Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN), Vol. 1 (2011) pp. 175-192.

ANISOTROPIC 'HELMHOLTZ' EQUATIONS: MASSIVELY PARALLEL STRUCTURED MULTIFRONTAL SOLVER USING NESTED DISSECTION BASED DOMAIN DECOMPOSITION WITH SEPARATORS OF VARIABLE THICKNESS

SHEN WANG*, JIANLIN XIA†, MAARTEN V. DE HOOP‡, AND XIAOYE LI§

Abstract. We consider the discretization and approximate solution of inhomogeneous anisotropic 'Helmholtz' equations in 3D. The anisotropy comprises general (tilted) TI symmetries. In particular, we are concerned with solving these equations on a large domain, for a large number of different sources. We make use of a nested dissection based domain decomposition in a massively parallel multifrontal solver combined with Hierarchically SemiSeparable (HSS) matrix compression. The anisotropy requires 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.

Key words. Helmholtz equation, anisotropy, multifrontal solver, domain descomposition, Hierarchically SemiSeparable matrices

1. Introduction. We consider the discretization and approximate solution of inhomogeneous anisotropic 'Helmholtz' equations in 3D. In particular, we are concerned with solving these equations on a large domain, for a large number of different sources in the context of modeling seismic wave propagation with applications in so-called (local optimization based) full waveform inversion (FWI) in mind. The direct method of choice for solving this problem is the multifrontal factorization algorithm [8]. The central idea of the multifrontal algorithm is to reorganize the sparse factorization into a series of dense local factorizations. The algorithm is used together with the method of nested dissection [6] to obtain a nested hierarchical structure and generate the LU factorization from the bottom up to minimize fill-ins.

In nested dissection, separators are exploited to recursively divide the mesh into subdomains of smaller sizes. Each separator consists of a small set of mesh points. The nested partitioning leads to a sequence of separators at different levels, which can form a tree. This tree is used in the multifrontal method as the merge tree. The anisotropy requires the introduction of separators with variable thickness, or separators consisting of multiple single planes. This development and the integration with the multifrontal solver, generalizing earlier work [15], is the main topic of this paper. We follow the approach developed by Xia *et al.* [18, 17] of integrating the multifrontal method with structured matrices. The main implication is that the fill-in blocks of the factorization are highly compressible using the framework of hierarchically semiseparable (HSS) matrices. The key issue, indeed, 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 introduce a general 125-point finite-difference stencil for discretizing the 'Helmholtz' operator on a regular mesh. This stencil accommodates anisotropy. Furthermore, it also allows the use of higher-order schemes in the isotropic case. We invoke PML boundaries. We note that the resulting matrix is non-Hermitian, indefinite, and relatively poorly conditioned. The algorithm presented here is designed for low to intermediate frequencies and extends an earlier algorithm with single layer separators. We exploit the regularity of the mesh leading to a complete binary assembly tree.

The introduction of anisotropic 'Helmholtz' equations is motivated by the modelling polarized elastic waves. If the elastic system of equations is of principal type, indeed, pseudodifferential equations can be constructed yielding time-harmonic polarized wave solutions; see, for example, Stolk and De Hoop [11]. The symbols of the pseudodifferential operators can be expanded to yield

*Center for Computational and Applied Mathematics, Purdue University, IN 47907 (wang273@math.purdue.edu).
†Center for Computational and Applied Mathematics, Purdue University, IN 47907 (xiaj@math.purdue.edu).
‡Center for Computational and Applied Mathematics, Purdue University, IN 47907 (mdehoop@math.purdue.edu).
§Lawrence Berkeley National Laboratory, Berkeley, CA 94720 (xsli@lbl.gov).

a (low-rank) separated representations in general. In the case of special symmetries, in coordinates aligned with the symmetry axes, one can factor out a quadratic wave number component and expand the symbol as a rational function in the other wave number components [10]. Thus an expansion in terms of partial differential operators is obtained.

The so-called acoustic equation for qP waves for VTI symmetry given in [1] is a special case of the above mentioned expansion. We use this equation as an example of our general anisotropic 'Helmholtz' equation. It is a fourth-order partial differential equation, the symbol of which can be obtained as a special limit of the dispersion relation. This particular equation is attractive, because it can be discretized with a 27-point finite-difference stencil (as in the isotropic case) admitting a separator thickness of 1. However, it is dynamically inaccurate. Also, 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 weak.

The structure of the acoustic equation for qP waves no longer yields the 27-point stencil in the case of general TTI symmetries. Time-domain strategies derived from the acoustic equation following the constuction of coupled pairs of partial differential equations which are second-order in time, for VTI, can be found in Grechka *et al.* [7] and in Zhou *et al.* [19]. The analogous construction for TTI can be found in [3, 5]. The propagator approach (in time) has been implemented by Crawley *et al.* [2]. In the frequency domain, considered here, there is no need for the standard introduction of coupled systems of partial differential equations to lower the order in time. The complication arises in the spatial part of the 'Helmholtz' operator which, in our algorithm, is addressed by introducing separators of variable thickness. We give complexity and interprocessor communication estimates. In particular, we compare these for separator thickness 2 with the ones for separator thickness 1, and verify the estimates by numerical experiments.

2. Anisotropic 'Helmholtz' equations.

2.1. Formal equations. We consider the following, anisotropic, generalization of the inhomogeneous, standard Helmholtz equation [9]:

$$\left[\Gamma(x,\partial_x,\omega)-\omega^2\right]u(x,\omega)=f(x,\omega),\quad x\in\mathbb{R}^3, \tag{2.1}$$

where

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

in which $C_{11},\ldots,C_{3312}$ are coefficients which can be dependent on x; the coefficients corresponding with (mixed) fourth-order derivatives are also dependent on ω^2. The principal symbol of this Helmholtz operator can reproduce the dispersion relaton for qP polarized elastic waves, which is discussed in the next section. (The density of mass is assumed to be constant.)

