$ is shown in Figure **??** (middle). Figure (**??**) (bottom) displays the angle of the symmetry axis measured away from the vertical direction. The step size is $h_3 = h_1 = 12.5$m. As an example we take a frequency of 5 Hz, and a 1801×12596 mesh, with $\mathbf{l}_{\max} = 19$. We compute the fields on 64 cores; the computation for the ISO case is 68s while the computation time for the TTI case is 243s. We confirm the estimates given in the previous section. We also show the results on a 1801×5097 mesh, in Figure **??** (top) for the ISO case and in Figure **??** (bottom) for the TTI case. The explosive point source is located at $(2.5, 18)$km. We note that the PML and the strategy to suppress erroneous qSV waves work satisfactorily, even in regions with large variations in $\epsilon - \delta$.

Finally, we show the time-harmonic field computations in a 3D TTI model provided by PGS. We conduct the computations on a $241 \times 241 \times 241$ mesh with $\mathbf{l}_{\max} = 18$ and the spatial step sizes $h_1 = h_2 = h_3 = 25$m, yielding the subsurface domain $[-3, 3]$km $\times$ $[-3, 3]$km $\times$ $[0, 6]$km. Figure **??** top displays the Thomsen's parameter ϵ; Figure **??** bottom displays the Thomsen's parameter δ. The tilt angle θ and the azimuth angle φ are shown in Figure **??** top and bottom respectively. Figure **??** top displays the P-wave velocity along the symmetry axis. We carry out three types of time-harmonic simulations. Figure **??** bottom displays a 10 Hz time-harmonic wavefield computed on the isotropic model using only the P-wave velocity along the symmetry axis, with the explosive point source location at $(0, 0, 1.25)$km. Figure **??** top shows a 10 Hz time-harmonic wavefield computed on the VTI model using the P-wave velocity, ϵ and δ. Figure **??** bottom displays a 10 Hz time-harmonic wavefield computed on the TTI model using all the model parameters. We note the uplift of the wavefront in the TTI simulation due to the rotation of the symmetry axis in 3D. We used

512 cores on `hopper.nersc.gov` to carry out the isotropic and VTI simulations, with wall time 1792s, and 1024 cores to carry out the TTI simulation, with wall time 7314s.

DISCUSSION

We developed a massively parallel structured direct solver for fourth-order partial equations describing time-harmonic qP-polarized waves in TTI media. We invoke the finite-difference method. The re-ordering of the relevant matrix follows a nested dissection based domain decomposition. The order of the equations necessitates the introduction of separators of variable thickness. The construction and implementation of these, integrated with our massively parallel structured direct solver, comprise the main results of this paper. In numerical examples in 3D, we find that with our choice of finite difference scheme, the computational complexity associated with the TTI case is about five times larger than the one associated with the VTI or isotropic case. We note that, with our algorithm, it is possible to exploit the tradeoff between matrix size and separator thickness in the framework of higher-order finite difference schemes for a given accuracy. The incorporation of anisotropy in RTM and FWI has been widely recognized as important in real-world applications.

ACKNOWLEDGEMENTS

The authors thank the members ConocoPhillips, ExxonMobil, Total, PGS and Statoil of the Geo-Mathematical Imaging Group (GMIG) for partial financial support. Especially, they thank PGS for providing the 3D TTI model. The authors also thank the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National

Laboratory (LBNL) for providing the computing resources, and Tariq Alkhalifah for his invaluable comments and suggestions.

REFERENCES

Agullo, E., A. Guermouche, and J. L'Excellent, 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 Computing, **34**, 296–317.

Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631.

———, 2000, An acoustic wave equation for anisotropic media: Geophysics, **65**, 1239–1250.

Alkhalifah, T., and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: Geophysics, **60**, 1550–1566.

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

Crawley, S., S. Brandsberg-Dahl, and J. McClean, 2010, 3D TTI RTM using the pseudo-analytic method: SEG expanded abstract, **29**, 3216–3220.

Duveneck, E., and P. Bakker, 2011, Stable p-wave modeling for reverse-time migration in tilted ti medium: Geophysics, **76**, S65–S75.

Duveneck, E., P. Milcik, and P. Bakker, 2008, Acoustic VTI wave equations and their application for anisotropic reverse-time migration: SEG expanded abstract, **27**, 2186–2190.

Fletcher, R., X. Du, and P. Fowler, 2008, A new pseudo-acoustic wave equation for TI media: SEG expanded abstract, **27**, 2082–2086.

———, 2009, Stablizing acoustic reverse-time migration in TTI medium: SEG expanded abstract, **28**, 2985–2989.

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

Grechka, V., L. Zhang, and J. Rector, 2004, Shear waves in acoustic anisotropic media: Geophysics, **69**, 576–582.

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

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

Operto, S., J. Virieux, A. Ribodetti, and J. Anderson, 2009, Finite-difference frequency-domain modeling of viscoacoustic wave propagation in 2D Tilted Transversely Isotropic (TTI) media: Geophysics, **74**, T75–T95.

Stolk, C., and M. De Hoop, 2002, Microlocal analysis of seismic inverse scattering in anisotropic elastic media: Communications on Pure and Applied Mathematics, **55**, 261–301.

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

Wang, S., M. De Hoop, and J. Xia, 2010, Seismic inverse scattering via Helmholtz operator factorization and optimization: Journal of Computational Physics, **229**, 8445–8462.

———, 2011, On 3D modeling of seismic wave propagation via a structured parallel multifrontal direct helmholtz solver: Geophysical Prospecting, **59**, 857–873.

Xia, J., 2010, On the complexity of some hierarchical structured matrices: sumitted to SIAM. J. Matrix Anal. Appl, `http://www.math.purdue.edu/~xiaj/hsscost.pdf`.

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

———, 2010, Fast algorithms for hierarchically semiseparable matrices: Numer. Linear Algebra Appl., **2010**, 953–976.

Zhou, H., G. Zhang, and R. Bloor, 2006, An anisotropic acoustic wave equation for modeling and migration in 2d TTI media: SEG expanded abstract, **25**, 194–198.

