# Element-by-element parallel spectral-element methods for 3-D teleseismic wave modeling 

Shaolin Liu ${ }^{1}$, Dinghui Yang ${ }^{1}$, Xingpeng Dong ${ }^{1}$, Qiancheng Liu ${ }^{2}$, and Yongchang Zheng ${ }^{3}$<br>${ }^{1}$ Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, China<br>${ }^{2}$ King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia<br>${ }^{3}$ Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100730, China<br>Correspondence to: Dinghui Yang (dhyang@math.tsinghua.edu.cn)

Received: 14 April 2017 - Discussion started: 8 May 2017
Revised: 17 August 2017 - Accepted: 30 August 2017 - Published: 28 September 2017


#### Abstract

The development of an efficient algorithm for teleseismic wave field modeling is valuable for calculating the gradients of the misfit function (termed "misfit gradients") or Fréchet derivatives when the teleseismic waveform is used for adjoint tomography. Here, we introduce an element-byelement parallel spectral-element method (EBE-SEM) for the efficient modeling of teleseismic wave field propagation in a reduced geology model. Under the plane-wave assumption, the frequency-wavenumber (FK) technique is implemented to compute the boundary wave field used to construct the boundary condition of the teleseismic wave incidence. To reduce the memory required for the storage of the boundary wave field for the incidence boundary condition, a strategy is introduced to efficiently store the boundary wave field on the model boundary. The perfectly matched layers absorbing boundary condition (PML ABC) is formulated using the EBE-SEM to absorb the scattered wave field from the model interior. The misfit gradient can easily be constructed in each time step during the calculation of the adjoint wave field. Three synthetic examples demonstrate the validity of the EBE-SEM for use in teleseismic wave field modeling and the misfit gradient calculation.


## 1 Introduction

The increasing demand for high-resolution imaging of deep lithospheric structures requires the utilization of teleseismic datasets for waveform inversion. Teleseismic waves provide tremendous amounts of information for the detection of
crustal and upper mantle structures (Rondenay, 2009). Over the past forty years, many techniques have been developed to analyze teleseismic wave datasets, including receiver function analysis (Langston, 1977; Kind et al., 2012), teleseismic wave travel-time tomography based on ray theory (Zhang et al., 2011), teleseismic migration (Shragge et al., 2006), and teleseismic scattering tomography (Roecker et al., 2010; Tong et al., 2014a). To achieve the high-resolution imaging of lithospheric structures, the adjoint-state method has become the state-of-the-art technique for teleseismic wave imaging (Tong et al., 2014a; Monteiller et al., 2015).

The adjoint tomography technique constructs the Fréchet derivatives of the objective function with respect to the model parameters by numerically solving the full seismic wave equation twice if the forward wave fields are stored on a computer disk at every given time interval (Tromp et al., 2005; Liu and Tromp, 2006; Fichtner, 2011). Adjoint tomography has been successfully implemented to investigate crustal (Tape et al., 2009) and continental subsurface heterogeneity (Chen et al., 2015). Seismic adjoint waveform tomography, which has a greater resolution than the ray-based travel-time tomography for the same dense seismic ray coverage (Liu and Gu, 2012), is able to image small-scale (half of the minimum wavelength) heterogeneity (Virieux and Operto, 2009). The main drawback of the adjoint tomography is its large computational burden. The computational requirement is linearly related to the number of earthquakes used for the tomographic inversion and the iterations required by the optimization technique. For a typical local-scale model, such as the southern California region, several thousand 3-D
full-wave field simulations are required to perform an adjoint tomography inversion (Tape et al., 2007, 2009).

Because most earthquakes occur in the crust, and seismic rays cannot easily illuminate the deep lithosphere in local seismic tomography, imaging the deep lithospheric structures can be difficult (Tong et al., 2014b, c). Increasing the model size enables more deep reflections and refractions to be included in the inversion dataset; as a result, deep structures can be inverted by fitting these reflected and refracted waveforms (Chen et al., 2015). However, for continental-scale models, it is difficult to invert short-period seismic data on a standard computing cluster, such as $1-2 \mathrm{~s}$ for P waves and $3-6 \mathrm{~s}$ for S waves (Tong et al., 2015).

To reduce the amount of computation involved in solving the full seismic wave equation, many hybrid methods have been developed to localize 2-D/3-D numerical solvers. The boundary conditions for the reduced simulation model are provided by rapid 1-D analytical solutions for the 1-D background Earth model (Capdeville et al., 2003; Monteiller et al., 2013, 2015; Tong et al., 2014a, 2015). The 2-D/3-D responses to the heterogeneity inside the reduced model contribute to the coda waves of the teleseismic phases, and the 2-D/3-D effects outside the model are neglected. This assumption of a 1-D background layered Earth model is similar to that in receiver function analysis and is often effective for a station that is sufficiently far from the source (Langston, 1977; Rondenay, 2009).

Although the computational requirements can be efficiently reduced by these hybrid methods, the computational costs are still excessive for a small workstation when we are faced with the several thousand forward simulations required in 3-D teleseismic adjoint tomography. To further reduce the computational costs, Roecker et al. (2010) constructed a frequency domain $2.5-\mathrm{D}$ hybrid method for teleseismic wave simulations. To simplify the teleseismic wave field invariant along a particular axis, the $2.5-\mathrm{D}$ formulation can significantly reduce the computational demands. However, the $2.5-\mathrm{D}$ formulation may restrict the application of the method in arbitrarily heterogeneous media (Tong et al., 2015).

In addition to the large computational demand (CPU time) associated with the 3-D hybrid methods, satisfying the memory requirements for storing the boundary wave fields to construct teleseismic incident boundary conditions is another important issue that should be carefully considered. Tong et al. (2015) adapted the Clayton and Engquist-type (CE-type) boundary condition (Clayton and Engquist, 1977) to interface the 1-D background analytical solution with a numerical solver on the boundary of a reduced model. This treatment can both assign the teleseismic incident condition for the computational domain and absorb the scattered wave field from the interior of the heterogeneous model. Implementing the CE-type boundary condition is extremely simple and does not substantially increase the required CPU time. However, the CE-type boundary condition can efficiently absorb waves only at approximately normal incident angles, and in-
cident waves at grazing angles may be reflected back to the model (Yang et al., 2003), which may reduce the accuracy of the forward and adjoint wave fields in teleseismic adjoint tomography and thus decrease the accuracy of the constructed Fréchet derivatives. Note that the CE-type boundary condition requires nine wave field components (three displacement components and six stress components) to be stored on the model boundary; this requirement may be a significant burden on the computer memory for a relatively large-scale model decomposed by a dense numerical mesh.

Here, we introduce the element-by-element parallel spectral-element method (EBE-SEM) for the efficient modeling of teleseismic plane-wave propagation in reduced models. A significant advantage of the EBE-SEM is the easy parallelization of the spectral-element algorithm, which does not require the assembly of the global stiffness matrix. The spectral elements are equally allocated to every CPU core, and the product of the stiffness matrix and the solution vector is calculated element by element; these aspects are quite useful for ensuring load balance among the CPU cores. The element stiffness matrix can be written as the tensor product of the submatrices, which greatly reduce the computer memory. In addition, the misfit gradient can be efficiently constructed because the element matrices for calculating the misfit gradient can also be formulated from the tensor products of the element submatrices. The perfectly matched layers (PMLs; Collino and Tsogka, 2001; Komatitsch and Tromp, 2003; Liu et al., 2014) are formulated by the EBE-SEM to effectively absorb scattered waves. A detailed discussion is presented to incorporate the teleseismic incident boundary condition in the EBE-SEM, and only six components on the interface between the computational domain and the PML domain must be stored in the computer memory. The high efficiency of the EBE-SEM for teleseismic wave modeling and misfit gradient construction is demonstrated by using three numerical examples.

## 2 EBE-SEM

A schematic of a teleseismic plane wave entering a localized model is depicted in Fig. 1, where the localized model is delineated by the blue lines. We first introduce the EBESEM for isotropic elastic wave propagation in an infinite half space, which includes the localized model. We denote the total wave field $\boldsymbol{u}_{\mathrm{t}}$, which is the summation of the background wave field $\boldsymbol{u}_{\mathrm{FK}}$ in the layered media and the scattered wave field $\boldsymbol{u}_{\mathrm{s}}$. The wavenumber-frequency (FK) domain method (Haskell, 1953; Zhu and Rivera, 2002; Tong et al., 2014a) can be used to efficiently calculate $\boldsymbol{u}_{\mathrm{FK}}$, and $\boldsymbol{u}_{\mathrm{s}}$ is excited by the incident plane wave because of the heterogeneity of the media.


Figure 1. Schematic of a teleseismic plane wave entering a local two-layered crust-upper-mantle model. The study area is delineated by the blue lines. The origin of the Cartesian system is located on the surface at corner A of the model. The positive $x$ and $y$ directions point to the east and south, respectively, and the positive $z$ direction points downward. The inverted red triangle located at ( 60 km , $60 \mathrm{~km}, 0 \mathrm{~km}$ ) represents a seismic station. The plane in the lower left of the figure denotes the wave front of the incident teleseismic plane wave. The purple arrows indicates the normal vector of the plane wave. The incident angle is $\theta$, which is the angle between the normal vector and $-z$. The azimuth angle is $\varphi$, which is the angle between the projection of the normal vector into the $x-y$ plane and $-y$. The azimuth angle is not explicitly shown in this figure.

### 2.1 EBE-SEM for total wave simulation

We assume that the total wave field obeys the following elastic wave equation (Tong et al., 2014a):
$\rho \ddot{\boldsymbol{u}}_{\mathrm{t}}=\nabla \cdot\left[\lambda\left(\nabla \cdot \boldsymbol{u}_{\mathrm{t}}\right) \mathrm{I}+\mu\left(\nabla \boldsymbol{u}_{\mathrm{t}}+\nabla \boldsymbol{u}_{\mathrm{t}}^{T}\right)\right]$,
where the two dots above $\boldsymbol{u}_{\mathrm{t}}$ denote the second-order time derivative, $\mathbf{I}$ is the $3 \times 3$ identity matrix, $\lambda$ and $\mu$ are Lamé parameters, $\nabla=\left(\boldsymbol{e}_{x} \partial_{x}+\boldsymbol{e}_{y} \partial_{y}+\boldsymbol{e}_{z} \partial_{z}\right)$ is the gradient operator, and the superscript $T$ denotes the matrix transpose operation. Following the classical SEM (Seriani, 1997; Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 2002), using an arbitrary test vector $v$ to multiply both sides of Eq. (1) and integrating over the domain $\Omega$, we obtain the following weak
form wave equation:

$$
\begin{align*}
\int_{\Omega} \boldsymbol{v} \cdot \ddot{\boldsymbol{u}}_{\mathrm{t}} \mathrm{~d} \Omega & +\int_{\Omega} \nabla \boldsymbol{v}:\left(\lambda\left(\nabla \cdot \boldsymbol{u}_{\mathrm{t}}\right) \mathbf{I}+\mu\left(\nabla \boldsymbol{u}_{\mathrm{t}}+\nabla \boldsymbol{u}_{\mathrm{t}}^{T}\right)\right) \mathrm{d} \Omega \\
& =\int_{\partial \Omega} \boldsymbol{v} \cdot \sigma_{\mathrm{t}} \cdot \boldsymbol{n} \mathrm{~d} s \tag{2}
\end{align*}
$$

where $\boldsymbol{n}$ is the normal vector of the boundary $\partial \Omega$. The definition of the double dot product can be found in De Basabe and Sen (2007). Under the natural boundary condition, the right side of Eq. (2) is 0 . If the infinite space is decomposed into $N$ nonoverlapping hexahedral elements, then Eq. (2) can be written as

$$
\begin{align*}
\sum_{e=1}^{N} \int_{\Omega_{e}} \boldsymbol{v} \cdot \ddot{\boldsymbol{u}}_{\mathrm{t}} \mathrm{~d} \Omega & +\sum_{e=1}^{N} \int_{\Omega_{e}} \nabla \boldsymbol{v}:\left(\lambda\left(\nabla \cdot \boldsymbol{u}_{\mathrm{t}}\right) \mathbf{I}\right. \\
& \left.+\mu\left(\nabla \boldsymbol{u}_{\mathrm{t}}+\nabla \boldsymbol{u}_{\mathrm{t}}^{T}\right)\right) \mathrm{d} \Omega=0 \tag{3}
\end{align*}
$$