We introduce a Perfectly Matched Layer (PML) [13] contained in the computational domain, $[0, L_1] \times [0, L_2] \times [0, L_3]$ say: Let $0 < L_{11} < L_1$, then the damping function S_1 is defined as

$$S_1 = S_1(x_1, \omega) = \begin{cases} 1 & \text{if} \quad 0 \le x_1 \le L_{11}, \\ 1 - \mathrm{i}\dfrac{\sigma_0}{\omega} \cos^2\left(\dfrac{\pi}{2} \dfrac{x_1 - L_{11}}{L_1 - L_{11}}\right) & \text{if } L_{11} < x_1 \le L_1; \end{cases} \tag{2.2}$$

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. 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

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

2.2. Basic finite-difference stencil. We introduce a regular mesh and lattice,

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

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

$$\begin{aligned} &([\Gamma^\mu(., \partial_x, \omega) - \omega^2]\, u)(x_{1,i}, x_{2,j}, x_{3,k}, \omega) \\ &\quad = \sum_{m_1, m_2, m_3 = -2}^{2} \mathcal{D}^\mu_{m_1 m_2 m_3}(i, j, k)\ u(x_{1,i} + m_1 h_1, x_{2,j} + m_2 h_2, x_{3,k} + m_3 h_3, \omega) + \mathcal{O}(h^4). \end{aligned} \tag{2.3}$$

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

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

while

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

and

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

and

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

where

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

and

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

and

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

For example, in the term $\Gamma^{14}(x,\partial_x,\omega)=C_{1222}(x,\omega)\frac{\partial^4}{\partial x_1\partial x_2^3}$ the coefficient $C_{1222}(x,\omega)$ 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,

$$\mathbf{u}_{(k-1)N_1N_2+(j-1)N_1+i}(\omega)=u(x_{1,i},x_{2,j},x_{3,k},\omega), i=1,\ldots,N_1,\ j=1,\ldots,N_2,\ k=1,\ldots,N_3, \tag{2.4}$$

with $(N_1-1)\,h_1=L_1$, $(N_2-1)\,h_2=L_2$, $(N_3-1)\,h_3=L_3$, $N_1,N_2,N_3\approx N$, which applies to f in a similar fashion, and cast the discretized Helmholtz equation in corresponding matrix form:

$$\mathbf{A}(\omega)\ \mathbf{u}(\omega)=\mathbf{f}(\omega). \tag{2.5}$$

Naturally, the matrix $\mathbf{A}(\omega)$ is of size $\sim N^3\times N^3$, and shares the same nonzero pattern for different values of ω. The matrix is non-definite, non-Hermitian, and poorly conditioned. For a prescribed accuracy and given computational domain, the N grows linearly with increasing frequency. Our approach addresses the complications associated with these properties.

3. Tranverse isotropy; Multi-frequency – Propagating waves. The propagation of singularities by the solution operator of the wave equation corresponding with the anisotropic Helmholtz equation (2.1) is governed by the solutions to

$$\Gamma(x, \mathrm{i}\xi, \omega) = \omega^2; \tag{3.1}$$

if we write $\xi = \omega p$ we find the equation defining the slowness surface, $\Gamma(x, \mathrm{i}p, 1) = 1$. This equation is quartic in the components of p, and typically generates two sheets.

3.1. The polarized Helmholtz equation and approximations. We return to revisiting the original general polarized wave equations. The polarized (pseudodifferential) 'Helmholtz' equations are given by (Stolk and De Hoop [11])

$$[\omega^2 - A(x, D_x)]\, u(x, \omega) = -f(x, \omega). \tag{3.2}$$

We will restrict the discussion in this section to the principal parts, $A^{\mathrm{prin}}(x, \xi)$, of the symbols of $A(x, D_x)$. Comparing this equation with the formal equation (2.1), we identify $\Gamma(x, \mathrm{i}D_x, \omega)$ with $A(x, D_x)$.

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

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

here

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

This quantity defines $\epsilon_A = E^2/E^2_{\max}$, given by

$$\epsilon_A = \frac{(c_{11} - c_{55})(c_{33} - c_{55}) - (c_{13} + c_{55})^2}{(c_{11} - c_{55})(c_{33} - c_{55})},$$

which parametrizes the anellipticity of the medium. Both symbols attain the form of the symbol of a standard Helmholtz operator if $E^2 = 0$.

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

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

where $\theta = \theta(x)$ is the local tilt and $\phi = \phi(x)$ is the local azimuth. We then generate

$$A^{\mathrm{prin}}_{qP;\mathrm{TTI}}(x, \xi) = A^{\mathrm{prin}}_{qP}(x, R^{-1}_{\theta(x),\phi(x)}\xi).$$

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

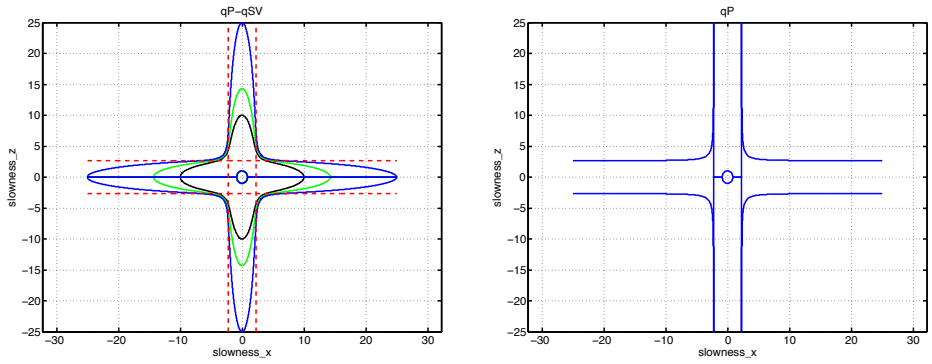

FIG. 1. *The sheets of the slowness surface associated with (3.9); left: qP-qSV for* $\epsilon = 0.2$, $\delta = 0.1$, *and three different values of* v_{sz}/v_{pz}*:* $0.1, 0.04, 0.07$*; right: the limit* $v_{sz} \downarrow 0$.

3.2. TTI 'scalar' wave equation. Let $\epsilon = \epsilon(x)$ and $\delta = \delta(x)$ denote Thomsen's parameters [12], and $v_{pz} = v_{pz}(x)$ denote the P-wave velocity along the symmetry axis. The so-called 'scalar' wave equation associated with TTI media is defined by

$$\Gamma = -v_{pz}^2 \left[\left(1 + 2(\epsilon - \delta)\frac{v_{pz}^2}{\omega^2}(H_1 + H_2) \right) H_3 + (1 + 2\epsilon)(H_1 + H_2) \right], \tag{3.3}$$

where

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

in which $\theta = \theta(x)$ is the dip angle measured away from the vertical, and $\varphi = \varphi(x)$ is the azimuth.