Nprocs	32	64	128	256	512
$t = 2$, TTI (s)	4483	2384	1411	810	498
$t = 1$, isotropic/VTI (s)	821	449	281	163	102
ratio	5.46	5.31	5.02	4.97	4.88

Table 2: Strong Scaling for a fixed 3D mesh $128 \times 128 \times 128$, for both isotropic/VTI case ($t = 1$) and TTI case ($t = 2$).

LIST OF TABLES AND FIGURES

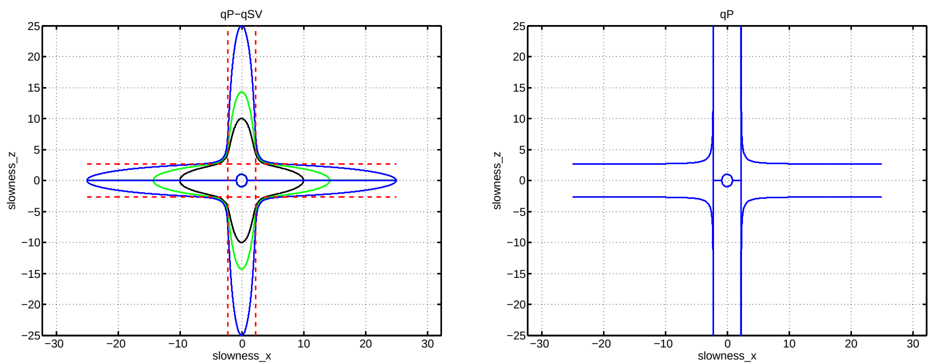


Figure 1: The sheets of the slowness surface associated with eq.(8); left: qP-qSV for $\epsilon = 0.2$, $\delta = 0.1$, and three different values of v_{sz}/v_{pz}: $0.1, 0.04, 0.07$; right: the limit $v_{sz} \downarrow 0$.

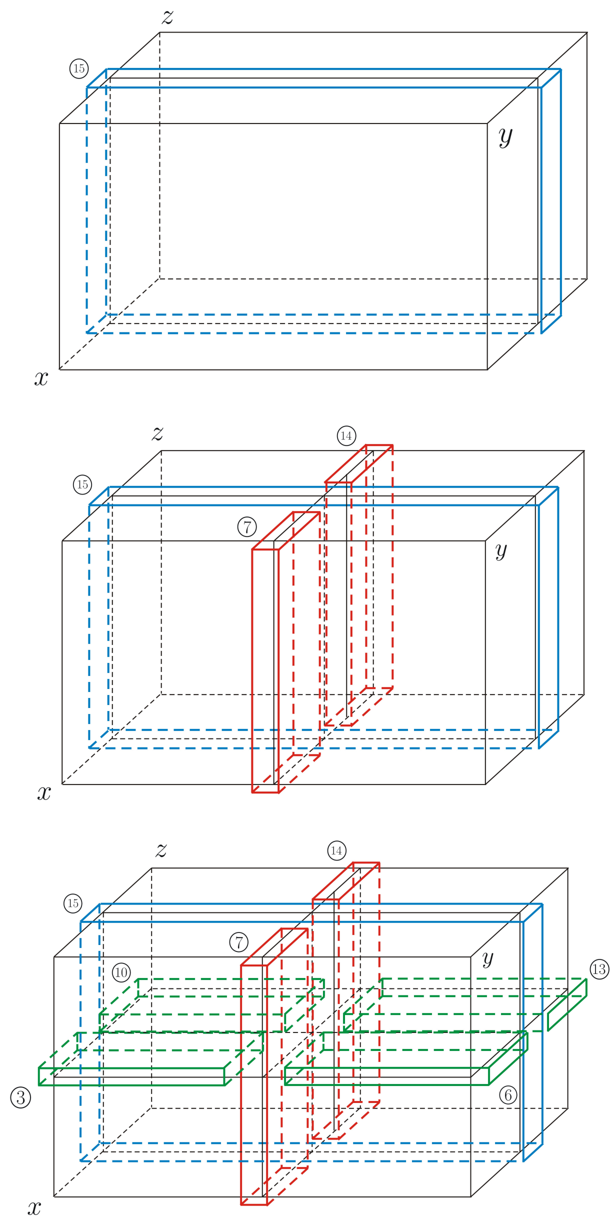


Figure 2: 3D Nested dissection with separators of variable thickness for different levels. **top**: level one; **middle**: level two; **bottom**: level three.

