Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

Solid Earth Discussions

Element-by-element parallel spectral-element methods for 3-D acoustic-wave-equation-based teleseismic wave modeling

Shaolin Liu[a], Dinghui Yang[a, *], Xingpeng Dong[a], Qiancheng Liu[b], and Yongchang Zheng[c]

a Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, China

b King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

c Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100730, China

Correspondence to: Dinghui Yang (dhyang@math.tsinghua.edu.cn).

Abstract

The increasing demand for the high-resolution imaging of deep lithosphere structures requires the utilization of a teleseismic dataset for waveform inversion. The construction of an efficient algorithm for the teleseismic wavefield modeling is valuable for the calculation of misfit kernels or Fréchet derivatives when the teleseismic waveform is used for adjoint tomography. Here, we introduce an element-by-element parallel spectral-element method (EBE-SEM) for the efficient modeling of teleseismic wavefield propagation in a localized geology model. Under the assumption of the plane wave, the frequency-wavenumber (FK) technique is implemented to compute the boundary wavefield used for constructing the boundary condition of the teleseismic wave incidence. To reduce the memory required for the storage of the boundary wavefield for the incidence boundary condition, an economical strategy is introduced to store the boundary wavefield on the model boundary. The perfectly matched layers absorbing boundary condition (PML ABC) is formulated by the EBE-SEM to

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

Solid Earth Discussions

absorb the scattered wavefield from the model interior. The misfit kernel (derivatives of the waveform misfit with respect to model parameters) can be easily constructed without extra computational effort for the calculation of the element stiffness matrix per time step during the calculation of the adjoint wavefield. Three synthetic examples demonstrate the validity of EBE-SEM for use in teleseismic wavefield modeling and the misfit kernel calculation.

1 Introduction

Teleseismic waves provide tremendous amount of information for the detection of crust and upper mantle structures (Rondenay, 2009). In the past fourth years, many techniques have been established 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). For the requirements of high-resolution teleseismic wave imaging, 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 technique of adjoint tomography constructs the Fréchet derivatives of the objective function with respect to model parameters by two times of numerical solving full seismic wave equation (Tromp et al., 2005; Liu and Tromp, 2006). Adjoint tomography has been successfully implemented to investigate crust (Tape et al., 2009) and continental subsurface heterogeneity (Chen et al., 2015). Indeed, the seismic wave-equation-based adjoint waveform tomography, which has a greater resolution than the ray-based traveltime tomography for the same dense seismic ray coverage (Liu and Gu, 2012), is able to image the small-scale (half of the minimum wavelength) heterogeneity (Virieux and Operato, 2009). The main drawback of the adjoint tomography is its huge computational burden. The computational requirement is linearly related to the earthquake events used for tomography inversion and the iterations required by the optimization technique. For a typical local scale model, such as southern California region, to perform adjoint tomography inversion, several thousand of 3-D full-wavefield simulations are required (Tape et al., 2007, 2009).

Because most earthquakes occur in the crust and seismic rays cannot easily illuminate the deep lithosphere in local region seismic tomography, it can be difficult to image the deep lithosphere structures (Tong et al., 2014b). Increasing the model size enables more deep reflections and refractions to be included in the inversion dataset; as a result, deep structures can be inversed 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 s for P waves and 3-6 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 the 2-D/3-D numerical solvers. The boundary conditions for the localized simulation model are provided by fast-computing 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 localized 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 computation efforts can be efficiently reduced by these hybrid methods, the

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

computational costs are still excessive for a small workstation when we are faced with the several thousand of 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-D hybrid method for teleseismic wave simulations. To simplify the teleseismic wavefield invariant along a particular axis, the 2.5-D formulation can significantly reduce the computational demands. However, the 2.5-D formulation may restrict the application of the method in an arbitrarily heterogeneous media (Tong et al., 2015).

In addition to the great computational demand (CPU time) associated with the 3-D hybrid methods, satisfying the memory requirements for storing the boundary wavefields 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 (1977) to interface the 1-D background analytical solution with a numerical solver on the boundary of a localized model. This treatment can not only assign the teleseismic incident condition for the computational domain, but also absorb the scatter wavefield from the interior of the heterogeneous model. The implementation of the CE-type boundary condition is extremely simple and dese not substantially increase the required CPU time. However, the CE-type boundary condition can efficiently absorb the waves only at approximately normal incident angles, and the incident 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 wavefields in teleseismic adjoint tomography and thus decrease the accuracy of the constructed Fréchet derivatives. Note that the CE-type boundary condition requires 9 wavefield components (3 displacement components and 6 stress components) to be stored on the model boundary; this requirement may be a great burden to the computer memory for the

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

case of a relatively large scale model decomposed by a dense numerical mesh.