This scalar equation is obtained by the composition $[\omega^2 - A_{qSV}^{\rm prin}(x, -\mathrm{i}\partial_x)][\omega^2 - A_{qP}^{\rm prin}(x, -\mathrm{i}\partial_x)]$ leading to the quartic equation

$$-\frac{1}{\omega^2}\, [\omega^2 - A_{qSV}^{\rm prin}(x, -\mathrm{i}\partial_x)][\omega^2 - A_{qP}^{\rm prin}(x, -\mathrm{i}\partial_x)] u(.,\omega) = 0,$$

or (up to principal parts)

$$\begin{aligned} &\left[-\left(1 + 2(\epsilon - \delta)\frac{v_{pz}^2}{\omega^2}(H_1 + H_2) \right) H_3 - (1 + 2\epsilon)(H_1 + H_2) \right. \\ &- \frac{v_{sz}^2}{v_{pz}^2}(H_3 + H_1 + H_2) - \frac{v_{sz}^2}{\omega^2}(H_3^2 + (1 + 2\epsilon)(H_1 + H_2)^2 + (2 + 2\delta)(H_1 + H_2)H_3) \\ &\left. - \frac{\omega^2}{v_{pz}^2} \right] u(.,\omega) = 0, \end{aligned} \tag{3.4}$$

and taking the limit of the SV velocity along the symmetry axis, $v_{sz} \downarrow 0$; see also [4, 5].

The slowness surface associated with eq.(3.4) consists of a qP and a qSV sheet, and is illustrated in Figure (1) left; in Figure (1) 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, however, the qP sheet remains intact; this observation was already made by Grechka *et al.* [7].

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. To mitigate numerical instability, we smoothly adapt the coefficients near the boundary, such that the medium is elliptic in the PML.

Instead of taking the mentioned limit, one can also re-order the terms in (3.9):

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

This equation can be re-factorized, according to

$$
(3.6)\qquad -\frac{v_{sz}^2}{\omega^2}\,[\partial_{x_3}^2 + B_{qSV}^{\mathrm{prin}}(x,\omega,-\mathrm{i}\partial_{x_1},-\mathrm{i}\partial_{x_2})][\partial_{x_3}^2 + B_{qP}^{\mathrm{prin}}(x,\omega,-\mathrm{i}\partial_{x_1},-\mathrm{i}\partial_{x_2})]u(.,\omega) = 0,
$$

up to principal parts. This can be viewed as a counterpart to the original form, factorized in H_3 (the derivatives along the symmetry axis) rather than ω^2. We introduce a multiplication by v_{pz}^2 and manipulate (3.6) up to principal symbols to obtain the form

$$
\begin{aligned}
-\frac{1}{\omega^2}\,&[v_{sz}^2(x)H_3 + Z_{qSV}^{\mathrm{prin}}(x,\omega^2,-v_{sz}^2(x)(H_1+H_2))] \\
&\qquad [v_{pz}^2(x)H_3 + Z_{qP}^{\mathrm{prin}}(x,\omega^2,-v_{ph}^2(x)(H_1+H_2))]u(.,\omega) = 0,
\end{aligned}
$$

where $v_{ph} = v_{ph}(x)$ denotes the P phase velocity in the plane perpendicular to the symmetry axis, which matches equations (18)-(19) in [10]. The operators Z are pseudodifferential operators of the form

$$
Z_{qP}^{\mathrm{prin}}(.,\omega^2,-v_{ph}^2(H_1+H_2)) = \omega^2 + v_{ph}^2(H_1+H_2) + \omega^2 F(-\omega^{-2}v_{ph}^2(H_1+H_2);E^2),
$$

and similarly for Z_{qSV}^{prin}. (We note that, in dimension 3, the factorization of qSH does not hold on the symmetry axis.) Their symbols, containing square roots through F, can be expanded about the elliptic case ($E^2 = 0$).

These expansions are "rational" in the symbol of $H_1 + H_2$; see (19) and (27) in Schoenberg and De Hoop [10]. One can then write down an equation for qP waves, and, by composition, obtain higher-order partial differential equations for the successive approximations. The first-order expansion thus yields a fourth-order partial differential equation as before.

4. 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 [14, 15], but we use separators of variable thickness. Thus, the massively parallel multifrontal factorization and solution methods in [14, 15] are used here also, and we only need to focusing on the situation with variable separators.

To reduce the fill-in issue of a direct solver, the matrix $\mathbf{A}(\omega)$ in eq. (2.5) generally needs to be reordered prior to the factorization stage. The nested dissection reordering [6, 8] 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 separators consists of t striaght lines or t planes in the mesh, where t is called the thickness of the separator. In [14, 15], $t = 1$, because the FD stencil was 27-point compact. In the anisotropic case discussed in Section 2, the FD stencil is 125-point and $t = 2$ 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:

(4.1) The number of grid points in the 3D compact stencil is $(2t + 1)^3$.

The variable t will play 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 tridiagonal 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 2 top illustrates the first level nested dissection. Figures 3 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 2 middle illustrates that two y direction separators are introduced. The further reordered matrix patterns are displayed in figure 3 row three. Figure 2 bottom together with figure 3 row four display the nested dissection at level three, when four x direction separators are introduced.

After the nested dissection with a predefined total number of levels $\mathbf{l}_{\max}$, the matrix $\mathbf{A}(\omega)$ in eq. (2.5) are reordered into the pattern as illustrated by Figure 4(left). At the same time, an assembly tree which is a postordered binary tree which defines the order of the Gaussian elimination is also formed. Figure 4(right) displays this assembly tree.

The further parallel multifrontal solver together with HSS approximation introduced in [14, 15] 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 4(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}, X^i_{tail}, Z^i_{head}, Z^i_{tail}. Figure 5 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 $X^p_{head} = X^i_{tail}$, 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 plain, 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 5).