Each hexahedral element is mapped to the reference cube $[-1,1]^{3}$, and the chosen interpolation points are consistent with the Gauss-Legendre-Lobatto (GLL) quadrature points (Cohen, 2002). The interpolation basis functions in each element are as follows:
$\phi_{i}(\xi, \gamma, \eta)=\varphi_{i_{1}}(\xi) \varphi_{i_{2}}(\gamma) \varphi_{i_{3}}(\eta)$,
where the subscript of $\phi$ denotes the $i$ th basis function, $\xi$, $\gamma$, and $\eta$ are the three coordinates in the isoparametric coordinate system, and $\varphi$ is the Lagrange polynomial. The $i$ th interpolation point in the physical element is mapped to the $\left(i_{1}, i_{2}, i_{3}\right)$ th node in the cube $[-1,1]^{3}$. Based on Eq. (4) and the discrete values $\boldsymbol{u}_{\mathrm{t}}^{e, i}$ and $\boldsymbol{v}^{e, i}$ on the interpolation points, the continuous values $\boldsymbol{u}_{\mathrm{t}}^{e}$ and $\boldsymbol{v}^{e}$ in element $e$ can be approximated by

$$
\left\{\begin{array}{l}
\boldsymbol{u}_{\mathrm{t}}^{e} \approx \sum_{i=1}^{(n+1)^{3}} \boldsymbol{u}_{\mathrm{t}}^{e, i} \phi_{i}(x, y, z)  \tag{5}\\
\boldsymbol{v}^{e} \approx \sum_{i=1}^{(n+1)^{3}} \boldsymbol{v}^{e, i} \phi_{i}(x, y, z)
\end{array}\right.
$$

where $n$ is the polynomial order of the interpolation basis. By substituting Eq. (5) into Eq. (2) and using the GLL quadrature rule, we obtain the ordinary differential equation in terms of time:
$\sum_{e=1}^{N} \mathbf{M}^{e} \ddot{\boldsymbol{U}}_{\mathrm{t}}^{e}+\sum_{e=1}^{N} \mathbf{K}^{e} \boldsymbol{U}_{\mathrm{t}}^{e}=0$,
where $\mathbf{M}^{e}$ is the element mass matrix, $\mathbf{K}^{e}$ is the element stiffness matrix, and $\boldsymbol{U}_{\mathrm{t}}^{e}$ is the element solution vector. The elements of $\mathbf{M}^{e}$ and $\mathbf{K}^{e}$ are presented in Appendix A. Because the interpolation points are consistent with the GLL quadrature points, $\mathbf{M}^{e}$ is diagonal, which is one of the advantages
of the Legendre polynomial-based SEM because the explicit inverse of the mass matrix can be easily obtained. The assembled global diagonal mass is as follows:
$\mathbf{M}=\sum_{e=1}^{N} \mathbf{M}^{e}$.
We define the projection operator $T^{e}$ to map the corresponding elements of $\mathbf{M}$ back to form an element matrix:
$\tilde{\mathbf{M}}^{e}=T^{e}(\mathbf{M})$,
where the tilde above $\tilde{\mathbf{M}}^{e}$ is used to distinguish the difference between the back-projected mass matrix and the element mass matrix in Eq. (6). Because the elements of the element mass matrix that correspond to the common points shared by the adjacent elements are summed to form the global mass matrix in Eq. (7), some elements of $\widetilde{\mathbf{M}}^{e}$ are greater than those of $\mathbf{M}^{e}$. The notation of Eq. (8) is important for the discussion below. If we denote the global solution vector as $\boldsymbol{U}_{\mathrm{t}}$, then the back-projected element solution $T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)$ is still $\boldsymbol{U}_{\mathrm{t}}^{e}$ because $\boldsymbol{u}_{\mathrm{t}}$ is continuous cross elements. From Eq. (6), we have
$\ddot{\boldsymbol{U}}_{\mathrm{t}}=-\mathbf{M}^{-1} \sum_{e=1}^{N} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)$.
An explicit temporal scheme, such as the Newmark scheme (Liu et al., 2017a), and structure-preserving schemes (Liu et al., 2017b) can be adopted for time integration. Here, we simply use the second-order central difference method (Dablain, 1986) for the time evolution. The product of the stiffness matrix and the solution vector is calculated element by element at the element level rather than at the global level. The computational burden of SEM mainly stems from the matrix-vector product, and the significant advantage of the element-by-element scheme of Eq. (9) is the great computational balance among CPU cores when the parallel algorithm is performed on a workstation. Our parallel code is based on the Message Passing Interface (MPI) library (www.mpich.org/downloads), and we equally distribute the spectral elements to CPU cores.

To further increase the computational efficiency of SEM, the GLL quadrature rule is fully considered in the product of the stiffness matrix and the solution vector. To simplify the discussion, the product of $\mathbf{K}_{11}^{e}$ in Eq. (A4) and $\boldsymbol{U}_{\mathrm{t} x}$ is presented below, where the subscript $x$ denotes the $x$ component of the discrete displacement vector. Based on the tensor product of matrices (Seriani, 1998), the matrix-vector product can be written as

$$
\begin{align*}
\mathbf{K}_{11}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right)= & {\left[\int_{\Omega_{e}}(\lambda+2 \mu) \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial x}\right.} \\
& \left.+\mu\left(\frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y}+\frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial z}\right) \mathrm{d} \Omega\right] T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right) \\
& =(\lambda+2 \mu)\left[\int_{\Omega_{e}}\left(\frac{\partial \xi}{\partial x} \frac{\partial \phi_{i}}{\partial \xi}+\frac{\partial \gamma}{\partial x} \frac{\partial \phi_{i}}{\partial \gamma}+\frac{\partial \eta}{\partial x} \frac{\partial \phi_{i}}{\partial \eta}\right)\right. \\
& \left.\left(\frac{\partial \xi}{\partial x} \frac{\partial \phi_{j}}{\partial \xi}+\frac{\partial \gamma}{\partial x} \frac{\partial \phi_{j}}{\partial \gamma}+\frac{\partial \eta}{\partial x} \frac{\partial \phi_{j}}{\partial \eta}\right) \mathrm{d} \Omega\right] T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right) \\
& +\mu\left[\int_{\Omega_{e}}\left(\frac{\partial \xi}{\partial y} \frac{\partial \phi_{i}}{\partial \xi}+\frac{\partial \gamma}{\partial y} \frac{\partial \phi_{i}}{\partial \gamma}+\frac{\partial \eta}{\partial y} \frac{\partial \phi_{i}}{\partial \eta}\right)\right. \\
& \left.\left(\frac{\partial \xi}{\partial y} \frac{\partial \phi_{j}}{\partial \xi}+\frac{\partial \gamma}{\partial y} \frac{\partial \phi_{j}}{\partial \gamma}+\frac{\partial \eta}{\partial y} \frac{\partial \phi_{j}}{\partial \eta}\right) \mathrm{d} \Omega\right] T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right) \\
& +\mu\left[\int_{\Omega_{e}} \frac{\partial \xi}{\partial z} \frac{\partial \phi_{i}}{\partial \xi}+\frac{\partial \gamma}{\partial z} \frac{\partial \phi_{i}}{\partial \gamma}+\frac{\partial \eta}{\partial z} \frac{\partial \phi_{i}}{\partial \eta}\right) \\
& \left.\left(\frac{\partial \xi}{\partial z} \frac{\partial \phi_{j}}{\partial \xi}+\frac{\partial \gamma}{\partial z} \frac{\partial \phi_{j}}{\partial \gamma}+\frac{\partial \eta}{\partial z} \frac{\partial \phi_{j}}{\partial \eta}\right) \mathrm{d} \Omega\right] T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right) \\
= & (\lambda+2 \mu)\left[\sum_{i=1}^{3} \sum_{j=1}^{3} \mathbf{P}_{i}^{T} \mathbf{A}_{i, j}^{1,1} \mathbf{P}_{j}\right] T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right) \\
& +\mu\left[\sum_{i=1}^{3} \sum_{j=1}^{3} \mathbf{P}_{i}^{T} \mathbf{A}_{i, j}^{2,2} \mathbf{P}_{j}\right] T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right) \\
& +\mu\left[\sum_{i=1}^{3} \sum_{j=1}^{3} \mathbf{P}_{i}^{T} \mathbf{A}_{i, j}^{3,3} \mathbf{P}_{j}\right] T^{e}\left(\boldsymbol{U}_{\mathrm{t} x}\right), \tag{10}
\end{align*}
$$

where $\mathbf{P}_{1}=\mathbf{D} \otimes \mathbf{I} \otimes \mathbf{I}, \mathbf{P}_{2}=\mathbf{I} \otimes \mathbf{D} \otimes \mathbf{I}, \mathbf{P}_{3}=\mathbf{I} \otimes \mathbf{I} \otimes \mathbf{D}, \mathbf{I}$ is the $(n+1)$ th order identity matrix, and the $(n+1) \times(n+1)$ matrix $\mathbf{D}$ is $\left[\frac{\partial \varphi_{i}\left(\xi_{j}\right)}{\partial \xi}\right]$. The superscripts of $\mathbf{A}$ represent the physical coordinates, whereas the subscripts denote the isoparametric coordinates. $\mathbf{A}_{1,1}^{1,1}$ is given by
$\mathbf{A}_{1,1}^{1,1}=\left[\left.\delta_{r i} \delta_{s j} \delta_{t k} \omega_{i} \omega_{j} \omega_{k} \frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial x} \operatorname{Det}(\mathbf{J})\right|_{\left(\xi_{i}, \gamma_{j}, \eta_{k}\right)}\right]$,
where $\delta$ is the Kronecker-delta symbol, $r, s$, and $t$ represents matrix row indices, $i, j$, and $k$ denote matrix column indices, and $\operatorname{Det}(\mathbf{J})$ is the determinant of the Jacobi matrix evaluated at GLL point $\left(\xi_{i}, \gamma_{j}, \eta_{k}\right)$. Equation (11) shows that $\mathbf{A}$ is a diagonal matrix, and only $(n+1)^{3}$ elements need to be stored for each A. For every hexahedral element, 81 combinations may appear for A. However, 45 matrices need to be stored because of the symmetry of the matrices.