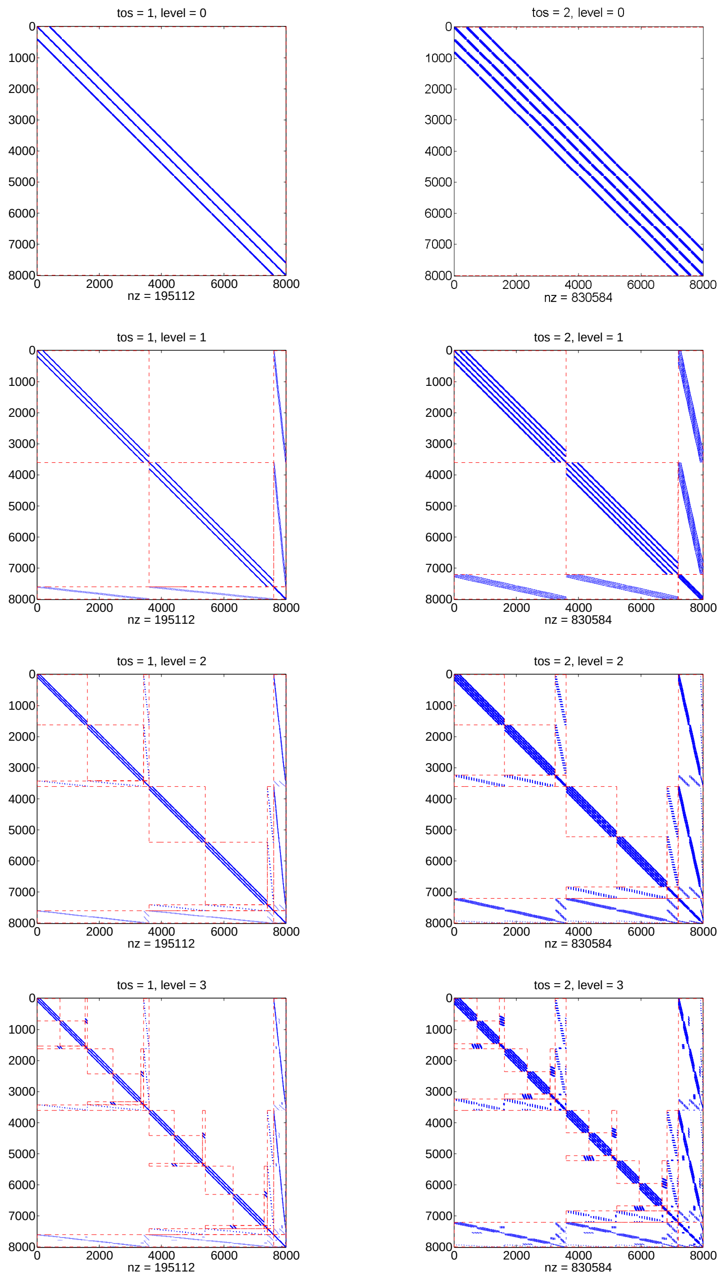


Figure 3: The pattern of the 3D matrix $\mathbf{A}(\omega)$ in eq.(4) discretized on a $20 \times 20 \times 20$ mesh, after nested dissection reordering for different levels and for variable thickness of separators (tos). **left column**: $tos = 1$; **right column**: $tos = 2$. **row one**: $level = 0$; **row two**: $level = 1$; **row three**: $level = 2$; **row four**: $level = 3$.

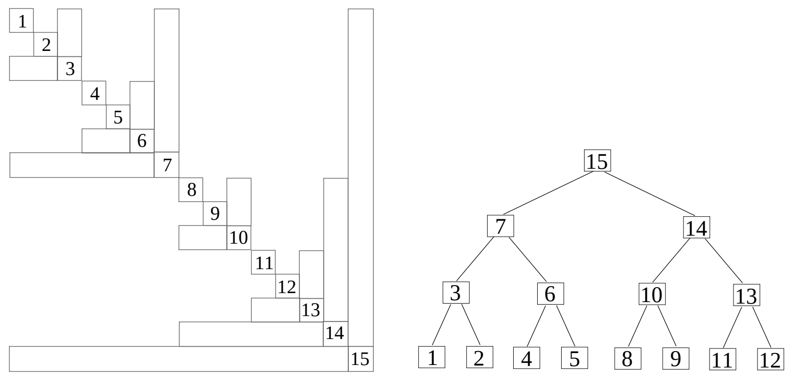


Figure 4: **Left**: the illustration of the matrix pattern after the nested dissection reordering displayed in figures (**??**) and (**??**); **right**: the corresponding assembly tree after the nested dissection.

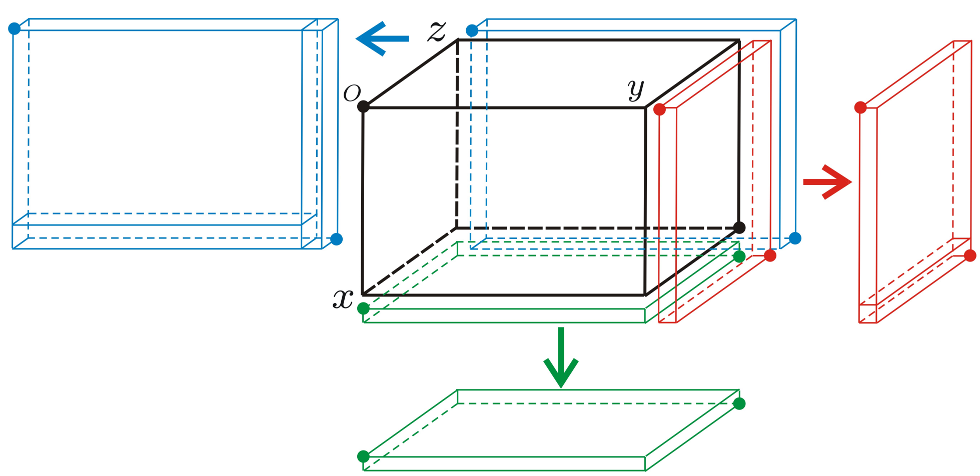


Figure 5: Neighbor determination in the 3D nested dissection with separators of variable thickness illustrated in figure (**??**).

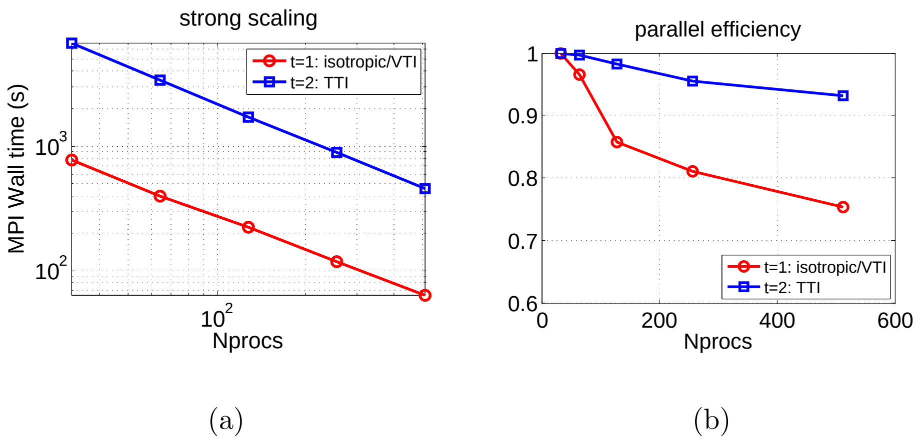


Figure 6: (a): strong scaling; (b) parallel efficiency; data is collected from table **??**.