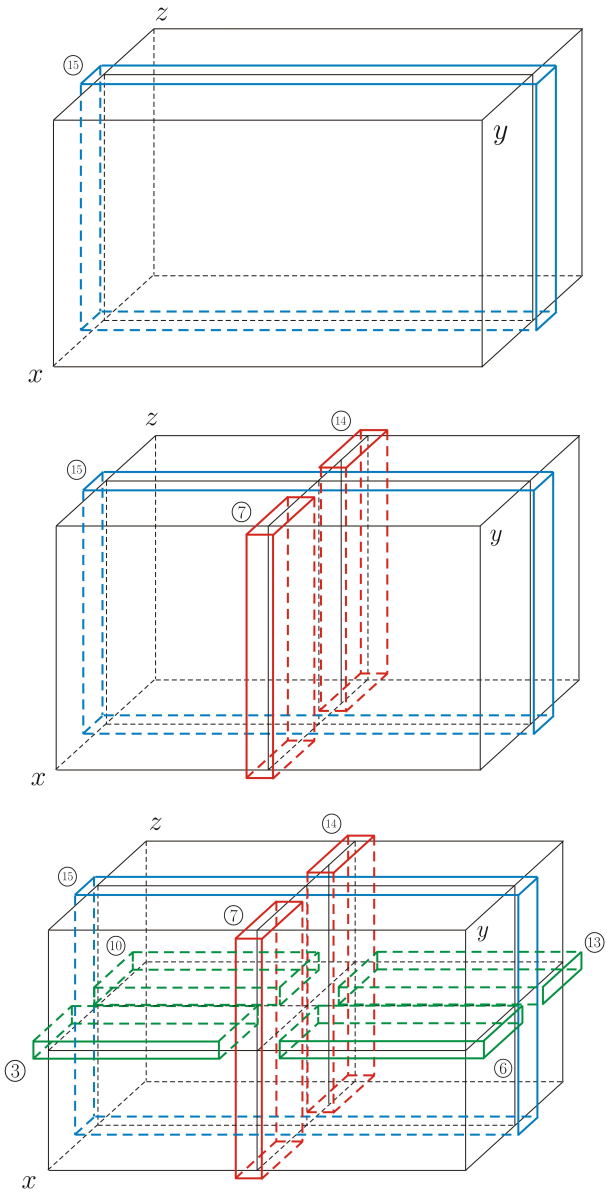

FIG. 2. *3D Nested dissection with separators of variable thickness for different levels.* ***top****: level one;* ***middle****: level two;* ***bottom****: level three.*

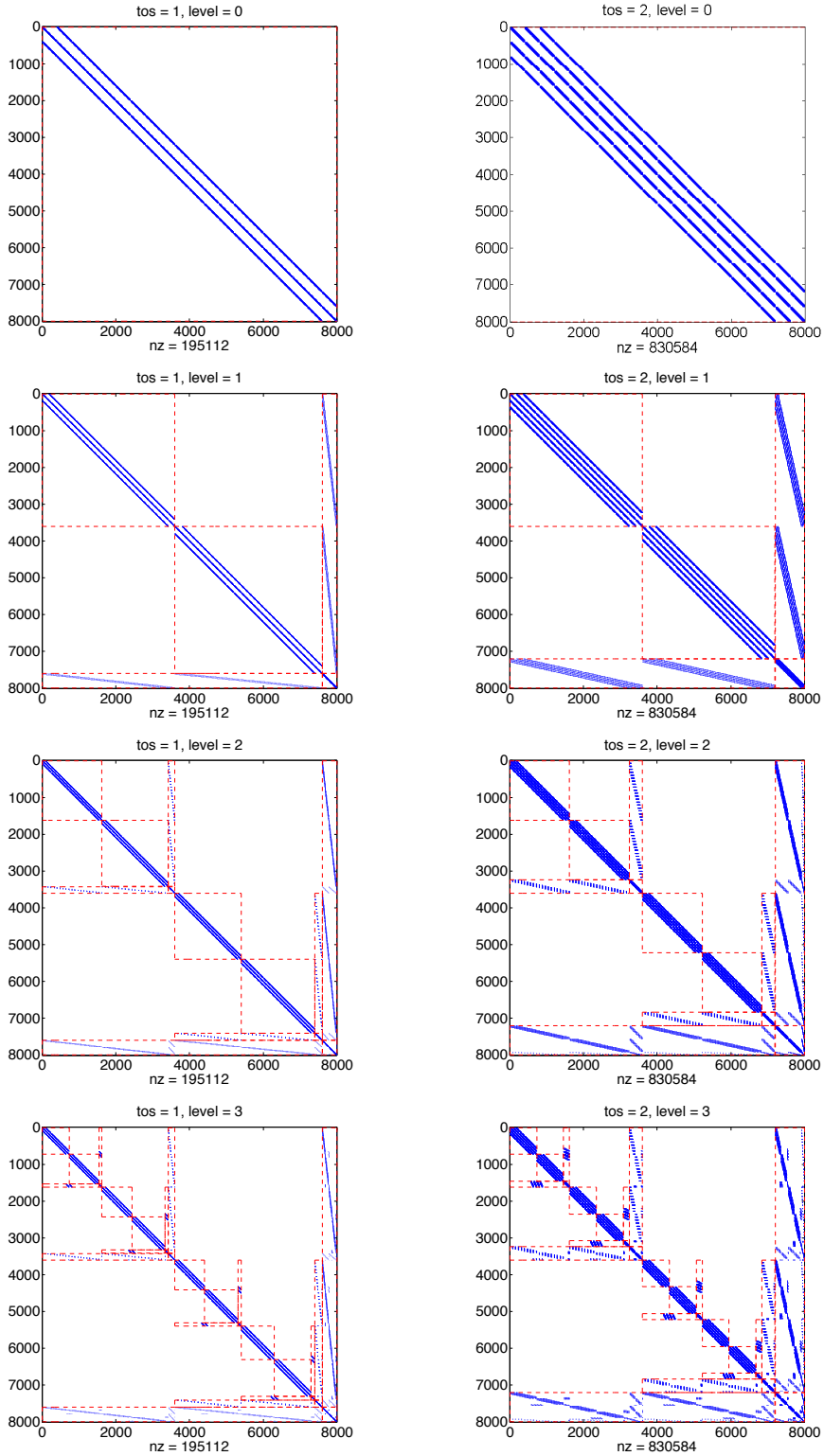

FIG. 3. *The pattern of the 3D matrix* $\mathbf{A}(\omega)$ *in eq.(2.5) discretized on a* $20 \times 20 \times 20$ *mesh, after nested dissection reordering for different levels and for variable thickness of separators (tos).* ***left column****:* $tos = 1$*;* ***right column****:* $tos = 2$*.* ***row one****:* $level = 0$*;* ***row two****:* $level = 1$*;* ***row three****:* $level = 2$*;* ***row four****:* $level = 3$*.*