### 2.2 EBE-SEM for the PML formula

In our previous work, a second-order PML absorbing boundary condition (PML ABC) was formulated using the mixedgrid finite-element method (Liu et al., 2014). Here, we use this type of PML ABC to absorb the scattered wave field $\boldsymbol{u}_{\mathrm{s}}=\left(\boldsymbol{u}_{\mathrm{s} x}, \boldsymbol{u}_{\mathrm{s} y}, \boldsymbol{u}_{\mathrm{s} z}\right)$. To simplify the discussion, only the for-
mula of the PML ABC for the displacement of the $x$ component is shown below:

$$
\left\{\begin{array}{l}
\rho\left(\ddot{u}_{\mathrm{s} x, 1}+2 d_{x} \dot{u}_{\mathrm{s} x, 1}+d_{x}^{2} u_{\mathrm{s} x, 1}\right)  \tag{12}\\
\quad=\frac{\partial}{\partial x}\left((\lambda+2 \mu) \frac{\partial u_{\mathrm{s} x}}{\partial x}\right)+\rho P_{x, 1}, \\
\rho \dot{P}_{x, 1}+\rho d_{x} P_{x, 1}=-(\lambda+2 \mu) d_{x}^{\prime} \frac{\partial u_{x}}{\partial x}, \\
\rho\left(\ddot{u}_{\mathrm{s} x, 2}+2 d_{y} \dot{u}_{\mathrm{s} x, 2}+d_{y}^{2} u_{\mathrm{s} x, 2}\right)=\frac{\partial}{\partial y}\left(\mu \frac{\partial u_{\mathrm{s} x}}{\partial y}\right)+\rho P_{x, 2}, \\
\rho \dot{P}_{x, 2}+\rho d_{y} P_{x, 2}=-\mu d_{y}^{\prime} \frac{\partial u_{\mathrm{s} x}}{\partial y}, \\
\rho\left(\ddot{u}_{\mathrm{s} x, 3}+2 d_{z} \dot{u}_{\mathrm{s} x, 3}+d_{z}^{2} u_{\mathrm{s} x, 3}\right)=\frac{\partial}{\partial z}\left(\mu \frac{\partial u_{\mathrm{s} x}}{\partial z}\right)+\rho P_{x, 3} \\
\rho \dot{P}_{x, 3}+\rho d_{z} P_{x, 3}=-\mu d_{z}^{\prime} \frac{\partial u_{x}}{\partial z}, \\
\rho\left(\ddot{u}_{\mathrm{s} x, 4}+\left(d_{x}+d_{y}\right) \dot{u}_{\mathrm{s} x, 4}+d_{x} d_{y} u_{\mathrm{s} x, 4}\right) \\
\quad=\frac{\partial}{\partial x}\left(\lambda \frac{\partial u_{\mathrm{s} y}}{\partial y}\right)+\frac{\partial}{\partial y}\left(\mu \frac{\partial u_{\mathrm{s} y}}{\partial x}\right) \\
\rho\left(\ddot{u}_{\mathrm{s} x, 5}+\left(d_{x}+d_{z}\right) \dot{u}_{\mathrm{s} x, 5}+d_{x} d_{z} u_{\mathrm{s} x, 5}\right) \\
\quad=\frac{\partial}{\partial x}\left(\lambda \frac{\partial u_{\mathrm{s} z}}{\partial z}\right)+\frac{\partial}{\partial z}\left(\mu \frac{\partial u_{\mathrm{s} z}}{\partial x}\right) \\
u_{\mathrm{s} x}=u_{\mathrm{s} x, 1}+u_{\mathrm{s} x, 2}+u_{\mathrm{s} x, 3}+u_{\mathrm{s} x, 4}+u_{\mathrm{s} x, 5}
\end{array}\right.
$$

where $d_{x}, d_{y}$, and $d_{z}$ are damping coefficients along the three coordinate axes, the subscript of $d$ denotes the first spatial derivative of the damping coefficient, $u_{\mathrm{s} x, 1}, u_{\mathrm{s} x, 2}, u_{\mathrm{s} x, 3}$, $u_{\mathrm{s} x, 4}$, and $u_{\mathrm{s} x, 5}$ are the five split components of $u_{x}$, and $P_{x, 1}$, $P_{x, 2}$, and $P_{x, 3}$ are the three intermediate variables. Based on the natural boundary condition, the corresponding weak form of Eq. (12) is

$$
\begin{align*}
& \int_{\Omega} \rho v_{x}\left(\ddot{u}_{\mathrm{s} x, 1}+2 d_{x} \dot{u}_{\mathrm{s} x, 1}+d_{x}^{2} u_{\mathrm{s} x, 1}\right) \mathrm{d} \Omega \\
& +\int_{\Omega}(\lambda+2 \mu) \frac{\partial v_{x}}{\partial x} \frac{\partial u_{\mathrm{s} x}}{\partial x} \mathrm{~d} \Omega=\int_{\Omega} \rho v_{x} P_{x, 1} \mathrm{~d} \Omega, \\
& \int_{\Omega} \rho v_{x}\left(\dot{P}_{x, 1}+d_{x} P_{x, 1}\right) \mathrm{d} \Omega \\
& +\int_{\Omega} v_{x}(\lambda+2 \mu) d_{x}^{\prime} \frac{\partial u_{\mathrm{s} x}}{\partial x} \mathrm{~d} \Omega=0, \\
& \int_{\Omega} \rho v_{x}\left(\ddot{u}_{\mathrm{s} x, 2}+2 d_{y} \dot{u}_{\mathrm{s} x, 2}+d_{y}^{2} u_{\mathrm{s} x, 2}\right) \mathrm{d} \Omega \\
& +\int_{\Omega} \mu \frac{\partial v_{x}}{\partial y} \frac{\partial u_{\mathrm{s} x}}{\partial y} \mathrm{~d} \Omega \\
& =\int_{\Omega} \rho v_{x} P_{x, 2} \mathrm{~d} \Omega, \\
& \int_{\Omega} \rho v_{x}\left(\stackrel{\Omega}{P}_{x, 2}+d_{y} P_{x, 2}\right) \mathrm{d} \Omega \\
& +\int_{\Omega} v_{x} \mu d_{y}^{\prime} \frac{\partial u_{\mathrm{s} x}}{\partial y} \mathrm{~d} \Omega=0, \\
& \int_{\Omega} \rho v_{x}\left(\ddot{u}_{\mathrm{s} x, 3}+2 d_{z} \dot{u}_{\mathrm{s} x, 3}+d_{z}^{2} u_{\mathrm{s} x, 3}\right) \mathrm{d} \Omega  \tag{13}\\
& +\int_{\Omega} \mu \frac{\partial v_{x}}{\partial z} \frac{\partial u_{\mathrm{s} x}}{\partial z} \mathrm{~d} \Omega=\int_{\Omega} \rho v_{x} P_{x, 3} \mathrm{~d} \Omega, \\
& \int_{\Omega} \rho v_{x}\left(\dot{P}_{x, 3}+d_{z} P_{x, 3}\right) \mathrm{d} \Omega \\
& +\int_{\Omega} v_{x} \mu d_{z}^{\prime} \frac{\partial u_{\mathrm{s} x}}{\partial z} \mathrm{~d} \Omega=0, \\
& \int \rho v_{x}\left(\ddot{u}_{\mathrm{s} x, 4}+\left(d_{x}+d_{y}\right) \dot{u}_{\mathrm{s} x, 4}+d_{x} d_{y} u_{\mathrm{s} x, 4}\right) \mathrm{d} \Omega \\
& +\int_{\Omega} \lambda \frac{\partial v_{x}}{\partial x} \frac{\partial u_{s y}}{\partial y} \mathrm{~d} \Omega+\int_{\Omega} \mu \frac{\partial v_{x}}{\partial y} \frac{\partial u_{s y}}{\partial x} \mathrm{~d} \Omega=0, \\
& \int_{\Omega} \rho v_{x}\left(\ddot{u}_{\mathrm{s} x, 5}+\left(d_{x}+d_{z}\right) \dot{u}_{\mathrm{s} x, 5}+d_{x} d_{z} u_{\mathrm{s} x, 5}\right) \mathrm{d} \Omega \\
& \begin{array}{l}
+\int_{\Omega} \lambda \frac{\partial v_{x}}{\partial x} \frac{\partial u_{\mathrm{s} z}}{\partial z} \mathrm{~d} \Omega+\int_{\Omega} \mu \frac{\partial v_{x}}{\partial z} \frac{\partial u_{\mathrm{s} z}}{\partial x} \mathrm{~d} \Omega=0, \\
=u_{\mathrm{s} x, 1}+u_{\mathrm{s} x, 2}+u_{\mathrm{s} x, 3}+u_{\mathrm{s} x, 4}+u_{\mathrm{s} x, 5}
\end{array}
\end{align*}
$$

The discretized version of Eq. (13) is

$$
\left\{\begin{array}{l}
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \ddot{\boldsymbol{U}}_{\mathrm{s} x, 1}^{e}+\sum_{e=1}^{N} \mathbf{D}_{x}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 1}^{e}+\sum_{e=1}^{N} \mathbf{D}_{x x}^{e} \boldsymbol{U}_{\mathrm{s} x, 1}^{e} \\
\quad+\sum_{e=1}^{N} \mathbf{K}_{x x}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \mathbf{P}_{x, 1}^{e}, \\
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \dot{\mathbf{P}}_{x, 1}^{e}+\frac{1}{2} \sum_{e=1}^{N} \mathbf{D}_{x}^{e} \mathbf{P}_{x, 1}^{e}+\sum_{e=1}^{N} \mathbf{K}_{x}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0, \\
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \ddot{U}_{\mathrm{s} x, 2}^{e}+\sum_{e=1}^{N} \mathbf{D}_{y}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 2}^{e}+\sum_{e=1}^{N} \mathbf{D}_{y y}^{e} \boldsymbol{U}_{\mathrm{s} x, 2}^{e} \\
\quad+\sum_{e=1}^{N} \mathbf{K}_{y y}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \mathbf{P}_{x, 2}^{e}, \\
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \dot{\mathbf{P}}_{x, 2}^{e}+\frac{1}{2} \sum_{e=1}^{N} \mathbf{D}_{y}^{e} \mathbf{P}_{x, 2}^{e}+\sum_{e=1}^{N} \mathbf{K}_{y}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0,  \tag{14}\\
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \ddot{\boldsymbol{U}}_{\mathrm{s} x, 3}^{e}+\sum_{e=1}^{N} \mathbf{D}_{z}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 3}^{e}+\sum_{e=1}^{N} \mathbf{D}_{z z}^{e} \boldsymbol{U}_{\mathrm{s} x, 3}^{e} \\
\quad+\sum_{e=1}^{N} \mathbf{K}_{z z}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \mathbf{P}_{x, 3}^{e}, \\
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \dot{\mathbf{P}}_{x, 3}^{e}+\frac{1}{2} \sum_{e=1}^{N} \mathbf{D}_{z}^{e} \mathbf{P}_{x, 3}^{e}+\sum_{e=1}^{N} \mathbf{K}_{z}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0, \\
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \ddot{\boldsymbol{U}}_{\mathrm{s} x, 4}^{e}+\sum_{e=1}^{N} \mathbf{D}_{x, y}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 4}^{e}+\sum_{e=1}^{N} \mathbf{D}_{x y}^{e} \boldsymbol{U}_{\mathrm{s} x, 4}^{e} \\
\quad+\sum_{e=1}^{N}\left(\mathbf{K}_{x y}^{e}+\mathbf{K}_{y x}^{e}\right) \boldsymbol{U}_{\mathrm{s} y}^{e}=0, \\
\sum_{e=1}^{N} \mathbf{M}_{1}^{e} \ddot{\boldsymbol{U}}_{\mathrm{s} x, 5}^{e}+\sum_{e=1}^{N} \mathbf{D}_{x, z}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 5}^{e} \\
\quad+\sum_{e=1}^{N} \mathbf{D}_{x z}^{e} \boldsymbol{U}_{\mathrm{s} x, 5}^{e}+\sum_{e=1}^{N}\left(\mathbf{K}_{x z}^{e}+\mathbf{K}_{z x}^{e}\right) \boldsymbol{U}_{\mathrm{s} z}^{e}=0, \\
\boldsymbol{U}_{\mathrm{s} x}=\boldsymbol{U}_{\mathrm{s} x, 1}+\boldsymbol{U}_{\mathrm{s} x, 2}+\boldsymbol{U}_{\mathrm{s} x, 3}+\boldsymbol{U}_{\mathrm{s} x, 4}+\boldsymbol{U}_{\mathrm{s} x, 5},
\end{array}\right.
$$