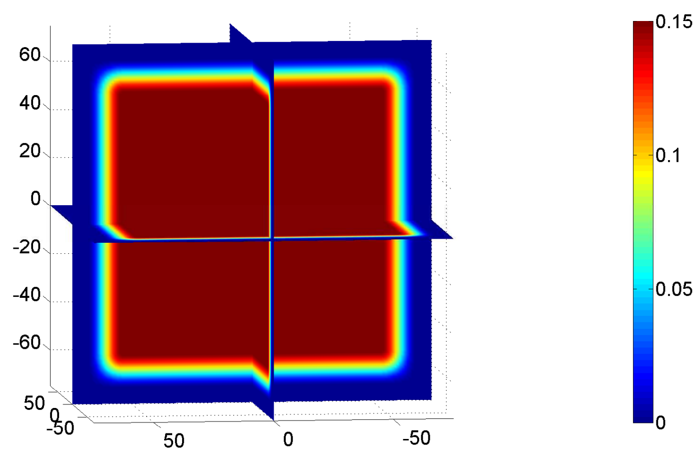


Figure 7: $\epsilon - \delta$ which is smoothly tapered to be zero in the PML region.

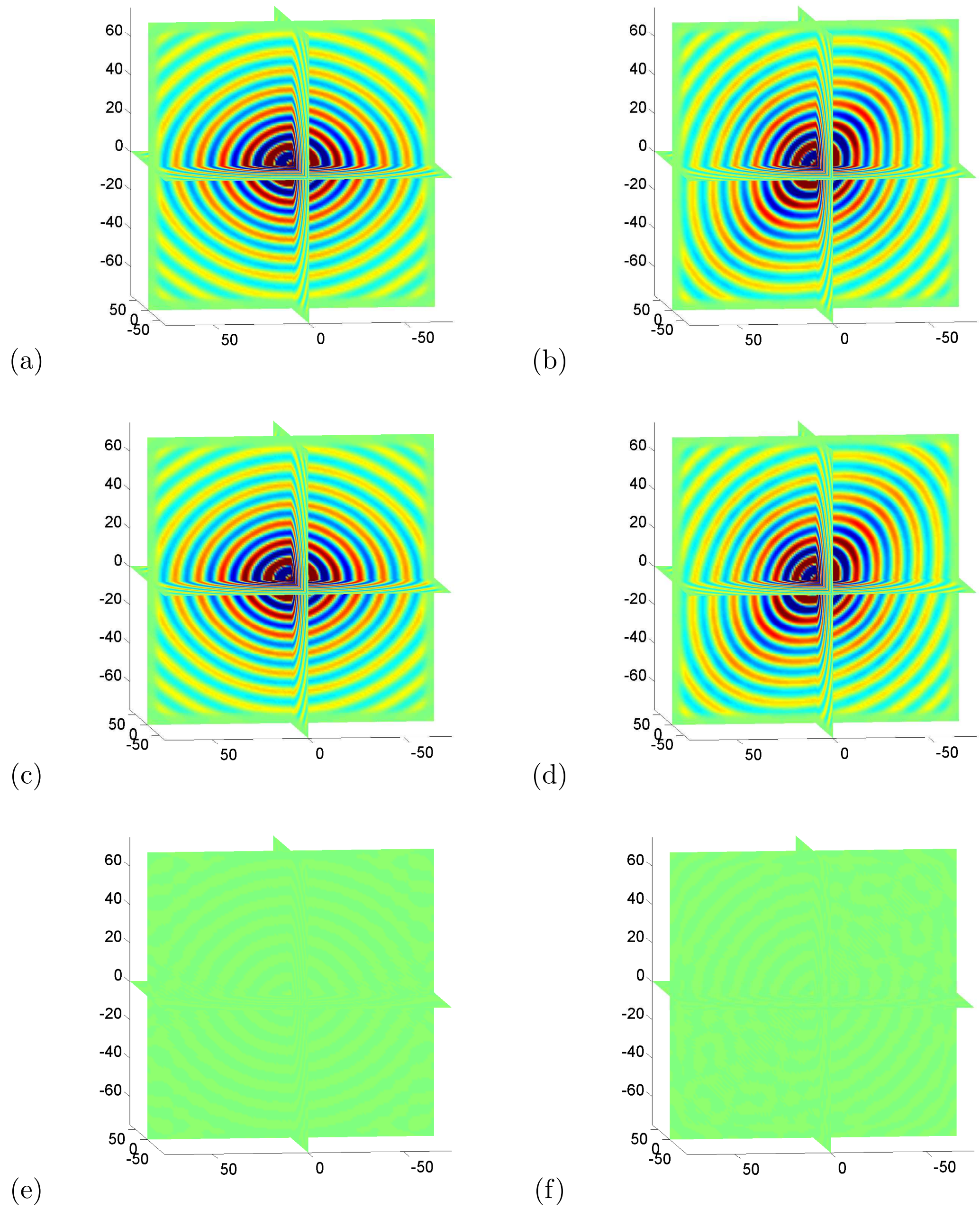
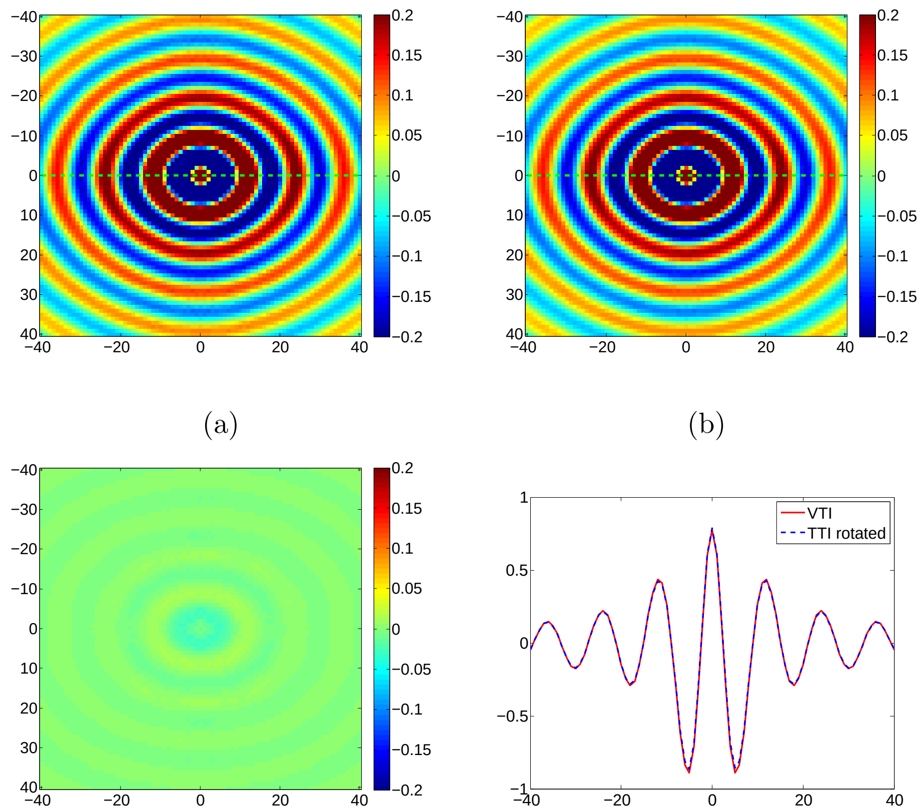