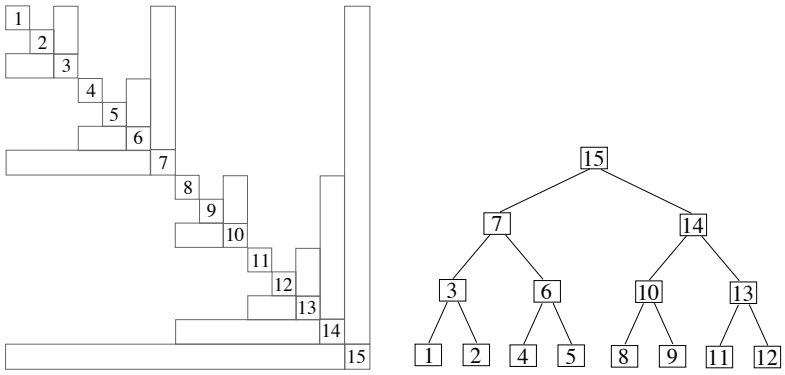

FIG. 4. ***Left**: the illustration of the matrix pattern after the nested dissection reordering displayed in figures (2) and (3); **right**: the corresponding assembly tree after the nested dissection.*

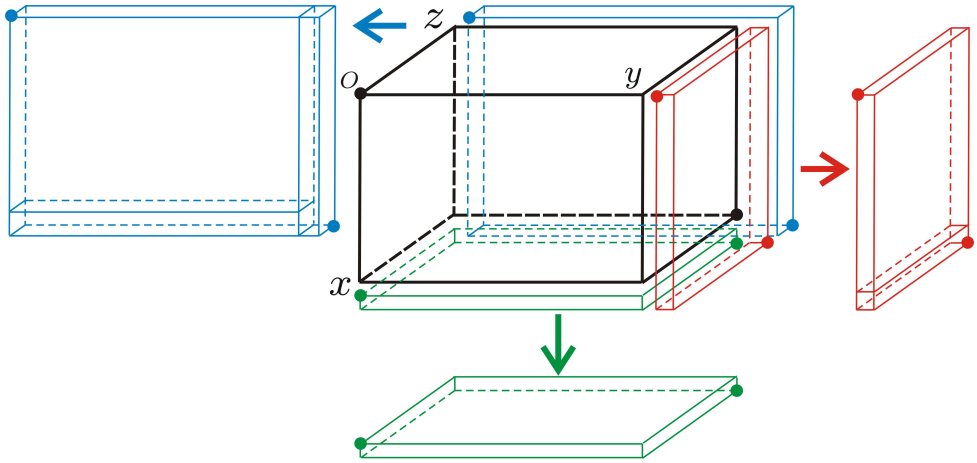

FIG. 5. *Neighbor determination in the 3D nested dissection with separators of variable thickness illustrated in figure (2).*

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.

5. 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, where $n = N^3$. The counts are improved with the incorporation of HSS structures.

The cost of the multifrontal method with HSS structures can be roughly analyzed as follows, with a way similar to the method in [17]. Assume the mesh size is $N \times N \times N$. First, we count the costs for $t = 1$, and then generalize to a variable t. The costs associated with the separators are listed in Table 5.1, where the counts in [16] are used.

Thus, we have the total cost for the algorithm

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

	Traditional factorizations	Structured factorizations
$\mathbf{l}_{\max} = O(\log_2 N)$ 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 $O(N/2^{\lfloor \mathbf{l}/2 \rfloor})$	
Cost (each separator)	$O((N/2^{\lfloor \mathbf{l}/2 \rfloor})^3)$	$O(r(N/2^{\lfloor \mathbf{l}/2 \rfloor})^2)$
Cost (subtotal)	$\sum_{\mathbf{l}=\mathbf{l}_s+1}^{\mathbf{l}_{\max}} 2^{\mathbf{l}} O((n/2^{\lfloor \mathbf{l}/2 \rfloor})^3)$	$\sum_{\mathbf{l}=0}^{\mathbf{l}_s} 2^{\mathbf{l}} O(r(n/2^{\lfloor \mathbf{l}/2 \rfloor})^2)$

TABLE 5.1

Flop count for the multifrontal method with intermediate HSS operations.

Choose $\mathbf{l}_s$ so that

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

That is,

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

In this situation, we have

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

Similarly, we can show that the solution cost is

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

and the storage requirement is

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

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

$$\mathcal{C}^t_{\text{fact}} \approx t^4 \mathcal{C}^1_{\text{fact}},\ \mathcal{S}^t_{\text{mem}} \approx t^3 \mathcal{S}^1_{\text{mem}}. \tag{5.1}$$

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

Here in our method, the error is proportional to N^{-r}, where 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 more expensive than the counts in (5.1), especially when t is large.

6. Numerical experiments.

6.1. Amplitudes and sampling rate. In the first example, we show the 3D time-harmonic wavefields in VTI and TTI homogeneous media, computed on a $100 \times 100 \times 100$ mesh with the step size $h_x = h_y = h_z = 10$m. The P-wave velocity along the symmetry axis is 4000 m/s, $\epsilon = 0.25, \delta = 0.0, \theta = 45^\circ, \omega/2\pi = 25$Hz. That is, we use a sampling rate of about 10 points per wave length. The (point) source is located at the center of the domain. Figure 6 (left) displays the wavefield computed in the VTI medium with the 27-point stencil, while Figure 6 (right) displays the wavefield computed in the TTI medium with the 125-point stencil.

We compare the relative accuracy of the TTI and VTI simulations in Figure 7, by rotating the TTI result to the orientation of the symmetry axis of the VTI medium. In view of the variable accuracy for the different derivatives in the stencil for the TTI case, to reach comparable accuracy, we increased the sampling rate from 5 points per wave length (VTI) to 10 points per wave length.