Here, we introduce EBE-SEM for the efficient modeling of teleseismic plane-wave propagation in localized models. One great advantage of EBE-SEM is the easy parallelization of the spectral-element algorithm, which does not require assembly of the global stiffness matrix. The spectral elements are equally allocated to every CPU core, and the product of the stiffness matrix with the solution vector is calculated element by element; these aspects are quite useful to ensure load balance among the CPU cores. In addition, the misfit kernel can be efficiently constructed, because each element stiffness matrix is stored in CPU memory and it is simultaneously available when calculating the misfit kernel. The perfectly matched layers (Collino and Tsogka, 2001; Komatitsch and Tromp, 2003; Liu et al., 2014) are formulated by EBE-SEM to effectively absorb scattered waves. A detailed discussion is presented to incorporate the teleseismic incident boundary condition for EBE-SEM, and only two layers of the boundary wavefield must be stored in computer memory. The high efficiency of EBE-SEM for teleseismic wave modeling and misfit kernel construction is demonstrated by three numerical examples.

2 EBE-SEM

A schematic of a teleseismic plane wave entering a localized model is depicted in Figure 1, where the localized mode is delineated by the blue lines. We first introduce EBE-SEM for acoustic wave propagation in an infinite half space, which includes the localized model. We denote the total wavefield u_t, which is the summation of the background wavefield u_{FK} in layered media and the scattered wavefield u_s. u_{FK} can be efficiently calculated with the wavenumber-frequency domain method (FK) (Haskell, 1953; Zhu and Rivera, 2002; Tong et al., 2014a). u_s is excited by the incident plane wave because of the heterogeneity of the media.

2.1 EBE-SEM for total wave simulation

We assume that the total wavefield (P or S wave) obeys the following acoustic wave equation (Tong et al., 2014c):

$$\ddot{u}_t - \nabla \cdot \left[c^2 \nabla u_t \right] = 0, \tag{1}$$

where the two dots above u_t denote the second-order time derivative, c is either the P or S wave velocity model, and $\nabla = \left(\mathbf{e}_x \partial_x + \mathbf{e}_y \partial_y + \mathbf{e}_z \partial_z \right)$ is the gradient operator. Following the classical spectral-element method (SEM) (Seriani, 1997; Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 2002), using an arbitrary test function v to multiply both sides of Eq. (1) and integrating over the domain Ω, we obtain the following weak form wave equation:

$$\int_\Omega v \ddot{u}_t d\Omega + \int_\Omega \nabla v \cdot c^2 \nabla u_t d\Omega = \int_{\partial\Omega} v \mathbf{n} \cdot c^2 \nabla u_t ds, \tag{2}$$

where $\mathbf{n}$ is the normal vector of the boundary $\partial\Omega$. Under the natural boundary condition, the right side of Eq. (2) is zero. If the infinite space is decomposed into N non-overlapping hexahedral elements, then Eq. (2) can be written as:

$$\sum_{e=1}^{N} \int_{\Omega_e} v \ddot{u}_t d\Omega + \sum_{e=1}^{N} \int_{\Omega_e} \nabla v \cdot c^2 \nabla u_t d\Omega = 0. \tag{3}$$

Each hexahedral element is mapped to the master cube $[-1, 1]^3$, and we choose the interpolation points to be consistent with the Gauss-Legendre-Lobatto (GLL) quadrature points (Cohen, 2002). The interpolation basis functions in each element is as follows:

$$\phi_i(x, y, z) = \varphi_{i_1}(\xi) \varphi_{i_2}(\gamma) \varphi_{i_3}(\eta), \tag{4}$$

where the subscript of ϕ denotes the i-th basis function; ξ, γ, and η are the three coordinates in the isoparametric coordinate system; and φ is the Lagrange polynomial. The i-th interpolation point in the physical element is mapped to the (i_1, i_2, i_3)-th node in the