Figure 8: 3D time-harmonic wavefields of the TI medium: $Vp = 3000$m/s, $\epsilon = 0.25$, $\delta = 0.1$, $\theta = 45^\circ$, $\omega/2\pi = 10Hz$, $h_1 = h_2 = h_3 = 30$m; The seismic source is at the center; (a): VTI pressure using eq.(11); (b): TTI pressure using eq.(11); (c): VTI pressure using eq.(10); (d): TTI pressure using eq.(10); (e): the true amplitude difference between (a) and (c); (f): the true amplitude difference between (b) and (d).



(a) (b)

(c) (d)

Figure 9: (a): 2D VTI slice extracted from figure (??c); (b): 2D TTI slice extracted from figure (??d), which is rotated to the orientation of the symmetry axis of the VTI medium; (c): the amplitude difference between (a) and (b); (d): the amplitude of the green dashed line in (a) and (b).

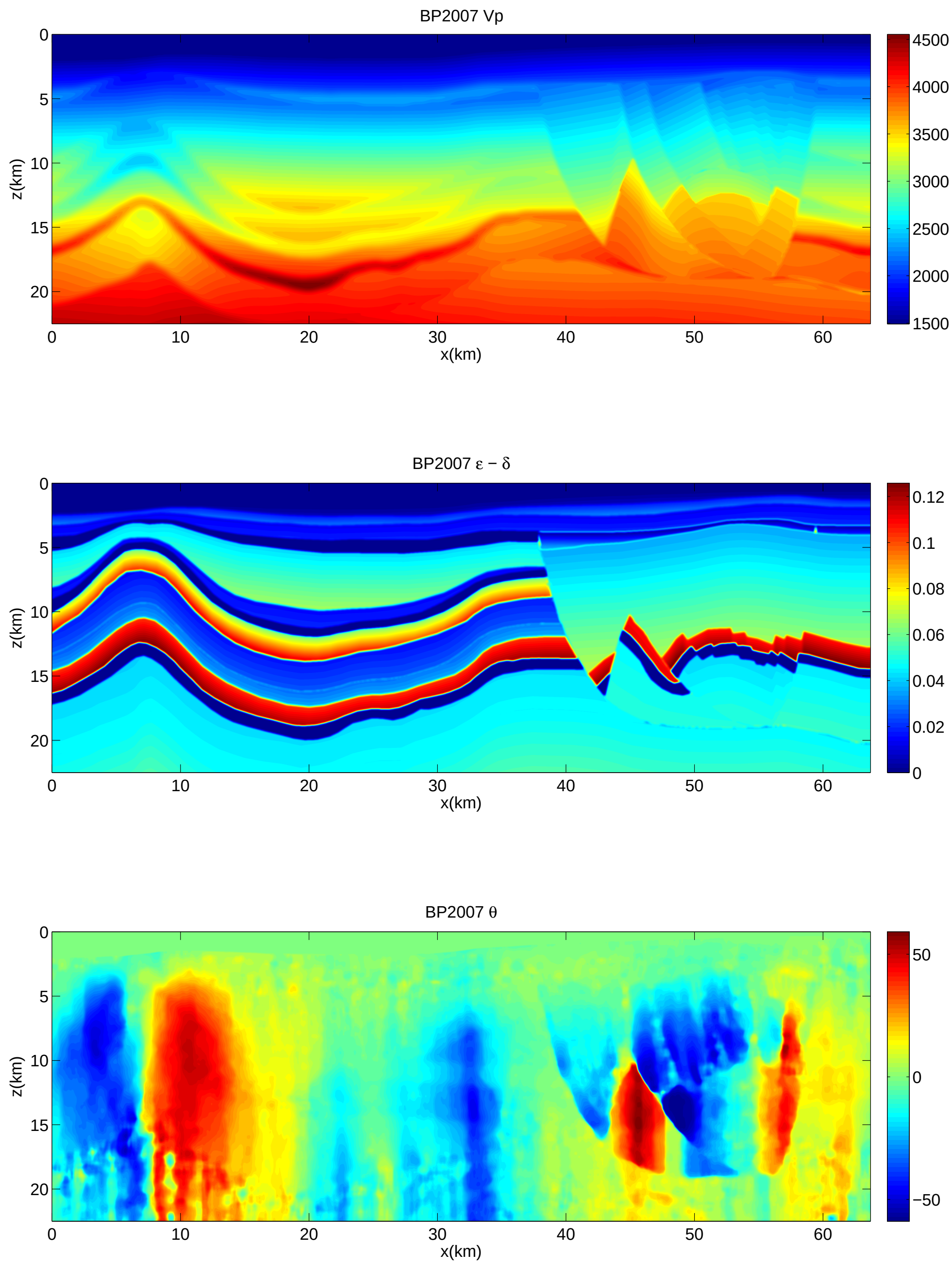
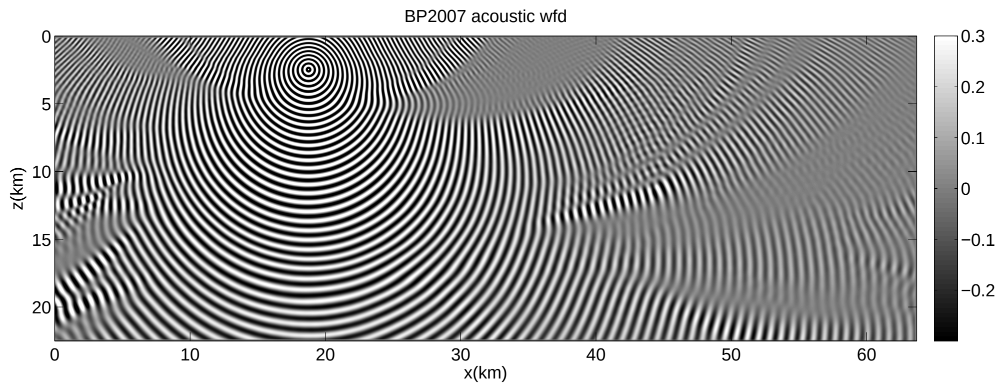