where $\mathbf{K}_{x}^{e}, \mathbf{K}_{y}^{e}, \mathbf{K}_{z}^{e}, \mathbf{K}_{x x}^{e}, \mathbf{K}_{y y}^{e}, \mathbf{K}_{z z}^{e}, \mathbf{K}_{x y}^{e}, \mathbf{K}_{y x}^{e}, \mathbf{K}_{x z}^{e}$, and $\mathbf{K}_{z x}^{e}$ are element stiffness matrices and $\mathbf{D}_{x}^{e}, \mathbf{D}_{y}^{e}, \mathbf{D}_{z}^{e}, \mathbf{D}_{x x}^{e}, \mathbf{D}_{y y}^{e}, \mathbf{D}_{z z}^{e}$, $\mathbf{D}_{x y}^{e}, \mathbf{D}_{x z}^{e}, \mathbf{D}_{x, y}^{e}$, and $\mathbf{D}_{x, z}^{e}$ are the element damping matrices. The detailed expressions of these element matrices are presented in Appendix B. The damping matrices are also diagonal matrices, whose element values are scaled by a constant compared with the element values of $\mathbf{M}^{e}$. If we denote the global damping matrices as $\mathbf{D}_{x}, \mathbf{D}_{y}, \mathbf{D}_{z}, \mathbf{D}_{x x}, \mathbf{D}_{y y}, \mathbf{D}_{z z}, \mathbf{D}_{x y}$, $\mathbf{D}_{x z}, \mathbf{D}_{x, y}$, and $\mathbf{D}_{x, z}$, then we obtain the element-by-element scheme for Eq. (14):

$$
\begin{aligned}
& \ddot{U}_{\mathrm{s} x, 1}+\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{x}\right) \dot{\boldsymbol{U}}_{\mathrm{s} x, 1}^{e} \\
& +\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{x x}\right) \boldsymbol{U}_{\mathrm{s} x, 1}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{x x}^{e} \boldsymbol{U}_{\mathbf{s} x}^{e}=\mathbf{P}_{x, 1}, \\
& \dot{\mathbf{P}}_{x, 1}+\frac{1}{2} \sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{x}\right) \mathbf{P}_{x, 1}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{x}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0, \\
& \ddot{U}_{\mathrm{s} x, 2}+\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{y}\right) \dot{\boldsymbol{U}}_{\mathrm{s} x, 2}^{e} \\
& +\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{y y}\right) \boldsymbol{U}_{\mathrm{sx}, 2}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{y y}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\mathbf{P}_{x, 2}, \\
& \dot{\mathbf{P}}_{x, 2}+\frac{1}{2} \sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{y}\right) \mathbf{P}_{x, 2}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{y}^{e} \boldsymbol{U}_{\mathbf{s} \boldsymbol{x}}^{e}=0, \\
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 3}+\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{z}\right) \dot{\boldsymbol{U}}_{\mathrm{s} x, 3}^{e} \\
& +\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{z z}\right) \boldsymbol{U}_{\mathrm{s} x, 3}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{z z}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\mathbf{P}_{x, 3}, \\
& \dot{\mathbf{P}}_{x, 3}+\frac{1}{2} \sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{z}\right) \mathbf{P}_{x, 3}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{z}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0, \\
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 4}+\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{x, y}\right) \dot{\boldsymbol{U}}_{\mathrm{s} x, 4}^{e} \\
& +\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{x y}\right) \boldsymbol{U}_{\mathrm{s} x, 4}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N}\left(\mathbf{K}_{x y}^{e}+\mathbf{K}_{y x}^{e}\right) \boldsymbol{U}_{\mathrm{s} y}^{e}=0, \\
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 5}+\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{x, z}\right) \dot{\boldsymbol{U}}_{\mathrm{s} x, 5}^{e} \\
& +\sum_{e=1}^{N} T^{e}\left(\mathbf{M}_{1}\right)^{-1} T^{e}\left(\mathbf{D}_{x z}\right) \boldsymbol{U}_{\mathrm{s} x, 5}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N}\left(\mathbf{K}_{x z}^{e}+\mathbf{K}_{z x}^{e}\right) \boldsymbol{U}_{\mathrm{sz}}^{e}=0, \\
& \boldsymbol{U}_{\mathrm{s} x}=\boldsymbol{U}_{\mathrm{s} x, 1}+\boldsymbol{U}_{\mathrm{s} x, 2}+\boldsymbol{U}_{\mathrm{s} x, 3}+\boldsymbol{U}_{\mathrm{s} x, 4}+\boldsymbol{U}_{\mathrm{s} x, 5} .
\end{aligned}
$$

From Eqs. (B11) to (A20), Eq. (15) can be rewritten as

$$
\begin{align*}
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 1}+\sum_{e=1}^{N} 2 d_{x}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 1}^{e}+\sum_{e=1}^{N}\left(d_{x}^{e}\right)^{2} \boldsymbol{U}_{\mathrm{s} x, 1}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{x x}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\mathbf{P}_{x, 1}, \\
& \dot{\mathbf{P}}_{x, 1}+\sum_{e=1}^{N} d_{x}^{e} \mathbf{P}_{x, 1}^{e}+\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{x}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0, \\
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 2}+\sum_{e=1}^{N} 2 d_{y}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 2}^{e}+\sum_{e=1}^{N}\left(d_{y}^{e}\right)^{2} \boldsymbol{U}_{\mathrm{s} x, 2}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{y y}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\mathbf{P}_{x, 2}, \\
& \dot{\mathbf{P}}_{x, 2}+\sum_{e=1}^{N} d_{y}^{e} \mathbf{P}_{x, 2}^{e}+\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{y}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0, \\
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 3}+\sum_{e=1}^{N} 2 d_{z}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 3}^{e}+\sum_{e=1}^{N}\left(d_{z}^{e}\right)^{2} \boldsymbol{U}_{\mathrm{s} x, 3}^{e}  \tag{16}\\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{z z}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=\mathbf{P}_{x, 3}, \\
& \dot{\mathbf{P}}_{x, 3}+\sum_{e=1}^{N} d_{z}^{e} \mathbf{P}_{x, 3}^{e}+\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N} \mathbf{K}_{z}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}=0, \\
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 4}+\sum_{e=1}^{N}\left(d_{x}^{e}+d_{y}^{e}\right) \dot{\boldsymbol{U}}_{\mathrm{s} x, 4}^{e}+\sum_{e=1}^{N} d_{x}^{e} d_{y}^{e} \boldsymbol{U}_{\mathrm{s} x, 4}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N}\left(\mathbf{K}_{x y}^{e}+\mathbf{K}_{y x}^{e}\right) \boldsymbol{U}_{\mathrm{s} y}^{e}=0, \\
& \ddot{\boldsymbol{U}}_{\mathrm{s} x, 5}+\sum_{e=1}^{N}\left(d_{x}^{e}+d_{z}^{e}\right) \dot{\boldsymbol{U}}_{\mathrm{s} x, 5}^{e}+\sum_{e=1}^{N} d_{x}^{e} d_{z}^{e} \boldsymbol{U}_{\mathrm{s} x, 5}^{e} \\
& +\left(\mathbf{M}_{1}\right)^{-1} \sum_{e=1}^{N}\left(\mathbf{K}_{x z}^{e}+\mathbf{K}_{z x}^{e}\right) \boldsymbol{U}_{\mathrm{s} z}^{e}=0 \\
& \boldsymbol{U}_{\mathrm{s} x}=\boldsymbol{U}_{\mathrm{s} x, 1}+\boldsymbol{U}_{\mathrm{s} x, 2}+\boldsymbol{U}_{\mathrm{s} x, 3}+\boldsymbol{U}_{\mathrm{s} x, 4}+\boldsymbol{U}_{\mathrm{s} x, 5} .
\end{align*}
$$

From Eq. (16), we observe that the element damping matrices do not need to be stored. It should be noted that the value of the damping coefficient is divided by $N_{e}$ if the node located on the element surface is shared by $N_{e}$ elements. In each element, the stiffness matrix can also be written as a tensor product of matrices, which is similar to Eq. (10).

### 2.3 Discussion of the EBE-SEM

Although the EBE-SEM is specially designed for teleseismic wave modeling (i.e., Eq. 9 is for teleseismic total wave field propagation if the proper teleseismic incident boundary condition is added and Eq. 16 is for absorbing the scatted wave field), EBE-SEM can be directly used for wave field simulation of an earthquake that occurred in the interior of the model (computational domain) if a source term is added to Eq. (9).

Seriani (1997) is the seminal work that introduced the EBE-SEM. In the 2-D case, the element-by-element scheme is combined with the Chebyshev orthogonal polynomialbased SEM. Because the Chebyshev orthogonal polynomial is orthogonal and associated with the weight $1 / \sqrt{1-\xi^{2}}$, the Chebyshev orthogonal polynomial-based SEM cannot lead
to a diagonal mass matrix and an iterative algorithm is generally used to solve a large spare linear system of equations, which may not be efficient for larger-scale seismic waveform modeling. Another aspect of the EBE algorithm of Seriani (1997) is that it is only appropriate for use with only rectangular elements of a fixed shape; this restriction may be problematic for curved elements. In contrast, EBE-SEM in the paper can be used for general cases.
The recent improvements in the SPECFEM3D software also incorporate the tensor products of the element stiffness matrix (Peter et al., 2011). Geology models can be decomposed by fully unstructured hexahedral meshes. Great load balancing is achieved based on graph partitioning. This software does not explicitly assemble the global stiffness matrix. The main contribution of this paper is the detailed introduction of EBE-SEM for high-efficient teleseismic wave modeling.

If appropriate boundary conditions are added for the computational domain and the PML domain, then the computation will be isolated in the counterpart domain, i.e., Eq. (9) for the computational domain and Eq. (16) for the PML domain. The boundary conditions are discussed presented below.

## 3 Teleseismic wave incident boundary conditions

For the completeness of this paper, we first simply introduce the FK method for determining the 3-D elastic-wave-equation-based plane-wave propagation in layered media. For more details, the reader is referred to Haskell (1953) and Tong et al. (2014a). Our focus is to construct the teleseismic wave incident boundary conditions and develop a highly efficient method for the storing the boundary wave fields.

### 3.1 Plane-wave propagation in 1-D layered media

Transforming the elastic wave equation into the $\left(\omega, k_{x}, k_{y}\right)$ domain, we obtain the following equation:
$-\rho \omega^{2}\left[\begin{array}{c}u_{\mathrm{FK} x} \\ u_{\mathrm{FK} y} \\ u_{\mathrm{FK} z}\end{array}\right]=\left[\begin{array}{l}i k_{x} \sigma_{x x}+i k_{y} \sigma_{x y}+\partial_{z} \sigma_{x z} \\ i k_{x} \sigma_{x y}+i k_{y} \sigma_{y y}+\partial_{z} \sigma_{y z} \\ i k_{x} \sigma_{x z}+i k_{y} \sigma_{y z}+\partial_{z} \sigma_{z z}\end{array}\right]$
where $\boldsymbol{u}_{\mathrm{FK}}=\left(u_{\mathrm{FK} x}, u_{\mathrm{FK} y}, u_{\mathrm{FK} z}\right)$ is the plane wave in layered media, $k_{x}$ and $k_{y}$ are the wavenumbers in the $x$ and $y$ directions, respectively, and $\omega$ is the angular frequency. We assign $\boldsymbol{u}_{\mathrm{FK}}$ the same notation in both the time-space and frequency-wavenumber domains to avoid clustering. If we define $u_{\mathrm{FKH}} \boldsymbol{e}_{\mathrm{H}}=u_{\mathrm{FK} x} \boldsymbol{e}_{x}+u_{\mathrm{FK} y} \boldsymbol{e}_{y}$ and $k \boldsymbol{e}_{\mathrm{H}}=k_{x} \boldsymbol{e}_{x}+k_{y} \boldsymbol{e}_{y}$, where $k=\sqrt{k_{x}^{2}+k_{y}^{2}}$ and $\boldsymbol{e}_{\mathrm{H}}$ is the horizontal normal vector, then Eq. (17) can be written as

$$
-\omega^{2}\left[\begin{array}{c}
u_{\mathrm{FKH}}  \tag{18}\\
u_{\mathrm{FK} z}
\end{array}\right]=\left[\begin{array}{c}
i k_{\mathrm{H}} \sigma_{\mathrm{HH}}+\partial_{z} \sigma_{\mathrm{H} z} \\
i k_{\mathrm{H}} \sigma_{\mathrm{H} z}+\partial_{z} \sigma_{z z}
\end{array}\right]
$$



Figure 2. 1-D layered background model. The thickness, P wave velocity, $S$ wave velocity, and density of each layer are shown.