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

cube $[-1, 1]^3$. Based on Eq. (4) and the discrete values $u_t^{e,i}$ and $v^{e,i}$ on the interpolation points, the continuous values u_t^e and v^e in element e can be approximated by

$$\begin{cases} u_t^e \approx \sum_{i=1}^{(n+1)^3} u_t^{e,i} \phi_i(x, y, z), \\ v^e \approx \sum_{i=1}^{(n+1)^3} v^{e,i} \phi_i(x, y, z), \end{cases} \tag{5}$$

where n is the polynomial order of the interpolation basis. Substituting Eq. (5) into Eq. (2), and using the GLL quadrature rule, we obtain the ordinary differential equation in terms of time:

$$\sum_{e=1}^{N} \mathbf{M}^e \ddot{\mathbf{U}}_t^e + \sum_{e=1}^{N} \mathbf{K}^e \mathbf{U}_t^e = 0, \tag{6}$$

where $\mathbf{M}^e$ is the element mass matrix, $\mathbf{K}^e$ is the element stiffness matrix, and $\mathbf{U}_t^e$ 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; this is one great advantageous of Legendre or polynomial-based SEM because explicit inverse of the mass matrix can be easily obtained. The assembled global diagonal mass is as follows:

$$\mathbf{M} = \sum_{e=1}^{N} \mathbf{M}^e. \tag{7}$$

We define the projection operator T^e to map the corresponding elements of $\mathbf{M}$ back to form an element matrix:

$$\tilde{\mathbf{M}}^e = T^e(\mathbf{M}), \tag{8}$$

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 corresponding to the common points shared by the adjacent elements are

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

summed together to form the global mass matrix in Eq. (7), some elements of $\tilde{\mathbf{M}}^e$ are greater than those of $\mathbf{M}^e$. If we denote the global solution vector as $\mathbf{U}_t$, then the projected back element solution $T^e(\mathbf{U}_t)$ is still $\mathbf{U}_t^e$ because u_t is continuous cross elements. From Eq. (6), we have

$$\ddot{\mathbf{U}}_t = -\sum_{e=1}^{N}\left(T^e(\mathbf{M})^{-1}\mathbf{K}^e T^e(\mathbf{U}_t)\right). \tag{9}$$

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 performed 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 that the great computational balance among CPU cores when the parallel algorithm is performed on a workstation. Our parallel code is based on Message Passing Interface (MPI) library (www.mpich.org/downloads), and we equally distribute spectral-elements to each of the CPU cores. Because $\mathbf{K}^e$ is symmetric, only the upper triangle of the $\mathbf{K}^e$ must be stored in computer memory.

2.2 EBE-SEM for PML formula

In our previous work, a second-order PML absorbing boundary condition (PML ABC) are formulated by the mixed-grid finite-element method (Liu et al., 2014). Here, we use this type of PML ABC to absorb the scattered wavefield u_s. The formula of the PML ABC can be written as

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

$$\begin{cases} \ddot{u}_{s,1} + 2d_x\dot{u}_{s,1} + d_x{}^2u_{s,1} = \dfrac{\partial}{\partial x}\left(c^2\dfrac{\partial u_s}{\partial x}\right) + P_x, \\ \dot{P}_x + d_xP_x = -c^2d'_x\dfrac{\partial u_s}{\partial x}, \\ \ddot{u}_{s,2} + 2d_y\dot{u}_{s,2} + d_y{}^2u_{s,2} = \dfrac{\partial}{\partial y}\left(c^2\dfrac{\partial u_s}{\partial y}\right) + P_y, \\ \dot{P}_y + d_yP_y = -c^2d'_y\dfrac{\partial u_s}{\partial y}, \\ \ddot{u}_{s,3} + 2d_z\dot{u}_{s,3} + d_z{}^2u_{s,3} = \dfrac{\partial}{\partial z}\left(c^2\dfrac{\partial u_s}{\partial z}\right) + P_z, \\ \dot{P}_z + d_zP_z = -c^2d'_z\dfrac{\partial u_s}{\partial z}, \end{cases} \tag{10}$$