Figure 10: Part of the BP2007 TTI model; **top**: P-wave velocity along the symmetry axis; **middle**: $\epsilon - \delta$; **bottom**: the angle of the symmetry axis measured away from z direction.



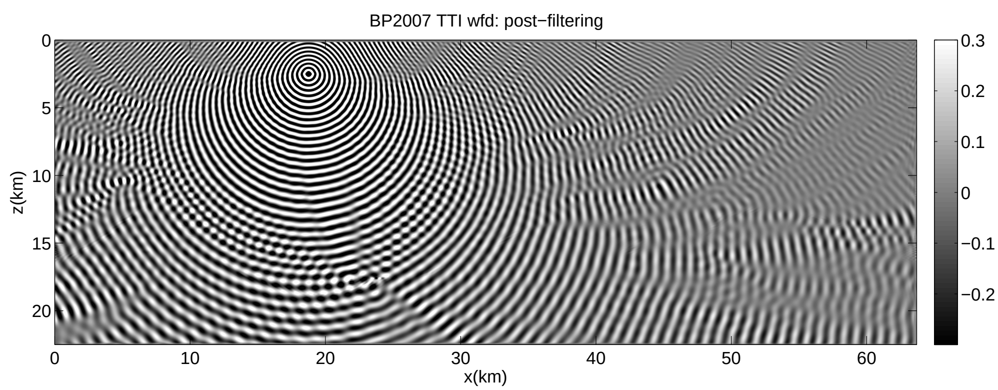


Figure 11: 5Hz time-harmonic wavefields computed in the model shown in Figure **??** and its isotropic counterpart, with the source located at $(2.5\text{km}, 18\text{km})$. Top: the wavefield in the ISO model. Bottom: the wavefield in the TTI model.