If we define $y_{1}=i u_{\mathrm{FKH}}, y_{2}=u_{\mathrm{FK} z}, y_{3}=i \sigma_{\mathrm{H} z}$, and $y_{4}=$ $\sigma_{z z}$, Eq. (18) can be changed to the following ordinary differential equations:

$$
\begin{align*}
\frac{d}{\mathrm{~d} z}\left[\begin{array}{l}
y_{1} \\
y_{2} \\
y_{3} \\
y_{4}
\end{array}\right] & =\left[\begin{array}{llll}
0 & k & \frac{1}{\mu} & 0 \\
-\frac{\lambda k}{\lambda+2 \mu} & 0 & 0 & \frac{1}{\lambda+2 \mu} \\
\frac{4 k^{2} \mu(\lambda+\mu)}{\lambda+2 \mu}-\rho \omega^{2} & 0 & 0 & \frac{k \lambda}{\lambda+2 \mu} \\
0 & -\rho \omega^{2} & -k & 0
\end{array}\right] \\
& \cdot\left[\begin{array}{l}
y_{1} \\
y_{2} \\
y_{3} \\
y_{4}
\end{array}\right]=\boldsymbol{F} y . \tag{19}
\end{align*}
$$

By calculating the eigenvalues and eigenvectors of $\boldsymbol{F}$, the general solution of Eq. (19) can be written as

$$
\begin{align*}
\boldsymbol{y} & =\left[\begin{array}{cccc}
1 & 1 & 1 & 1 \\
-i \frac{k}{v_{S}} & i \frac{k}{v_{S}} & i \frac{v_{P}}{k} & -i \frac{v_{P}}{k} \\
i \frac{\mu \gamma}{v_{S}} & -i \frac{\mu \gamma}{v_{S}} & -2 i \mu v_{P} & 2 i \mu v_{P} \\
-2 k \mu & -2 k \mu & -\frac{\mu \gamma}{k} & -\frac{\mu \gamma}{k}
\end{array}\right] \\
& \cdot\left[\begin{array}{llll}
e^{-i v_{S} z} & 0 & 0 & 0 \\
0 & e^{i v_{S} z} & 0 & 0 \\
0 & 0 & e^{-i v_{P} z} & 0 \\
0 & 0 & 0 & e^{i v_{P} z}
\end{array}\right]\left[\begin{array}{l}
C_{1} \\
C_{2} \\
C_{3} \\
C_{4}
\end{array}\right], \tag{20}
\end{align*}
$$

where $C_{1}, C_{2}, C_{3}$, and $C_{4}$ correspond to the amplitudes of the upgoing S wave, downgoing S wave, upgoing P wave, downgoing P wave, respectively; $p=\frac{k}{\omega}, v_{P}=\omega \sqrt{\frac{1}{\alpha^{2}}-p^{2}}$, $v_{S}=\omega \sqrt{\frac{1}{\beta^{2}}-p^{2}}$, and $\gamma=\left(2 k^{2}-\frac{\omega^{2}}{\beta^{2}}\right)$ are the ray parameters. As shown in Fig. 2, the velocities, density, and thickness of the $m$ th layer are denoted as $V_{P m}, V_{S m}, \rho_{m}$, and $h_{m}$, respectively. We define the following matrices for the $m$ th
layer:
$\boldsymbol{R}_{m}=\left[\begin{array}{cccc}1 & 1 & 1 & 1 \\ -i \frac{k}{v_{S m}} & i \frac{k}{v_{S m}} & i \frac{v_{P m}}{k} & -i \frac{v_{P m}}{k} \\ i \frac{\mu_{m} \gamma}{v_{S}} & -i \frac{\mu_{m} \gamma}{v_{S}} & -2 i \mu_{m} v_{P m} & 2 i \mu_{m} v_{P m} \\ -2 k \mu_{m} & -2 k \mu_{m} & -\frac{\mu_{m} \gamma_{m}}{k} & -\frac{\mu_{m} \gamma_{m}}{k}\end{array}\right]$,
$\boldsymbol{H}\left(h_{m}\right)=\left[\begin{array}{llll}e^{-i v_{S} z} & 0 & 0 & 0 \\ 0 & e^{i v_{S} z} & 0 & 0 \\ 0 & 0 & e^{-i v_{P} z} & 0 \\ 0 & 0 & 0 & e^{i v_{P} z}\end{array}\right]$.
The relationship between the wave fields at $z=z_{m-1}$ and $z=$ $z_{m}$ can be written as
$\boldsymbol{y}_{m-1}=\boldsymbol{R}_{m} \mathbf{H}\left(-h_{m}\right) \mathbf{R}_{m}^{-1} \boldsymbol{y}_{m}$.
In Eq. (23), $\mathbf{L}_{m}=\mathbf{R}_{m} \mathbf{H}\left(-h_{m}\right) \mathbf{R}_{m}^{-1}$ is the propagation matrix. The wave field at the free surface can be represented by
$\boldsymbol{y}_{0}=\mathbf{L}_{1} \cdots \mathbf{L}_{n} \mathbf{R}_{n+1} \mathbf{C}$.
If we consider that the incident plane wave is an upgoing P wave with unit amplitude, then we have
$\left\{\begin{array}{l}C_{1}=0, \\ C_{3}= \pm i \sin (\theta) .\end{array}\right.$
We define
$\mathbf{A}=\mathbf{L}_{1} \cdots \mathbf{L}_{n} \mathbf{R}_{n+1}=\left[\begin{array}{cccc}a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44}\end{array}\right]$.
The values of $y_{3}$ and $y_{4}$ at the free surface can be written as
$\left[\begin{array}{l}y_{3} \\ y_{4}\end{array}\right]_{z=0}=\left[\begin{array}{llll}a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44}\end{array}\right]\left[\begin{array}{l}0 \\ C_{2} \\ C_{3} \\ C_{4}\end{array}\right]$.
Based on the free surface boundary condition $y_{3}=y_{4}=0$, we have
$\left[\begin{array}{l}C_{2} \\ C_{4}\end{array}\right]=-C_{3}\left[\begin{array}{ll}a_{32} & a_{34} \\ a_{42} & a_{44}\end{array}\right]^{-1}\left[\begin{array}{l}a_{33} \\ a_{43}\end{array}\right]$.
Substituting Eqs. (25) and (28) into Eq. (23), we obtain the wave field at $z=z_{m}$. Based on Eq. (23), we can calculate the wave field in each layer.

### 3.2 Incident boundary conditions

We assume that the infinite space is composed of the following hexahedron set:
$B=\left\{e_{1}, \ldots, e_{N_{1}}, e_{N_{1}+1}, \ldots, e_{N_{1}+N_{2}}, \ldots\right\}$.

The first $N_{1}$ elements of $B$ compose the computational domain (the domain bounded by the blue lines in Fig. 1), and the elements from $N_{1}+1$ to $N_{1}+N_{2}$ compose the PML domain. We define the set of elements in the computational domain as
$B_{C}=\left\{e_{1}, \ldots, e_{N_{1}}\right\}$.

A total of $N_{3}$ elements in $B_{C}$ that contact the boundary of PML domain are collected in the following set:
$B_{C, B}=\left\{e_{j_{1}}, \ldots, e_{j_{N_{3}}}\right\}$.

The set of $N_{2}$ elements that compose the PML domain is given by
$B_{P}=\left\{e_{N_{1}+1}, \ldots, e_{N_{1}+N_{2}}\right\}$.

A total of $N_{4}$ elements in $B_{P}$ that contact the boundary of the computational domain are gathered in the following set:
$B_{P, B}=\left\{e_{k_{1}}, \ldots, e_{k_{N_{4}}}\right\}$.

We define an operator that maps the elements of vector $\boldsymbol{A}$ to a target vector $\boldsymbol{A}_{\mathrm{t}}$ according to the numbering set $\boldsymbol{A}_{\mathrm{n}}$. The index number of each elements of set $\boldsymbol{A}_{\mathrm{t}}$ in set $\boldsymbol{A}$ is gathered in $\boldsymbol{A}_{\mathrm{n}}$, and we define
$\langle\boldsymbol{A}\rangle_{A_{\mathrm{n}}}=\boldsymbol{A}_{\mathrm{t}}$.

From Eq. (9), we have

$$
\begin{align*}
\left\langle\ddot{\boldsymbol{U}}_{\mathrm{t}}\right\rangle_{A_{C}} & =\left\langle-\sum_{e \in B}\left(T^{e}\left(\mathbf{M}_{1}\right)^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C}} \\
& =\left\langle-\sum_{e \in B_{C} \cup B_{P, B}}\left(T^{e}\left(\mathbf{M}_{1}\right)^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C}}, \tag{35}
\end{align*}
$$

where $A_{C}$ is the numbering set of the nodes in the computational domain. Equation (35) shows that only matrix-vector products in the elements of the computational domain and the elements of the PML domain in contact with the boundary of the computational domain contribute to the acceleration wave field on the nodes in the computational domain.

Based on Eq. (35), we have

$$
\begin{align*}
\left\langle\ddot{\boldsymbol{U}}_{\mathrm{t}}\right\rangle_{A_{C, B}} & =\left\langle-\sum_{e \in B_{C} \cup B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C, B}} \\
& =\left\langle-\sum_{e \in B_{C, B} \cup B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C, B}} \\
& =\left\langle-\sum_{e \in B_{C, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C, B}} \\
& +\left\langle-\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C, B}}, \tag{36}
\end{align*}
$$

where $A_{C, B}$ is the numbering set of nodes of the computational domain located on the interface between the computational and PML domains. The second term of the right side of Eq. (36) can be written as

$$
\begin{align*}
\langle- & \left.\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C, B}} \\
& =\left\langle-\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{FK}+\boldsymbol{U}_{\mathrm{s}}}\right)\right)\right\rangle_{A_{C, B}} \\
& =\left\langle-\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{FK}}\right)\right)\right\rangle_{A_{C, B}}  \tag{37}\\
& +\left\langle-\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{s}}\right)\right)\right\rangle_{A_{C, B}} .
\end{align*}
$$

Equation (37) is the plane-wave incident condition for the computational domain. Based on Eq. (37), we only need to store the boundary wave field $\left\langle-\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{FK}}\right)\right)\right\rangle_{A_{C, B}} \quad$ to construct the incident boundary condition for the computational domain. The length of $\left\langle-\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{FK}}\right)\right)\right\rangle_{A_{C, B}}$ is 3 times as long as the number of nodes located on the interface between the computational and PML domains; and it is not necessary to store $\boldsymbol{U}_{\mathrm{FK}}$ on the nodes of the elements in $B_{P, B}$. This storage technique can decreased the amount of memory required by $\frac{n-1}{n} \times 100 \%$, which is extremely important for large-scale models. This storage technique is similar to the method of the linear combination of a boundary wave field, which was proposed in our previous work (Liu et al., 2015).