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; $u_{s,1}$, $u_{s,2}$, and $u_{s,3}$ are the three split components of u_s; and P_x, P_y, and P_z are the three intermediate variables. Based on the natural boundary condition, the corresponding wake form of Eq. (10) is

$$\begin{cases} \int_\Omega v\left(\ddot{u}_{s,1} + 2d_x\dot{u}_{s,1} + d_x{}^2u_{s,1}\right)d\Omega + \int_\Omega \dfrac{\partial v}{\partial x}c^2\dfrac{\partial u_s}{\partial x}d\Omega = \int_\Omega vP_x d\Omega, \\ \int_\Omega v\left(\dot{P}_x + d_xP_x\right)d\Omega + \int_\Omega vc^2d'_x\dfrac{\partial u_s}{\partial x}d\Omega = 0, \\ \int_\Omega v\left(\ddot{u}_{s,2} + 2d_y\dot{u}_{s,2} + d_y{}^2u_{s,2}\right)d\Omega + \int_\Omega \dfrac{\partial v}{\partial y}c^2\dfrac{\partial u_s}{\partial y}d\Omega = \int_\Omega vP_y d\Omega, \\ \int_\Omega v\left(\dot{P}_y + d_yP_y\right)d\Omega + \int_\Omega vc^2d'_y\dfrac{\partial u_s}{\partial y}d\Omega = 0, \\ \int_\Omega v\left(\ddot{u}_{s,3} + 2d_z\dot{u}_{s,3} + d_z{}^2u_{s,3}\right)d\Omega + \int_\Omega \dfrac{\partial v}{\partial z}c^2\dfrac{\partial u_s}{\partial z}d\Omega = \int_\Omega vP_z d\Omega, \\ \int_\Omega v\left(\dot{P}_z + d_zP_z\right)d\Omega + \int_\Omega vc^2d'_z\dfrac{\partial u_s}{\partial z}d\Omega = 0. \end{cases} \tag{11}$$

The discretized version of Eq. (11) is

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

$$\begin{cases} \sum_{e=1}^{N}\mathbf{M}^e\ddot{\mathbf{U}}_{s,1}^e + \sum_{e=1}^{N}\mathbf{D}_x^e\dot{\mathbf{U}}_{s,1}^e + \sum_{e=1}^{N}\mathbf{D}_{xx}^e\mathbf{U}_{s,1}^e + \sum_{e=1}^{N}\mathbf{K}_{xx}^e\mathbf{U}_{s,1}^e = \sum_{e=1}^{N}\mathbf{M}^e\mathbf{P}_x^e, \\ \sum_{e=1}^{N}\mathbf{M}^e\dot{\mathbf{P}}_x^e + \frac{1}{2}\sum_{e=1}^{N}\mathbf{D}_x^e\mathbf{P}_x^e + \sum_{e=1}^{N}\mathbf{K}_x^e\mathbf{P}_x^e = 0, \\ \sum_{e=1}^{N}\mathbf{M}^e\ddot{\mathbf{U}}_{s,2}^e + \sum_{e=1}^{N}\mathbf{D}_y^e\dot{\mathbf{U}}_{s,2}^e + \sum_{e=1}^{N}\mathbf{D}_{yy}^e\dot{\mathbf{U}}_{s,2}^e + \sum_{e=1}^{N}\mathbf{K}_{yy}^e\mathbf{U}_{s,2}^e = \sum_{e=1}^{N}\mathbf{M}^e\mathbf{P}_y^e, \\ \sum_{e=1}^{N}\mathbf{M}^e\dot{\mathbf{P}}_y^e + \frac{1}{2}\sum_{e=1}^{N}\mathbf{D}_y^e\mathbf{P}_y^e + \sum_{e=1}^{N}\mathbf{K}_y^e\mathbf{P}_y^e = 0, \\ \sum_{e=1}^{N}\mathbf{M}^e\ddot{\mathbf{U}}_{s,3}^e + \sum_{e=1}^{N}\mathbf{D}_z^e\dot{\mathbf{U}}_{s,3}^e + \sum_{e=1}^{N}\mathbf{D}_{zz}^e\dot{\mathbf{U}}_{s,3}^e + \sum_{e=1}^{N}\mathbf{K}_{zz}^e\mathbf{U}_{3}^e = \sum_{e=1}^{N}\mathbf{M}^e\mathbf{P}_z^e, \\ \sum_{e=1}^{N}\mathbf{M}^e\dot{\mathbf{P}}_z^e + \frac{1}{2}\sum_{e=1}^{N}\mathbf{D}_z^e\mathbf{P}_z^e + \sum_{e=1}^{N}\mathbf{K}_z^e\mathbf{P}_z^e = 0, \\ \mathbf{U}_s = \mathbf{U}_{s,1} + \mathbf{U}_{s,1} + \mathbf{U}_{s,1}, \end{cases} \tag{12}$$