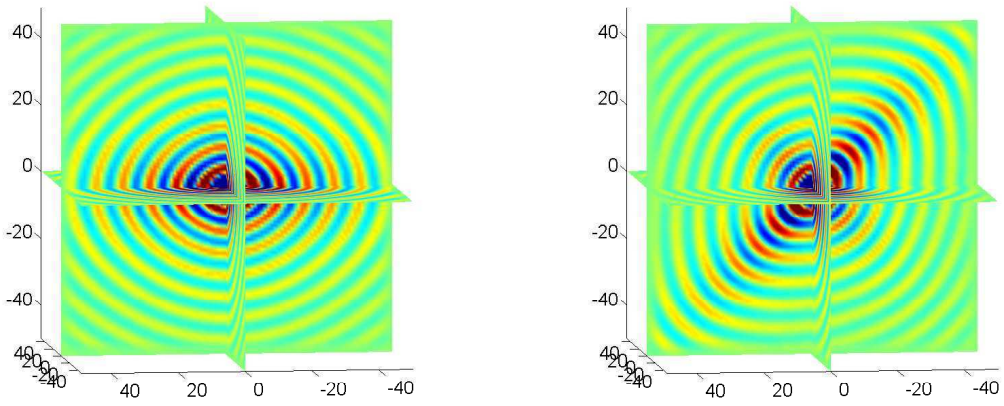

FIG. 6. ***left**: 3D VTI time-harmonic wavefield with the source at the center: $\epsilon = 0.25, \delta = 0.0, \omega/2\pi = 25Hz$; **right**: 3D TTI time-harmonic wavefield with the source at the center: $\epsilon = 0.25, \delta = 0.0, \theta = 45^\circ, \omega/2\pi = 25Hz$.*

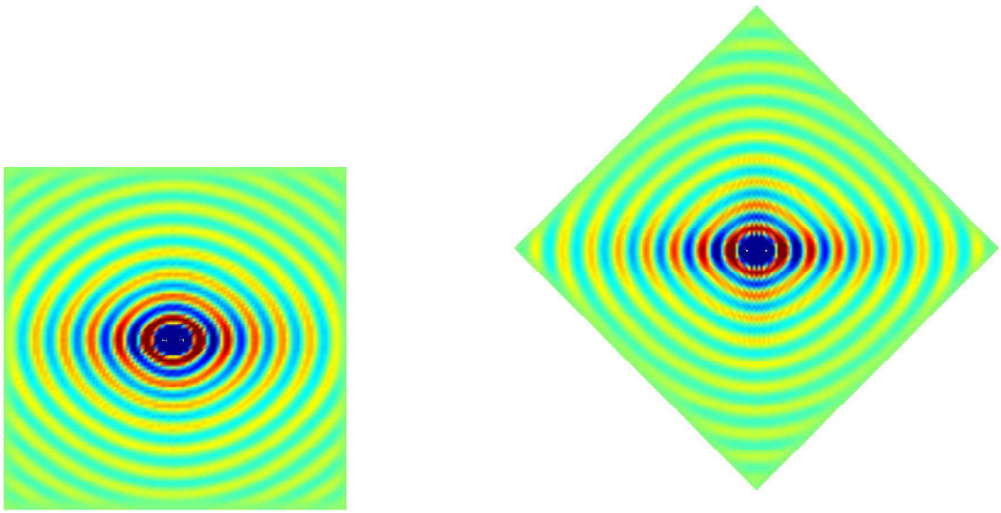

FIG. 7. ***left**: 2D VTI slice extracted from figure (6); **right**: 2D TTI slice extracted from figure (6), which is rotated to the orientation of the symmetry axis of the VTI medium.*

6.2. Computation times and interprocessor communication. We test the relatve efficiency between separator thicknesses using (part of) 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 8 (top) displays the P-wave velocity along the symmetry axis; $\epsilon - \delta$ is shown in Figure 8 (middle). Figure (8) (bottom) displays the angle of the symmetry axis measured away from the vertical direction. The step size is $h_z = h_x = 12.5$m.

As an example we take a frequency of5 Hz, and a 1801×12596 mesh. We compute the wavefields on 64 cores; the computation for the ISO case is 68 s while the computation time for the TTI case is 243 s. We confirm the estimates given in Section 4. We also show the results, on a 1801×5097 mesh, in Figure 9 (top) for the ISO case and in Figure 9 (bottom) for the TTI case. The source is located at (2.5km, 18km). We note that the PML and the strategy to suppress erroneous qSV waves work satisfactory, even in regions with large variations in $\epsilon - \delta$.