To simplify the discussion, the boundary condition of the first equation of Eq. (16) is discussed. From Eq. (16), we have

$$
\begin{align*}
\left\langle\ddot{\boldsymbol{U}}_{\mathrm{s} x, 1}\right\rangle_{A_{P}} & +\left\langle\sum_{e \in B} 2 d_{x}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 1}^{e}\right\rangle_{A_{P}} \\
& +\left\langle\sum_{e \in B}\left(d_{x}^{e}\right)^{2} \boldsymbol{U}_{\mathrm{s} x, 1}^{e}\right\rangle_{A_{P}} \\
& +\left\langle\sum_{e \in B} T^{e}(\mathbf{M})^{-1} \mathbf{K}_{x x}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}\right\rangle_{A_{P}}=\left\langle\boldsymbol{P}_{x, 1}\right\rangle_{A_{P}} \tag{38}
\end{align*}
$$

where $A_{P}$ is the numbering set of nodes in the PML domain. The scattered wave on the interface between the computational domain and the PML domain can be obtained by the following equation:

$$
\begin{equation*}
\left\langle\boldsymbol{U}_{\mathrm{s} x}\right\rangle_{A_{P, B}}=\left\langle\boldsymbol{U}_{\mathrm{t} x}\right\rangle_{A_{P, B}}-\left\langle\boldsymbol{U}_{\mathrm{FK} x}\right\rangle_{A_{P, B}} \tag{39}
\end{equation*}
$$

where $A_{P, B}$ is the numbering set of the nodes of the PML domain on the interface. $\left\langle\boldsymbol{U}_{\mathrm{s} x}\right\rangle_{A_{P, B}}$ does not need to be calculated by Eq. (16). From Eqs. (38) and (39), we have

$$
\begin{align*}
\left\langle\ddot{\boldsymbol{U}}_{\mathrm{s} x, 1}\right\rangle_{A_{P}-A_{P, B}} & +\left\langle\sum_{e \in B_{P}}\left(2 d_{x}^{e} \dot{\boldsymbol{U}}_{\mathrm{s} x, 1}^{e}+\left(d_{x}^{e}\right)^{2} \boldsymbol{U}_{\mathrm{s} x, 1}^{e}\right)\right\rangle_{A_{P}-A_{P, B}} \\
& +\left\langle\sum_{e \in B_{P}} T^{e}(\mathbf{M})^{-1} \mathbf{K}_{P, x x}^{e} \boldsymbol{U}_{\mathrm{s} x}^{e}\right\rangle_{A_{P}-A_{P, B}} \\
& =\left\langle\boldsymbol{P}_{x}\right\rangle_{A_{P}-A_{P, B}} . \tag{40}
\end{align*}
$$

Considering the other equations of PML ABC , the boundary condition of PML ABC to absorb the scattered waves is

$$
\begin{equation*}
\left\langle\boldsymbol{U}_{\mathrm{s}}\right\rangle_{A_{P, B}}=\left\langle\boldsymbol{U}_{\mathrm{t}}\right\rangle_{A_{P, B}}-\left\langle\boldsymbol{U}_{\mathrm{FK}}\right\rangle_{A_{P, B}} . \tag{41}
\end{equation*}
$$

In addition to the boundary condition given by Eq. (41), the Dirichlet boundary condition is added to the outer boundaries of the PML domain, and the natural boundary condition is added to the planes that connect the free surface of the computational domain. Because of the boundary conditions, the element-level matrix-vector product is restricted to only the element in the computational domain and the PML domain. Because the boundary conditions in Eqs. (37) and (41) involve the plane-wave fields, we call Eqs. (37) and (41) teleseismic wave incident conditions.

## 4 Analysis of the computational costs

We use the model in Fig. 1 to quantitatively discuss the computational cost of EBE-SEM for teleseismic wave modeling. Because the thickness of the PML domain in our numerical examples is only three elements wide, the number of floating point operations is trivial compared with the computational domain, and the main computational cost of the PML domain
is from the storage requirement of the boundary condition (Eq. 41).

The model is decomposed into 75000 cubic elements with a size of $2 \mathrm{~km} \times 2 \mathrm{~km} \times 2 \mathrm{~km}$. The fourth-order interpolation polynomial is used in the space. The time interval is 0.01 s , and the total time for the numerical modeling is 100 s .

The floating point operations in the computational domain are mainly from the element matrix-vector product. A total of $1.0546875 \times 10^{10}$ floating point multiplication operations are required in each time step. If a global stiffness matrix is assembled, then the product of the global stiffness matrix and the global solution vector requires a total of $9.35974 \times 10^{9}$ floating point multiply operations. The computational burden of EBE-SEM is greater than that of the conventional SEM for two reasons: the element stiffness matrix and solution vector product are decomposed into a total of 189 submatrix and solution vector products (Eq. 10), and the acceleration on common nodes shared by the adjacent elements is calculated more than once. However, the number of operations of EBE-SEM increases by only $11.26 \%$. This increased computational amount may be compensated for by the high load balance of EBE-SEM in parallel computing, as will be discussed in the numerical examples.

The memory requirements of EBE-SEM and the conventional SEM for teleseismic wave modeling are presented in Table 1. A total of 13.12 GB is required to store the boundary wave field to construct the teleseismic incident conditions. If the classical compressed sparse row (CSR) storage format (Greathouse and Daga, 2014) is used to store the spare stiffness matrix of the conventional SEM, then the storage requirement is 38.76 GB , which is nearly 25 times as large as the storage requirement of EBE-SEM to store the 45 element submatrices. Two factors contribute to this storage difference. First, CSR must store the row and column information of the global stiffness matrix in addition to the storage of nonzero elements. Nearly 0.018 and 3.87 GB are required to store the row and column information, respectively. Second, the 45 submatrices for each element are diagonal, i.e., only $(n+1)^{3}$ nonzero elements of each submatrix must be stored.

## 5 Numerical examples

Three numerical examples are provided to validate the efficiency of EBE-SEM for teleseismic wave modeling. All three examples use the Gaussian source-time function with a cutoff frequency of 2 Hz (Tong et al., 2014a).

### 5.1 Benchmark for the 1-D crust-upper-mantle model

Except for the model parameters presented in Sect. 4, the material parameters are shown in Table 2; the incidence angle is $15^{\circ}$, and the azimuth angle is $100^{\circ}$. The cross section of the plane-wave front with the Moho is located at $(-400 \mathrm{~km}$, $0 \mathrm{~km}, 30 \mathrm{~km}$ ) at the initial time $t=0$. A seismic station is


Figure 3. Plane-wave incident on a 1-D crust-upper-mantle model. The snapshots from the top to the bottom are taken at $t=10,16,18$, and 22 s , respectively. The left, middle, and right panels are corresponding to snapshots of planes at $x=50 \mathrm{~km}, y=50 \mathrm{~km}$, and $z=30 \mathrm{~km}$, respectively. The yellow arrows in $(\mathbf{j})$ and $(\mathbf{k})$ denote the weakly transmitted and reflected S waves. The color scale is shown on the right side of the figure.


Figure 4. Comparison of the waveforms generated by EBE-SEM (dotted line) and FK (solid line). The waveform computed by FK is considered the reference solution. Panels (a), (c), and (e) are the waveforms in the $x, y$, and $z$ directions, respectively. The error curves presented in the right panels are obtained by subtracting of the numerical solution from the reference solution. For visualization purposes, the errors are multiplied by 20 .

Table 1. Memory requirements of EBE-SEM and the conventional SEM for teleseismic wave modeling. All the values are stored in memory with single float precision. GB denotes gigabyte.

| Method | Stiffness matrix | $\left\langle-\sum_{e \in B_{P, B}}\left(T^{e}(\mathbf{M})^{-1} \mathbf{K}^{e} T^{e}\left(\boldsymbol{U}_{\mathrm{t}}\right)\right)\right\rangle_{A_{C, B}}$ | $\left\langle\boldsymbol{U}_{\mathrm{FK}}\right\rangle_{A_{P, B}}$ |
| :--- | :--- | :--- | :--- |
| EBE-SEM | 1.57 GB | 6.56 GB | 6.56 GB |
| General SEM | 38.76 GB | 6.56 GB | 6.56 GB |



Figure 5. Parallel efficiency of the EBE-SEM. The red line denotes the theoretical CPU time, and the blue line with circles represents the actual CPU times. The red arrows designate the abnormal CPU times compared with the neighboring CPU times.

Table 2. Material parameters of the crust-upper-mantle model.

|  | Density $\rho$ <br> $\left(\mathrm{kg} \mathrm{m}^{3}\right)$ | $V_{P}$ <br> $\left(\mathrm{~m} \mathrm{~s}^{-1}\right)$ | $V_{S}$ <br> $\left(\mathrm{~m} \mathrm{~s}^{-1}\right)$ |
| :--- | :---: | :---: | :---: |
| Crust | 2700 | 5800 | 3200 |
| Mantle | 3500 | 8045 | 4500 |

located at ( $60 \mathrm{~km}, 60 \mathrm{~km}, 0 \mathrm{~km}$ ). The computed results are shown in Figs. 3-5.

Figure 3 shows snapshots at four instants. When $t=10 \mathrm{~s}$, the plane wave is still outside the domain (Fig. 3a-c). When $t=16 \mathrm{~s}$, the incident plane wave enters the lower layer of the model (Fig. 3d-f). When $t=18 \mathrm{~s}$, the transmitted and reflected P waves are very clear in the plane at $y=50 \mathrm{~km}$ (Fig. 3h). Figure 3i clearly illustrates that the wavelength in the upper mantle is greater than that in the crust because the velocity of the upper mantle is approximately 1.39 times the velocity of the crust. When $t=22 \mathrm{~s}$, the transmitted P wave is reflected by the free surface, and the Moho-reflected waves are propagating outside the model (Fig. 3k).

To qualitatively evaluate the accuracy of EBE-SEM for teleseismic wave modeling, the synthetic seismology at the station is compared with the reference solution, which is generated by the FK technique. The results are shown in Fig. 4. As depicted in Fig. 4a, the synthetic and reference waveforms show excellent agreement. The direct wave with the largest amplitude is followed by the converted S wave and the crust multiples with relatively small amplitudes. Although the amplitudes of the crust multiples are small, their phases, ampli-
tudes, and travel times are correctly modeled. The error curve is shown in the right panel of Fig. 4. The maximum relative difference between the numerical solution and the reference solution is less than $5 \%$, although the numerical method uses only fourth-order interpolation in the space and nearly one spectral-element sampling for the minimal wavelength.

The computation was performed on a workstation with 2 Intel Xeon CPUs (E5-2680 v3) and 128 GB of RAM, and 24 CPU cores were used. The master-slave communication pattern is used in our parallel algorithm, which is similar to that of Komatitsch and Tromp (2002). The parallel efficiency is shown in Fig. 5. Because the master-slave communication pattern requires at least two CPU cores to perform the parallel algorithm, the computation time of single CPU core is not shown in Fig. 5. When 2 CPU cores are used, the amount of communication between the CPU cores is negligible compared to the amount of computation. If we denote the CPU time when 2 CPU cores are used in parallel computation as $T_{2}$, the CPU time of $n$ CPU cores can be estimated by $T_{n}=\frac{2}{n} T_{2}$. The red line in Fig. 5 represents the estimated CPU times (theoretical times), and the blue line with circles is the practical CPU times. It can be clearly observed that the actual CPU times are extremely close to the theoretical times, even if the number of CPU cores is 24 . The red arrows in Fig. 5 indicate the abnormal CPU times. This phenomenon is attributed to the excessive communication amount when the numbers of CPU cores are 18 and 19. However the anomaly is not large compared with the neighboring CPU times.

### 5.2 Plane-wave incidence to a 3-D model

To demonstrate that EBE-SEM works well in 3-D heterogeneous models, an abnormal structure with cube shape of an additional $15 \%$ plus wave speeds and density are added at the center of the model in Fig. 1. The size of the abnormal structure is $20 \mathrm{~km} \times 20 \mathrm{~km} \times 20 \mathrm{~km}$. Except for velocity structure, the other computational parameters are the same as those of the first numerical example. The simulation results are shown in Fig. 6.
The black arrows in the upper plots of Fig. 6 indicate the distortion of the wave front because of the velocity anomaly in the media. Because of the positive velocity anomaly, the distorted wave front travels faster than the undistorted plane wave. The lower plots in Fig. 6 show strong scattered waves. As the yellow arrows indicates, the strong scattered waves


Figure 6. Snapshots in a heterogeneous model. The snapshots in the upper and lower panels are taken at $t=18$ and 22 s , respectively. The left, middle, and right panels are snapshots of planes at $x=50 \mathrm{~km}, y=50 \mathrm{~km}$, and $z=30 \mathrm{~km}$, respectively. The black arrows in the upper plots indicate the distortion of the wave front because of the velocity anomaly in the media. The yellow arrows in (a) and (b) indicate the scattered waves. The yellow arrows in (d) and (e) denote the scattered waves that are efficiently absorbed by the constructed PML ABC. The color scale is shown on the right side of the figure.
do not reflect back to the interior of the model because of the efficiency of the PML ABC used in this paper.