where $\mathbf{K}_x^e$, $\mathbf{K}_y^e$, $\mathbf{K}_z^e$, $\mathbf{K}_{xx}^e$, $\mathbf{K}_{yy}^e$; $\mathbf{K}_{zz}^e$ are element stiffness matrices; and $\mathbf{D}_x^e$, $\mathbf{D}_{xx}^e$, $\mathbf{D}_y^e$, $\mathbf{D}_{yy}^e$, $\mathbf{D}_z^e$, and $\mathbf{D}_{zz}^e$ are the element damping matrices. The detailed expressions of these element matrices are presented in Appendix A. 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 damping matrices as $\mathbf{D}_x$, $\mathbf{D}_y$, and $\mathbf{D}_z$, then we obtain the element-by-element scheme for Eq. (12):

$$\begin{cases} \ddot{\mathbf{U}}_{s,1} + \sum_{e=1}^{N}T^e\left(\mathbf{M}_P\right)^{-1}T^e\left(\mathbf{D}_x\right)\dot{\mathbf{U}}_{s,1}^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}_P\right)^{-1}T^e\left(\mathbf{D}_{xx}\right)\mathbf{U}_{s,1}^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}\mathbf{K}_{xx}^e\mathbf{U}_{s,1}^e = \mathbf{P}_x, \\ \dot{\mathbf{P}}_x + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}T^e\left(\mathbf{D}_x\right)\mathbf{P}_x^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}\mathbf{K}_x^e\mathbf{P}_x^e = 0, \\ \ddot{\mathbf{U}}_{s,2} + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}T^e\left(\mathbf{D}_y\right)\dot{\mathbf{U}}_{s,2}^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}T^e\left(\mathbf{D}_{yy}\right)\mathbf{U}_{s,2}^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}\mathbf{K}_{yy}^e\mathbf{U}_{s,2}^e = \mathbf{P}_y, \\ \dot{\mathbf{P}}_y + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}T^e\left(\mathbf{D}_y\right)\mathbf{P}_y^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}\mathbf{K}_y^e\mathbf{P}_y^e = 0, \\ \ddot{\mathbf{U}}_{s,3} + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}T^e\left(\mathbf{D}_z\right)\dot{\mathbf{U}}_{s,3}^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}T^e\left(\mathbf{D}_{zz}\right)\mathbf{U}_{s,3}^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}\mathbf{K}_{zz}^e\mathbf{U}_{s,3}^e = \mathbf{P}_z, \\ \dot{\mathbf{P}}_z + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}T^e\left(\mathbf{D}_z\right)\mathbf{P}_z^e + \sum_{e=1}^{N}T^e\left(\mathbf{M}\right)^{-1}\mathbf{K}_z^e\mathbf{P}_z^e = 0, \\ \mathbf{U}_s = \mathbf{U}_{s,1} + \mathbf{U}_{s,1} + \mathbf{U}_{s,1}. \end{cases} \tag{13}$$

From Eq. (A9) to (A14), Eq. (13) can be rewritten as

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

$$\begin{cases}
\ddot{\mathbf{U}}_{s,1} + \sum_{e=1}^{N} 2d_x^e \dot{\mathbf{U}}_{s,1}^e + \sum_{e=1}^{N} \left(d_x^e\right)^2 \mathbf{U}_{s,1}^e + \sum_{e=1}^{N} T^e \left(\mathbf{M}\right)^{-1} \mathbf{K}_{xx}^e \mathbf{U}_{s,1}^e = \mathbf{P}_x, \\
\dot{\mathbf{P}}_x + \sum_{e=1}^{N} d_x^e \mathbf{P}_x^e + \sum_{e=1}^{N} T^e \left(\mathbf{M}\right)^{-1} \mathbf{K}_x^e T\left(\mathbf{P}_x\right) = 0, \\
\ddot{\mathbf{U}}_{s,2} + \sum_{e=1}^{N} 2d_y^e \dot{\mathbf{U}}_{s,2}^e + \sum_{e=1}^{N} \left(d_y^e\right)^2 \mathbf{U}_{s,2}^e + \sum_{e=1}^{N} T^e \left(\mathbf{M}\right)^{-1} \mathbf{K}_{yy}^e \mathbf{U}_{s,2}^e = \mathbf{P}_y, \\
\dot{\mathbf{P}}_y + \sum_{e=1}^{N} d_y^e \mathbf{P}_y^e + \sum_{e=1}^{N} T^e \left(\mathbf{M}\right)^{-1} \mathbf{K}_y^e T\left(\mathbf{P}_y\right) = 0, \\
\ddot{\mathbf{U}}_{s,3} + \sum_{e=1}^{N} 2d_z^e \dot{\mathbf{U}}_{s,3}^e + \sum_{e=1}^{N} \left(d_z^e\right)^2 \mathbf{U}_{s,3}^e + \sum_{e=1}^{N} T^e \left(\mathbf{M}\right)^{-1} \mathbf{K}_{zz}^e \mathbf{U}_{s,3}^e = \mathbf{P}_z, \\
\dot{\mathbf{P}}_z + \sum_{e=1}^{N} d_z^e \mathbf{P}_z^e + \sum_{e=1}^{N} T^e \left(\mathbf{M}\right)^{-1} \mathbf{K}_z^e T\left(\mathbf{P}_z\right) = 0, \\
\mathbf{U}_s = \mathbf{U}_{s,1} + \mathbf{U}_{s,2} + \mathbf{U}_{s,3}.
\end{cases} \tag{14}$$