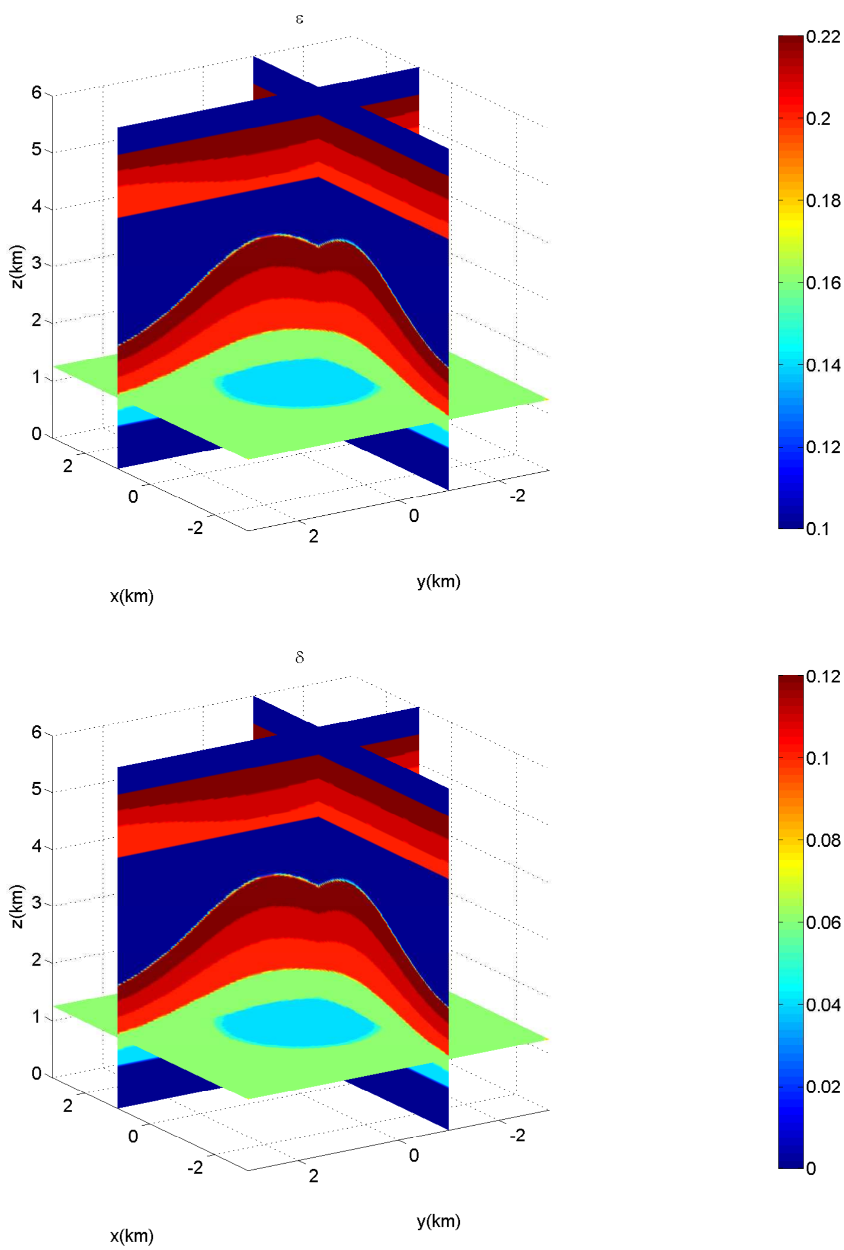


Figure 12: 3D PGS TTI model discretized on a $241 \times 241 \times 241$ mesh; **top**: Thomsen's parameter ϵ; **bottom**: Thomsen's parameter δ.

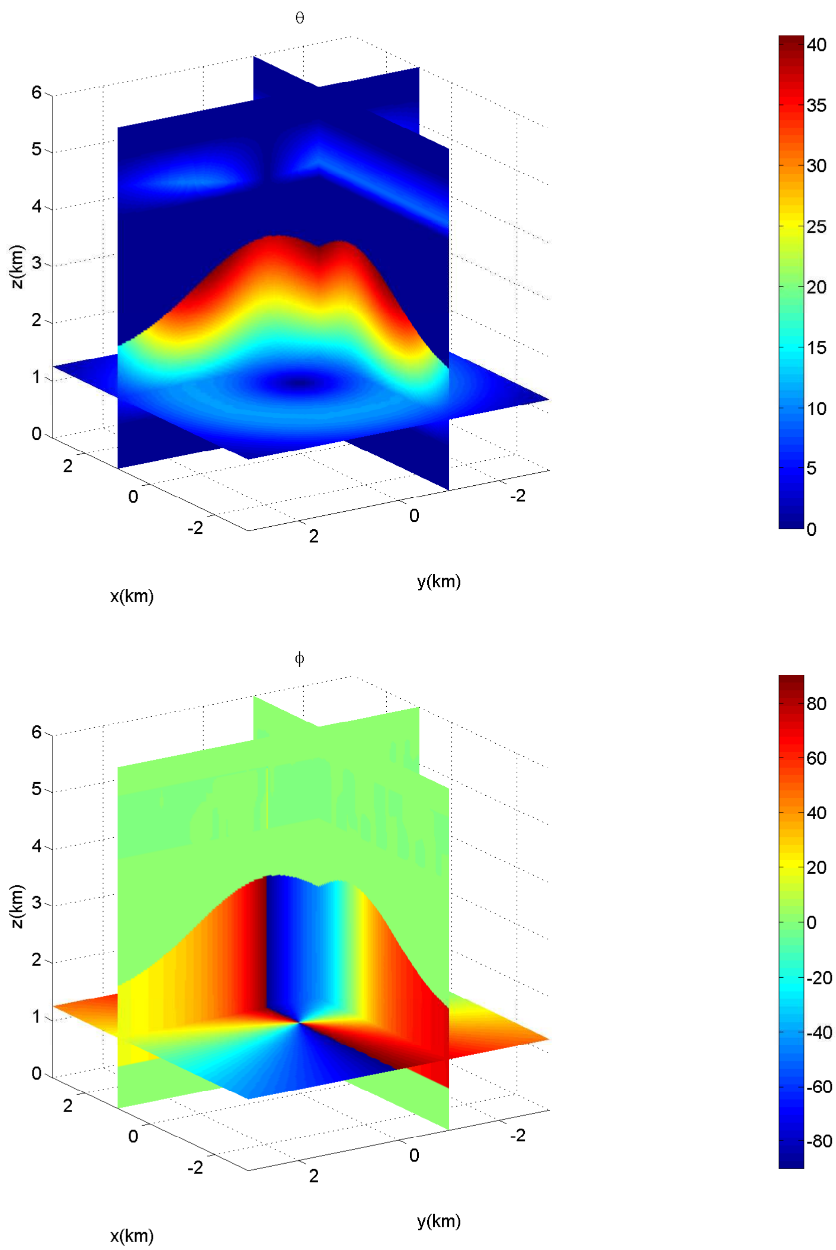


Figure 13: 3D PGS TTI model discretized on a $241 \times 241 \times 241$ mesh; **top**: dip angle θ; **bottom**: azimuth angle ϕ.