### 5.3 Teleseismic waveform misfit gradient

One key advantage of the EBE-SEM is its convenience for constructing the misfit gradient because the element stiffness matrix can easily be assembled based on the tensor product of the submatrices. To illustrate this advantage, we first define the misfit function:
$E(\boldsymbol{m})=\frac{1}{2} \sum_{s} \sum_{r} \int_{0}^{T}\left[\boldsymbol{d}\left(t, \boldsymbol{x}_{r} ; \boldsymbol{x}_{\mathrm{s}}\right)-\boldsymbol{s}\left(t, \boldsymbol{x}_{r} ; \boldsymbol{x}_{\mathrm{s}}\right)\right]^{2} \mathrm{~d} t$,
where $\boldsymbol{m}=(\ln \rho, \ln \lambda, \ln \mu)$ represents model parameters, $\boldsymbol{x}_{r}$ is the station location, $\boldsymbol{x}_{\mathrm{s}}$ is the teleseismic incident parameters, $\boldsymbol{d}$ is the observed dataset, and $\boldsymbol{s}$ is the synthetic dataset. The cross section between the plane wave and the Moho at $y=0$, the incident angle, and the azimuth angle constitute the teleseismic incident parameters. To simplify the discussion, this example only considers one source and one station. Based on the continuous adjoint method discussed by Fichtner (2011), the adjoint equation is

$$
\begin{align*}
\rho \ddot{\boldsymbol{w}}_{\mathrm{t}} & -\nabla \cdot\left[\lambda\left(\nabla \cdot \boldsymbol{w}_{\mathrm{t}}\right) \mathbf{I}+\mu\left(\nabla \boldsymbol{w}_{\mathrm{t}}+\nabla \boldsymbol{w}_{\mathrm{t}}^{T}\right)\right] \\
& =\left(\boldsymbol{d}\left(T-t, \boldsymbol{x}_{r} ; \boldsymbol{x}_{\mathrm{s}}\right)-\boldsymbol{s}\left(T-t, \boldsymbol{x}_{r} ; \boldsymbol{x}_{\mathrm{s}}\right)\right) \delta\left(\boldsymbol{x}-\boldsymbol{x}_{r}\right), \tag{43}
\end{align*}
$$

where $\boldsymbol{w}_{\mathrm{t}}$ is the adjoint total wave field. The misfit gradient $\nabla_{\mathrm{m}} E$ obeys the following equations (Monteiller et al., 2015):
$\frac{\partial E}{\partial(\ln \rho)_{e}}=-\sum_{k=0}^{N T}\left[\boldsymbol{W}_{\mathrm{t}}^{e}(T-k \Delta t) \mathbf{M}^{e} \ddot{\boldsymbol{U}}_{\mathrm{t}}^{e}(k \Delta t) \Delta t\right]$,
$\frac{\partial E}{\partial(\ln \lambda)_{e}}=-\sum_{k=0}^{N T}\left[\lambda_{e} \boldsymbol{W}_{\mathrm{t}}^{e}(T-k \Delta t) \mathbf{K}_{\lambda}^{e} \boldsymbol{U}_{\mathrm{t}}^{e}(k \Delta t) \Delta t\right]$,

$$
\begin{equation*}
\frac{\partial E}{\partial(\ln \mu)_{e}}=-\sum_{k=0}^{N T}\left[2 \mu_{e} \boldsymbol{W}_{\mathrm{t}}^{e}(T-k \Delta t) \mathbf{K}_{\mu}^{e} \boldsymbol{U}_{\mathrm{t}}^{e}(k \Delta t) \Delta t\right] \tag{46}
\end{equation*}
$$

where $e$ represents the $e$ th element; $\boldsymbol{W}$ is the discrete vector of $\boldsymbol{w} ; \mathbf{K}_{\lambda}^{e}$ and $\mathbf{K}_{\mu}^{e}$ are the element stiffness matrices for the misfit gradient calculation, which are presented in Appendix C ; and $N T$ is the total time step. $\mathbf{K}_{\lambda}^{e}$ and $\mathbf{K}_{\mu}^{e}$ can also be written as a combination of the tensor product of the submatrices, which leads to an easy construction of the misfit gradient. The misfit gradients can be written as the gradients with respect to $\ln \rho, \ln V_{P}$, and $\ln V_{S}$ by simple combinations of Eqs. (44)-(46).

We consider the model in the second numerical example to be the real model and the 1-D layered model in Fig. 1 as the initial model. The observed and synthetic waveforms at the station (red triangle in Fig. 1) are shown in Fig. 7a. As Fig. 7b shows, the time-reversed waveform differences between the observed data and the theoretical data act as the source term in Eq. (43).

Figure 8 shows the constructed misfit gradients in the plane at $y=61 \mathrm{~km}$ at two time slices. Because of the singularity of the source, large amplitudes are distributed in the vicinity of the adjoint source. The misfit gradient no longer has a banana-doughnut shape (Tromp et al., 2005) but rather is similar to the Greek letter $\Lambda$ (Fig. 8e, f) because of the plane-wave source. The misfit gradient with respect to $\ln \rho$ is much weaker than those for $\ln V_{P}$ and $\ln V_{S}$. The extremely strong negative misfit gradients of $\ln V_{P}$ distribute along the ray path of the direct $P$ wave.

## 6 Discussion and conclusions

Teleseismic wave adjoint tomography has the ability to image the deep structure of the lithosphere. Thus, a highly ef-


Figure 7. Construction of the adjoint source. In the left panels, the synthetic waveforms from the second example are considered the observed data (blue line), and the synthetic waveforms from the first example are treated as the theoretical data (red line). The rectangular boxes indicate the time window [20,24] to isolate the waveforms to construct the adjoint source. The top, middle, and bottom panels correspond to the $x, y$, and $z$ components of the displacement. The three components of the adjoint source-time function is the time-reversed waveform difference between the observed data and the theoretical data.


Figure 8. Misfit gradients in the plane at $y=61 \mathrm{~km}$ at two times. The top and bottom panels correspond to $t=18$ and 24 s , respectively. The left, middle, and right panels correspond to the misfit gradients in terms of $\rho, V_{P}$, and $V_{S}$, respectively. The color scale is shown on the right side of the figure.
ficient method for teleseismic wave forward-modeling and misfit calculation is important. In this work, the EBE-SEM was specially tailored to teleseismic wave modeling and misfit gradient calculation. In this approach, the PML ABC is discretized by EBE-SEM, and the method can efficiently absorb scattered teleseismic waves. Teleseismic wave incident conditions are constructed for the computational and PML domains. An economic technique for boundary wave field storage is introduced that can greatly reduce the required amount of computer memory.

The numerical results from the first and second numerical examples demonstrate not only the efficiency of EBESEM in modeling teleseismic waves but also the validity of the constructed teleseismic wave incident boundary condition. As shown in the third numerical example, hardly any
extra effort is required to construct the misfit gradient. The EBE-SEM has advantages over the traditional SEM in three respects: the reduction in the required computer memory requirement, easy calculation of the misfit gradient, and significant parallelization efficiency.

Code and data availability. The original source codes for the numerical examples are written in C. To obtain these source codes, please contact the first author (email: slliu@math.tsinghua.edu.cn) or the corresponding author by email. The MPI library was downloaded from http://www.mpich.org/downloads.

Appendix A: Element matrices for the computational domain

The element matrices for the computational domain are listed below:

$$
\begin{align*}
& \mathbf{M}^{e}=\left[\begin{array}{lll}
\mathbf{M}_{1}^{e} & & \\
& \mathbf{M}_{1}^{e} & \\
& & \mathbf{M}_{1}^{e}
\end{array}\right],  \tag{A1}\\
& \mathbf{M}_{1}^{e}=\left[\begin{array}{ccc}
\delta_{i j} \int_{\Omega_{e}} \rho \phi_{i} \phi_{j} \mathrm{~d} \Omega \\
\Omega_{e}
\end{array}\right],  \tag{A2}\\
& \mathbf{K}^{e}=\left[\begin{array}{lll}
\mathbf{K}_{11}^{e} & \mathbf{K}_{12}^{e} & \mathbf{K}_{13}^{e} \\
\mathbf{K}_{21}^{e} & \mathbf{K}_{22}^{e} & \mathbf{K}_{23}^{e} \\
\mathbf{K}_{31}^{e} & \mathbf{K}_{32}^{e} & \mathbf{K}_{33}^{e}
\end{array}\right], \tag{A3}
\end{align*}
$$

$\mathbf{K}_{11}^{e}=\left[\int_{\Omega}(\lambda+2 \mu) \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial x}+\mu\left(\frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y}+\frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial z}\right) \mathrm{d} \Omega\right]$,
$\mathbf{K}_{12}^{e}=\left[\int_{\Omega} \lambda \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial y}+\mu \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right]$,
$\mathbf{K}_{13}^{e}=\left[\int_{\Omega} \lambda \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial z}+\mu \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right]$,
$\mathbf{K}_{21}^{e}=\left[\int_{\Omega} \mu \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial y}+\lambda \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right]$,
$\mathbf{K}_{22}^{e}=\left[\int_{\Omega}(\lambda+2 \mu) \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y}+\mu\left(\frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial x}+\frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y}\right) \mathrm{d} \Omega\right]$,
$\mathbf{K}_{23}^{e}=\left[\int_{\Omega} \lambda \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial z}+\mu \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial y} \mathrm{~d} \Omega\right]$,
$\mathbf{K}_{33}^{e}=\left[\int_{\Omega}(\lambda+2 \mu) \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial z}+\mu\left(\frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial x}+\frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y}\right) \mathrm{d} \Omega\right]$.

## Appendix B: Element matrices for the PML domain

The element matrices for the PML domain are listed below:

$$
\begin{align*}
& \mathbf{K}_{x x}^{e}=\left[\int_{\Omega_{e}}(\lambda+2 \mu) \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right],  \tag{B1}\\
& \mathbf{K}_{y y}^{e}=\left[\int_{\Omega_{e}} \mu \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y} \mathrm{~d} \Omega\right],  \tag{B2}\\
& \mathbf{K}_{z z}^{e}=\left[\int_{\Omega_{e}} \mu \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial z} \mathrm{~d} \Omega\right] \text {, }  \tag{B3}\\
& \mathbf{K}_{x y}^{e}=\left[\int_{\Omega_{e}} \lambda \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial y} \mathrm{~d} \Omega\right] \text {, }  \tag{B4}\\
& \mathbf{K}_{y x}^{e}=\left[\int_{\Omega_{e}} \mu \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right] \text {, }  \tag{B5}\\
& \mathbf{K}_{x z}^{e}=\left[\int_{\Omega_{e}} \lambda \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial z} \mathrm{~d} \Omega\right],  \tag{B6}\\
& \mathbf{K}_{z x}^{e}=\left[\int_{\Omega_{e}} \mu \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right] \text {, }  \tag{B7}\\
& \mathbf{K}_{x}^{e}=\left[\int_{\Omega_{e}}(\lambda+2 \mu) d_{x}^{\prime} \phi_{i} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right] \text {, }  \tag{B8}\\
& \mathbf{K}_{y}^{e}=\int_{\Omega_{e}} \mu d_{y}^{\prime} \phi_{i} \frac{\partial \phi_{j}}{\partial y} \mathrm{~d} \Omega,  \tag{B9}\\
& \mathbf{K}_{z}^{e}=\left[\int_{\Omega_{e}} \mu d_{z}^{\prime} \phi_{i} \frac{\partial \phi_{j}}{\partial z} \mathrm{~d} \Omega\right],  \tag{B10}\\
& \mathbf{D}_{x}^{e}=\left[2 \delta_{i j} \int_{\Omega_{e}} d_{x}^{e} \rho \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]  \tag{B11}\\
& \mathbf{D}_{y}^{e}=\left[2 \delta_{i j} \int_{\Omega_{e}} d_{y}^{e} \rho \phi_{i} \phi_{j} \mathrm{~d} \Omega\right],  \tag{B12}\\
& \mathbf{D}_{z}^{e}=\left[2 \delta_{i j} \int_{\Omega_{e}} d_{z}^{e} \rho \phi_{i} \phi_{j} \mathrm{~d} \Omega\right], \tag{B13}
\end{align*}
$$

$\mathbf{D}_{x x}^{e}=\left[\delta_{i j} \int_{\Omega_{e}} \rho\left(d_{x}^{e}\right)^{2} \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]$,
$\mathbf{D}_{y y}^{e}=\left[\delta_{i j} \int_{\Omega_{e}} \rho\left(d_{y}^{e}\right)^{2} \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]$,
$\mathbf{D}_{z z}^{e}=\left[\delta_{i j} \int_{\Omega_{e}} \rho\left(d_{z}^{e}\right)^{2} \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]$,
$\mathbf{D}_{x y}^{e}=\left[\delta_{i j} \int_{\Omega_{e}} \rho d_{x}^{e} d_{y}^{e} \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]$,
$\mathbf{D}_{x z}^{e}=\left[\delta_{i j} \int_{\Omega_{e}} \rho d_{x}^{e} d_{z}^{e} \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]$,
$\mathbf{D}_{x, y}^{e}=\left[\delta_{i j} \int_{\Omega_{e}} \rho\left(d_{x}^{e}+d_{y}^{e}\right) \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]$,
$\mathbf{D}_{x, z}^{e}=\left[\delta_{i j} \int_{\Omega_{e}} \rho\left(d_{x}^{e}+d_{z}^{e}\right) \phi_{i} \phi_{j} \mathrm{~d} \Omega\right]$.
where $d^{e}$ represents the average value of the damping coefficient on the element $e$.