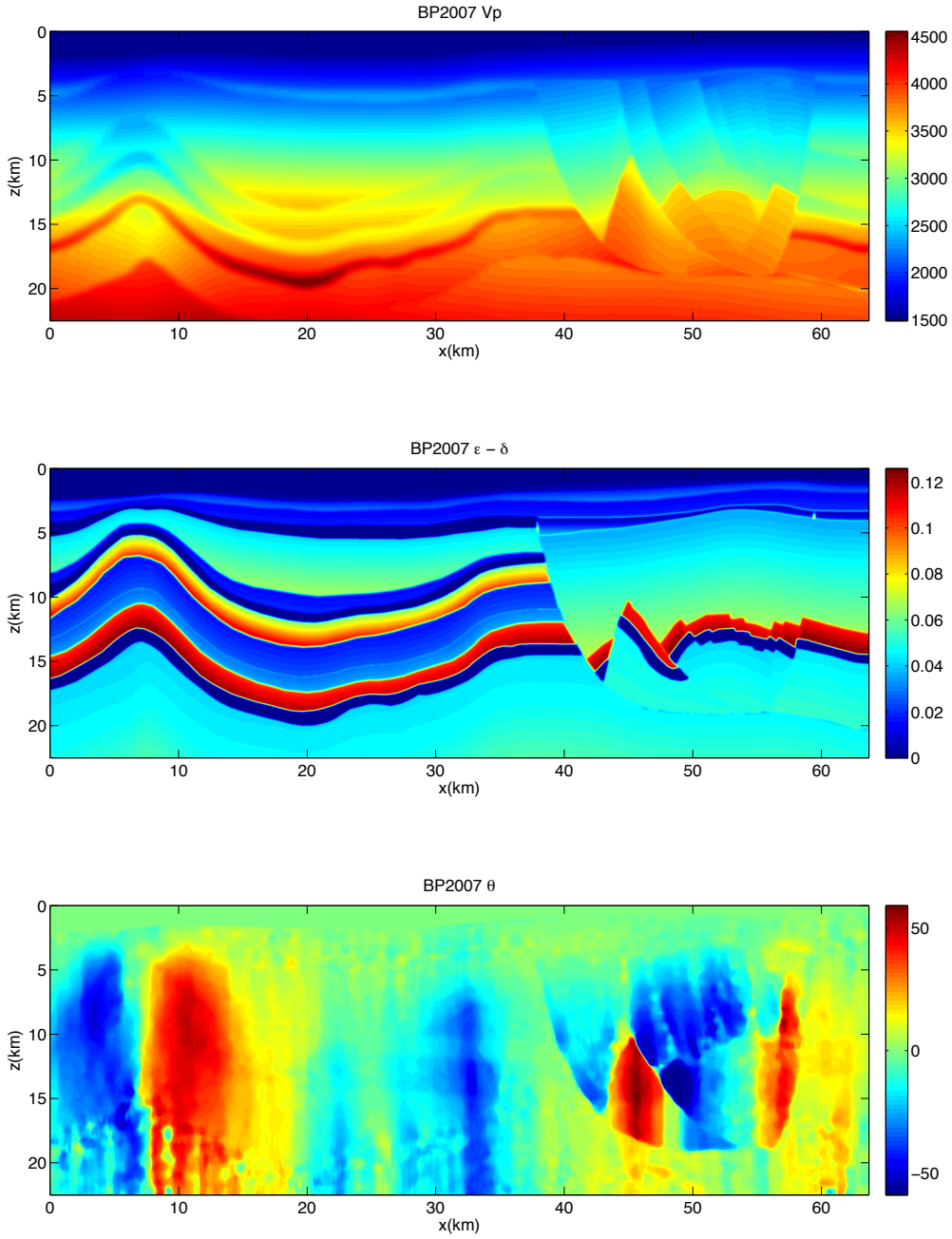

FIG. 8. *Part of the BP2007 TTI model;* ***top****: P-wave velocity along the symmetry axis;* ***middle****:* $\epsilon - \delta$*;* ***bottom****: the angle of the symmetry axis measured away from* z *direction.*

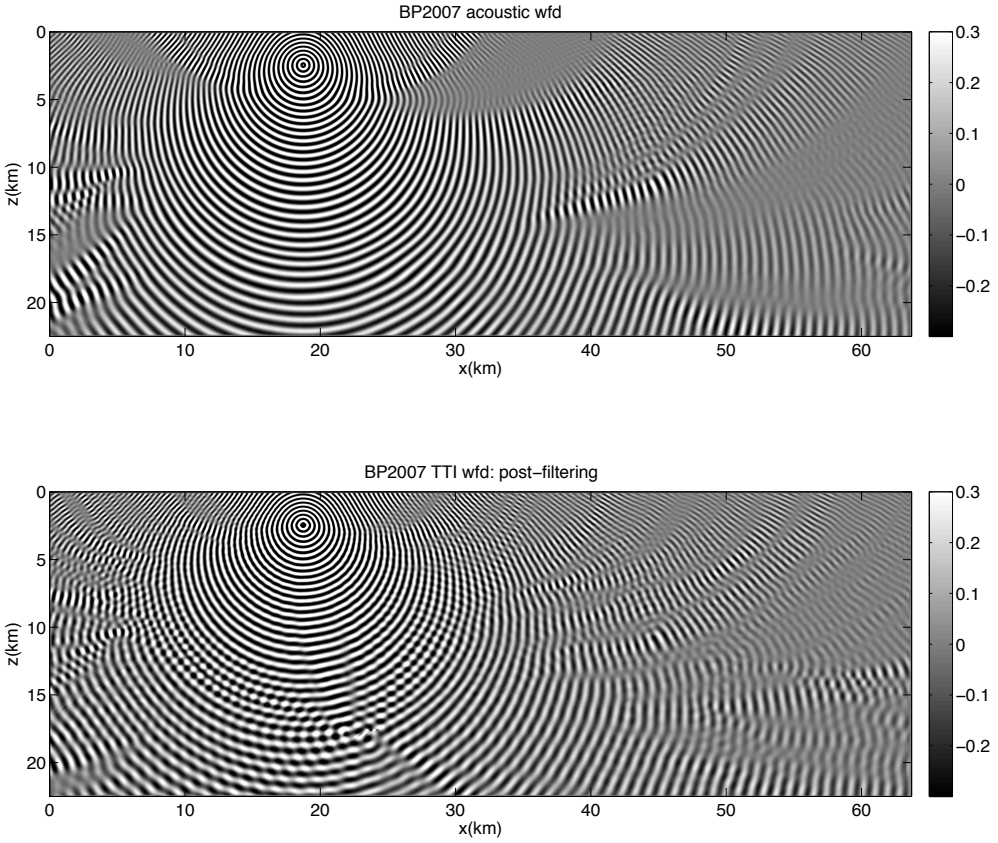

FIG. 9. *5Hz time-harmonic wavefields computed in the model shown in Figure 8 and its isotropic counterpart, with the source located at* $(2.5km, 18km)$. *Top: the wavefield in the ISO model. Bottom: the wavefield in the TTI model.*

We synthesize an artifical 3D model by using part of the BP2007 TTI model in a 2.5D fashion. We do the computations in 3D, on a $128\times128\times128$ mesh with spatial step size $h_x = h_y = h_z = 10$m. Figure 10 (upper left) displays the P-wave velocity along the symmetry axis; Figure 10 (upper right) shows $\epsilon - \delta$. Figure 10 (lower left) displays the angle of the symmetry axis measured away from the vertical direction. The frequency is 10 Hz and the source is located at the center of the domain. The wavefield is show in Figure 10 (lower right). Upon comparing computation times with a 3D model synthesized from the BP2007 ISO model, we, again, confirm the estimates given in Section 4.

7. Discussion. We developed a massively parallel approximate direct solver for anisotropic 'Helmholtz' equations. These are fourth order in space. 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 multifrontal solver equipped with a structured matrix approach, comprise the main results of this paper. 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. Our estimates indicate that, increasing the separator thickness by a factor of t is much more efficient than doubling the sampling rate, especially when t is large.