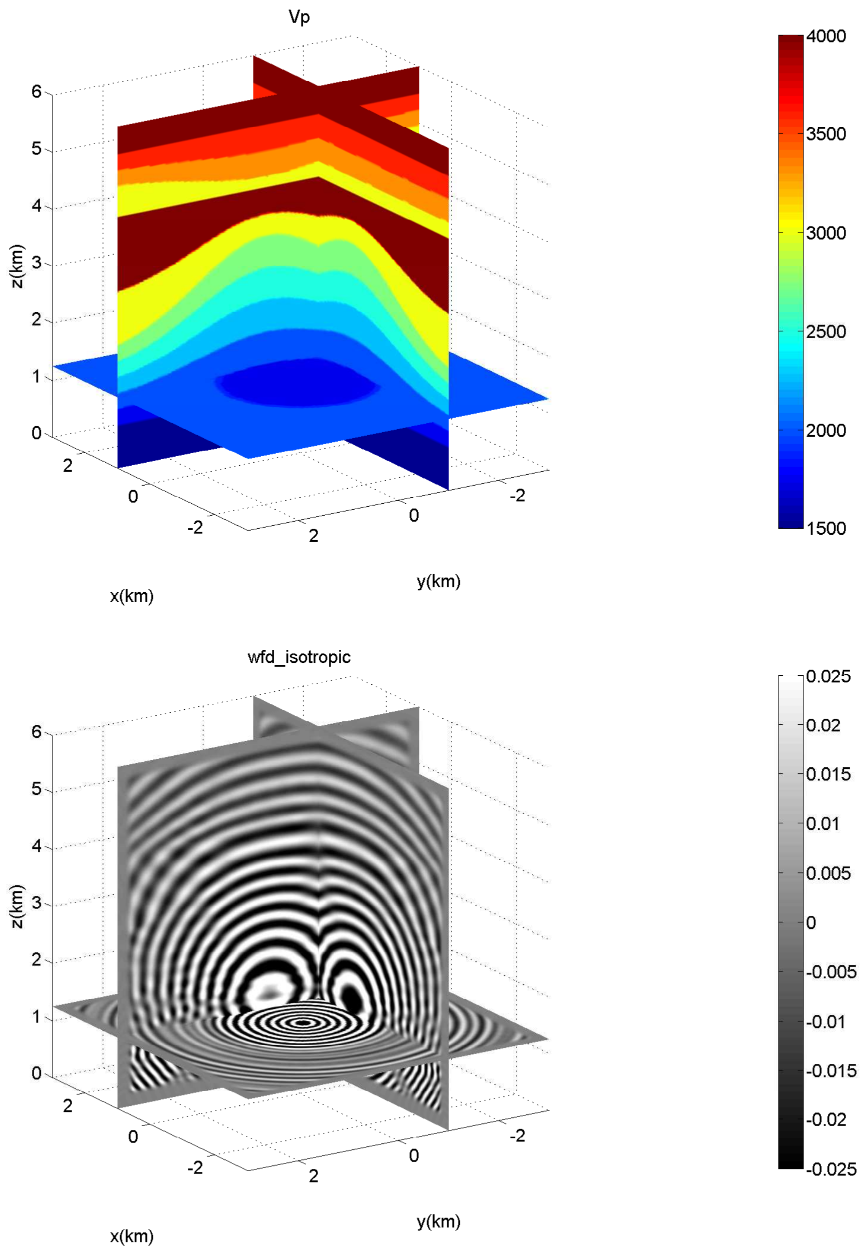


Figure 14: **top**: P wave velocity along the symmetry axis; **bottom**: 10 Hz time-harmonic wavefield for the isotropic model with the source location at $x_s = (0, 0, 1.25)$km.

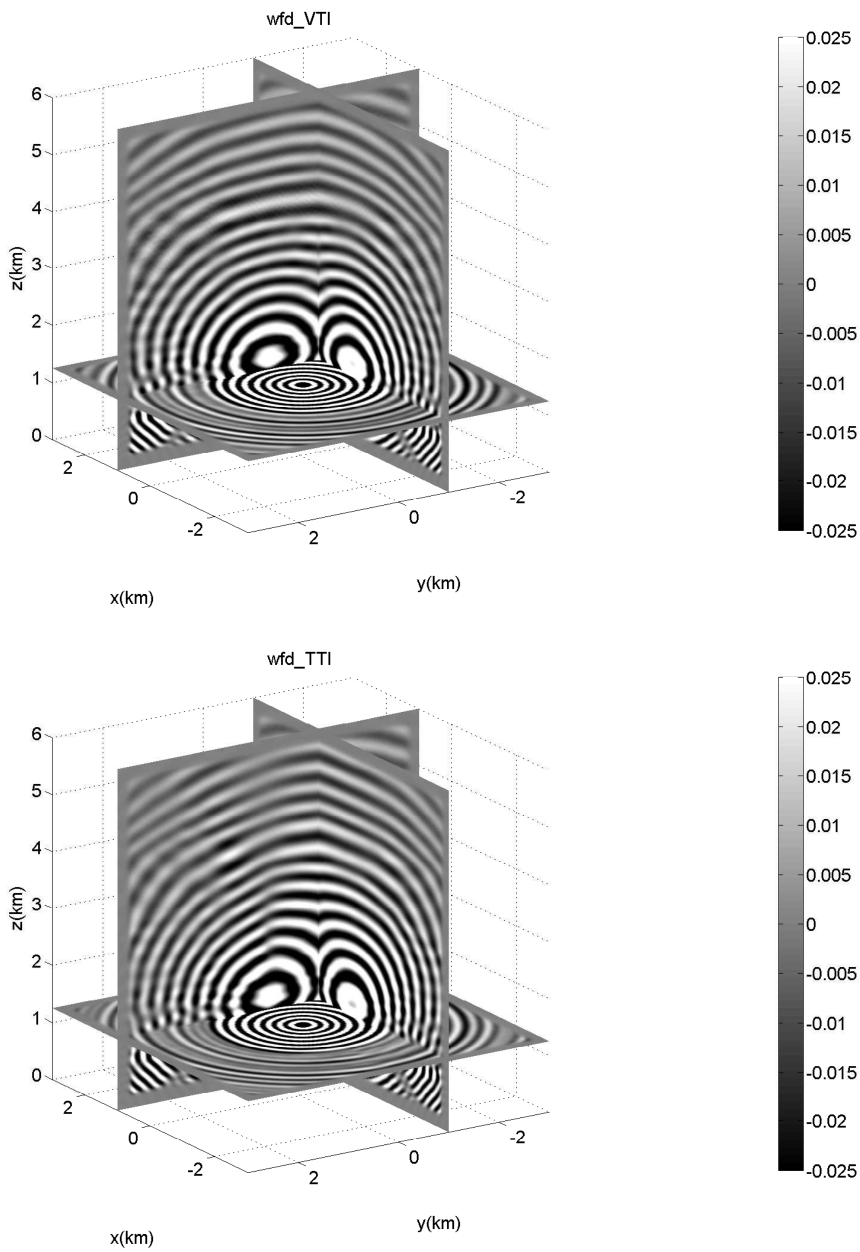


Figure 15: **top**: 10 Hz time-harmonic wavefield for the VTI model with the source location at $x_s = (0, 0, 1.25)$km; **bottom**: 10 Hz time-harmonic wavefield for the TTI model with the source location at $x_s = (0, 0, 1.25)$km.