## Appendix C: Element matrices for the misfit gradient

 In Eqs. (44) and (45), $\mathbf{K}_{\lambda}^{e}$ and $\mathbf{K}_{\mu}^{e}$ are given by$$
\left.\begin{array}{l}
\mathbf{K}_{\lambda}^{e}=\left[\begin{array}{lll}
\mathbf{K}_{\lambda 11}^{e} & \mathbf{K}_{\lambda 12}^{e} & \mathbf{K}_{\lambda 13}^{e} \\
\mathbf{K}_{\lambda 21}^{e} & \mathbf{K}_{\lambda 22}^{e} & \mathbf{K}_{\lambda 23}^{e} \\
\mathbf{K}_{\lambda 31}^{e} & \mathbf{K}_{\lambda 32}^{e} & \mathbf{K}_{\lambda 33}^{e}
\end{array}\right], \\
\mathbf{K}_{\mu}^{e}=\left[\begin{array}{lll}
\mathbf{K}_{\mu_{11}}^{e} & 0 & 0 \\
0 & \mathbf{K}_{\mu 11}^{e} & 0 \\
0 & 0 & \mathbf{K}_{\mu 11}^{e}
\end{array}\right], \\
\mathbf{K}_{\lambda 11}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right.
\end{array}\right],
$$

$$
\begin{equation*}
\mathbf{K}_{\lambda 13}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial z} \mathrm{~d} \Omega\right], \tag{C5}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{K}_{\lambda 21}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right], \tag{C6}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{K}_{\lambda 22}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y} \mathrm{~d} \Omega\right], \tag{C7}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{K}_{\lambda 23}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial z} \mathrm{~d} \Omega\right], \tag{C8}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{K}_{\lambda 31}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial x} \mathrm{~d} \Omega\right], \tag{C9}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{K}_{\lambda 32}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial y} \mathrm{~d} \Omega\right], \tag{C10}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{K}_{\lambda 33}^{e}=\left[\int_{\Omega_{e}} \frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial z} \mathrm{~d} \Omega\right], \tag{C11}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{K}_{\mu_{11}}^{e}=\int_{\Omega_{e}}\left(\frac{\partial \phi_{i}}{\partial x} \frac{\partial \phi_{j}}{\partial x}+\frac{\partial \phi_{i}}{\partial y} \frac{\partial \phi_{j}}{\partial y}+\frac{\partial \phi_{i}}{\partial z} \frac{\partial \phi_{j}}{\partial z}\right) \mathrm{d} \Omega . \tag{C12}
\end{equation*}
$$

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We greatly appreciate the detailed suggestions from Michal Afanasiev and the anonymous reviewer. Their valuable suggestions greatly improved the quality of the paper. This study was supported by the National Natural Science Foundation of China (grant nos. 41230210 and 41604034). Shaolin Liu was financially supported by the China Postdoctoral Science Foundation (grant no. 2015M580085).

## Edited by: Michal Malinowski

Reviewed by: Michael Afanasiev and one anonymous referee

## References

Capdeville, Y., Chaljub, E., Vilotte, J., and Montagner, J.: Coupling the spectral element method with a modal solution for elastic wave propagation in global Earth models, Geophys. J. Int., 152, 34-67, 2003.
Chen, M., Niu, F., Liu, Q., Tromp, J., and Zheng, X.: Multiparameter adjoint tomography of the crust and upper mantle beneath East Asia: 1. Model construction and comparisons, J. Geophys. Res., 120, 22-39, 2015.
Clayton, R. and Engquist, B.: Absorbing boundary condition for acoustic and elastic wave equation, B. Seismol. Soc. Am., 67, 1529-1540, 1977.
Cohen, G.: High-order numerical methods for transient wave equations, Scientific Computation, Springer-Verlag, Berlin Heidelberg, 2002.
Collino, F. and Tsogka, C.: Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media, Geophysics, 66, 294-307, 2001.

Dablain, M.: The application of high-order differencing to the scalar wave equation, Geophysics, 51, 54-66, 1986.
De Basabe, J. and Sen, M.: Grid dispersion and stability criteria of some common finite-element methods for acoustic and elastic wave equations, Geophysics, 72, T81-T95, 2007.
Fichtner, A.: Full seismic waveform modelling and inversion, Springer-Verlag Berlin Heidelberg, 2011.
Greathouse, J. and Daga, M.: Efficient sparse matrix-vector multiplication on GPUs using the CSR storage format, International Conference for High Performance Computing, Netwroking, Storage and Analysis, 769-780, 2014.
Haskell, N.: The dispersion of surface waves on multilayered media, B. Seismol. Soc. Am., 43, 17-34, 1953.

Kind, K., Yuan, X., and Kumar, P.: Seismic receiver functions and the lithosphere-astheosphere boundary, Technophysics, 536537, 25-43, 2012.
Komatitsch, D. and Tromp, J.: Spectral-element simulations of global seismic wave propagation - I. Validation, Geophys. J. Int., 149, 390-412, 2002.
Komatitsch, D. and Tromp, J.: A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation, Geophys. J. Int., 154, 146-153, 2003.

Komatitsch, D. and Vilotte, J.: The spectral element method: an efficient tool to simulation the seismic response of 2D and 3D geological structures, B. Seismol. Soc. Am., 88, 368-392, 1998.
Langston, C.: Corvallis, Oregon, crustal and upper mantel receiver structure from teleseismic P and S waves, B. Seismol. Soc. Am., 67, 713-724, 1977.
Liu, Q. and Gu, Y.: Seismic imaging: From classical to adjoint tomography, Tectonophysics, 566-567, 31-66, 2012.
Liu, Q. and Tromp, J.: Finite-frequency kernels based on adjoint method, B. Seismol. Soc. Am., 96, 2383-2397, 2006.
Liu, S., Li, X., Wang, W., and Liu, Y.: A mixed-grid finite element method with PML absorbing boundary conditions for seismic wave modelling, J. Geophys. Eng., 11, 055009, https://doi.org/10.1088/1742-2132/11/5/055009, 2014.
Liu, S., Li, X., Wang, W., and Zhu, T.: Source wavefield reconstruction using a linear combination of the boundary wavefield in reverse time migration, Geophysics, 80, S203-S212, 2015.
Liu, S., Yang, D., Lang, C., Wang, W., and Pan, Z.: Modified symplectic schemes with nearly-analytic discrete operators for acoustic wave simulations, Comp. Phys. Comm., 213, 52-63, 2017a.
Liu, S., Yang, D., and Ma, J.: A modified symplectic PRK scheme for seismic wave modeling, Comp. Geosci., 99, 28-36, 2017 b.
Monteiller, V., Chevrot, S., Komatitsch, D., and Fuji, N.: A hybrid method to compute short-period synthetic seismograms of teleseismic body waves in a 3-D regional model, Geophys. J. Int., 192, 230-247, 2013.
Monteiller, V., Chevrot, S., Komatitsch, D., and Wang, Y.: Threedimensional full waveform inversion of short-period teleseismic wavefields based upon the SEM-DSM hybrid method, Geophys. J. Int., 202, 811-827, 2015.

Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casrotti, E., Le Loher, P., Magnoni, F., Liu, Q., Blitz, C., Nissen-Meyer, T., Basini, P., and Tromp, J.: Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, Geophys. J. Int., 186, 721-739, 2011.
Roecker, S., Baker, D., and McLaughlin, J.: A finite-difference algorithm for full waveform teleseismic tomography, Geophys. J. Int., 181, 1017-1040, 2010.
Rondenay, S.: Upper mantle imaging with array recordings of converted and scattered teleseismic waves, Surv. Geophys., 30, 377405, 2009.
Seriani, G.: A parallel spectral element method for acoustic wave modeling, J. Comp. Acoust., 5, 53-69, 1997.
Seriani, G.: 3-D large-scale wave propagation modelling by spectral element method on Cray T3E multiprocessor, Comput. Method. Appl. M., 164, 235-247, 1998.
Shragge, J., Artman, B., and Wilson, C.: Teleseismic shot-profile migration, Geophysics, 71, S1221-S1229, 2006.
Tape, C., Liu, Q., and Tromp, J.: Finite-frequency tomography using adjoint methods-Methodology and examples using membrance surface waves, Geophys. J. Int., 168, 1105-1129, 2007.
Tape, C., Liu, Q., Maggi, A., and Tromp, J.: Adjoint tomography of the southern California crust, Science, 325, 988-992, 2009.
Tong, P., Chen, C., Komatitsch D., Basini, P., and Liu, Q.: Highresolution seismic array imaging based on an SEM-FK hybrid method, Geophys. J. Int., 197, 369-395, 2014a.
Tong, P., Zhao, D., Yang, D., Yang, X., Chen, J., and Liu, Q.: Wave-equation-based travel-time seismic tomography - Part 2:

Application to the 1992 Landers earthquake ( $M_{\mathrm{w}} 7.3$ ) area, Solid Earth, 5, 1169-1188, https://doi.org/10.5194/se-5-11692014, 2014b.
Tong, P., Zhao, D., Yang, D., Yang, X., Chen, J., and Liu, Q.: Wave-equation-based travel-time seismic tomography - Part 1: Method, Solid Earth, 5, 1151-1168, https://doi.org/10.5194/se-5-1151-2014, 2014c.
Tong, P., Komatitsch, D., Tseng, T., Hung, S., Chen, C., Basini, P., and Liu, Q.: A 3-D spectral-element and frequency-wave number hybrid method for high-resolution seismic array imaging, Geophys. Res. Lett., 41, 7025-7034, 2015.
Tromp, J., Tape, C., and Liu, Q.: Seismic tomography, adjoint method, time reversal, and banana-doughnut kernels, Geophys. J. Int., 160, 195-216, 2005.

Virieux, J. and Operto, S.: An overview of full-waveform inversion in exploration geophysics, Geophysics, 74, WCC1-WCC26, 2009.

Yang, D., Wang, S., Zhang, Z., and Teng, J.: n-times absorbing boundary conditions for compact finite-difference modeling of acoustic ad elastic wave propagation in 2D TI medium, B. Seismol. Soc. Am., 93, 2389-2401, 2003.
Zhang, F., Li, Y., Wu, Q., and Ding, Z.: The P wave velocity structure of upper mantle beneath the North China and surrounding regions from FMTT, Chin. J. Geophys., 54, 1233-1242, 2011 (in Chinese).
Zhu, L. and Rivera, A.: A note on the dynamic and static displacement from a point source in multilayered media, Geophys. J. Int., 148, 619-627, 2002.