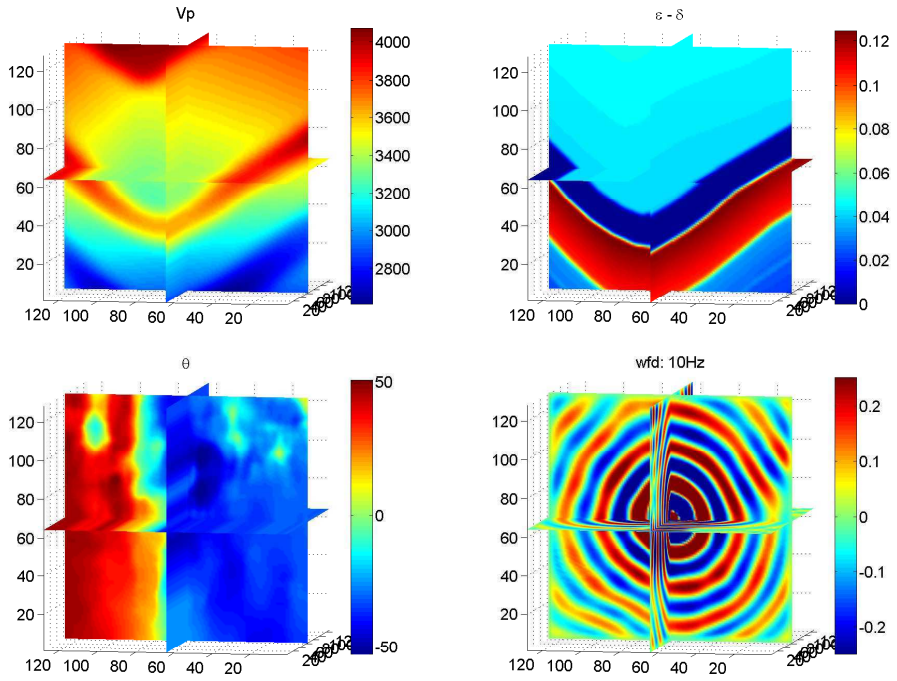

FIG. 10. *Partial 2.5D BP2007 TTI model:* ***upper left****: P-wave velocity along the symmetry axis;* ***upper right****:* $\epsilon - \delta$*;* ***lower left****: the angle of the symmetry axis measured away from the vertical direction;* ***lower right****: the 10Hz time-harmonic TTI wavefield with the source located at the center of the domain.*

Acknowledgments. This research was supported in part by the members, ConocoPhillips, ExxonMobil, PGS, Statoil and Total, of the Geo-Mathematical Imaging Group.

REFERENCES

[1] T. ALKHALIFAH, *An acoustic wave equation for anisotropic media*, Geophysics, 65 (2000), pp. 1239–1250.
[2] S. CRAWLEY, S. BRANDSBERG-DAHL, AND J. MCCLEAN, *3D TTI RTM using the pseudo-analytic method*, SEG expanded abstract, 29 (2010), pp. 3216–3220.
[3] E. DUVENECK, P. MILCIK, AND P.M BAKKER, *Acoustic VTI wave equations and their application for anisotropic reverse-time migration*, SEG expanded abstract, 27 (2008), pp. 2186–2190.
[4] R. FLETCHER, X. DU, AND P.J. FOWLER, *A new pseudo-acoustic wave equation for TI media*, SEG expanded abstract, 27 (2008), pp. 2082–2086.
[5] ———, *Stablizing acoustic reverse-time migration in TTI medium*, SEG expanded abstract, 28 (2009), pp. 2985–2989.
[6] J.A. GEORGE, *Nested dissection of a regular finite element mesh*, SIAM J. Numer. Anal., 10 (1973), pp. 345–363.
[7] V. GRECHKA, L. ZHANG, AND J.W. RECTOR, *Shear waves in acoustic anisotropic media*, Geophysics, 69 (2004), pp. 576–582.
[8] J.W.H. LIU, *The multifrontal method for sparse matrix solution: Theory and practice*, SIAM Review, 34 (1992), pp. 82–109.
[9] S. OPERTO, J. VIRIEUX, A. RIBODETTI, AND J.E. ANDERSON, *Finite-difference frequency-domain modeling of viscoacoustic wave propagation in 2D Tilted Transversely Isotropic (TTI) media*, Geophysics, 74 (2009), pp. T75–T95.
[10] M. SCHOENBERG AND M.V. DE HOOP, *Approximate dispersion relations for* $qP - qSV$ *waves in transversely isotropic media*, Geophysics, 65 (2000), pp. 919–933.
[11] C.C. STOLK AND M.V. DE HOOP, *Microlocal analysis of seismic inverse scattering in anisotropic elastic media*, Communications on Pure and Applied Mathematics, 55 (2002), pp. 261–301.
[12] L. THOMSEN, *Weak elastic anisotropy*, Geophysics, 51 (1986), pp. 1954–1966.
[13] E. TURKEL AND A. YEFET, *Absorbing PML boundary layers for wave-like equations*, Applied Numerical Mathematics, 27 (1998), pp. 533–557.
[14] S. WANG, M.V. DE HOOP, AND J.L XIA, *Seismic inverse scattering via Helmholtz operator factorization and*

optimization, Journal of Computational Physics, 229 (2010), pp. 8445–8462.
[15] S. Wang, M.V. de Hoop, and J. Xia, *On 3D modeling of seismic wave propagation via a structured massively parallel multifrontal direct helmholtz solver*, to appear on Geophysical Prospecting, (2011).
[16] J. Xia, *On the complexity of some hierarchical structured matrices*, preprint, (2010).
[17] J. Xia, S. Chandrasekaran, M. Gu, and X.S. Li, *Superfast multifrontal method for large structured linear systems of equations*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1382–1411.
[18] ———, *Fast algorithms for hierarchically semiseparable matrices*, Numer. Linear Algebra Appl., 2010 (2010), pp. 953–976.
[19] H. Zhou, G. Zhang, and R. Bloor, *An anisotropic acoustic wave equation for modeling and migration in 2d TTI media*, SEG expanded abstract, 25 (2006), pp. 194–198.