From Eq. (14), we observe that element damping matrices do not need to be stored, and only the mass matrix and the element stiffness matrix must be kept in computer memory. Note that because $\mathbf{K}_x^e$, $\mathbf{K}_y^e$, and $\mathbf{K}_z^e$ are not symmetric, and all the elements of each element matrix should be stored.

2.3 A discussion of EBE-SEM

Although EBE-SEM is specially designed for teleseismic wave modeling, i.e., Eq. (9) is for teleseismic total wavefield propagation if the proper teleseismic incident boundary condition is added and Eq. (14) is for absorbing the scatted wavefield, EBE-SEM can be directly used for wavefield simulation for an earthquake that occurred interior of the model (computation domain) if a source term is added in Eq. (9).

The seminal work for introducing EBE-SEM was that of Seriani (1997). In the 2-D case, the element-by-element scheme is combined with the Chebyshev orthogonal polynomial-based SEM. Because the Chebyshev orthogonal polynomial is orthogonal 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

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

large spare linear system of equations, which may be not efficient for larger-scale seismic waveform modeling. Another aspect of the BEB algorithm designed by Seriani (1997) is that this algorithm is for use with only rectangular elements; the restriction may be problematic for curved elements. In contrast, our EBE-SEM can be used for general cases.

If appropriate boundary conditions are added for the computation domain and the PML domain, then the computation will be isolated in the counterpart domain, i.e., Eq. (9) for the computation domain and Eq. (14) for the PML domain. The discussion of the boundary condition is presented below.

3 Teleseismic wave incident boundary conditions

For the completeness of the paper, we first simply introduce the FK method for determining the 3-D acoustic-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 storage of the boundary wavefields.

3.1 Plane wave propagation in 1-D layered media

Transforming the acoustic wave equation into the $\left(\omega,\, k_x,\, k_y\right)$ domain, we obtain the following equation:

$$-\omega^2 u_{FK} = \left(ik_x\mathbf{e}_x + ik_y\mathbf{e}_y + \partial_z\right)c^2\left(ik_x\mathbf{e}_x + ik_y\mathbf{e}_y + \partial_z\right)u_{FK}, \tag{15}$$

where u_{FK} is the plane wave in layered media; k_x and k_y are wavenumbers in the x and y directions, respectively; and ω is the angular frequency. We assign u_{FK} the same notation in both time-space and frequency-wavenumber domains to avoid clustering. If we define $k\mathbf{e}_H = k_x\mathbf{e}_x + k_y\mathbf{e}_y$, where $k = \sqrt{k_x^2 + k_y^2}$ and $\mathbf{e}_H$ is the horizontal normal vector, then Eq. (15) can be written as

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

$$-\omega^2 u_{FK} = \left(ik\mathbf{e}_H + \partial_z\right)\left(c^2 ik\mathbf{e}_H + c^2\partial_z\right)u_{FK}. \tag{16}$$

Eq. (16) is changed to the following ordinary differential equations:

$$\frac{\partial}{\partial z}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{c^2} \\ c^2k^2 - \omega^2 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{Fy}, \tag{17}$$

where $y_1 = u_{FK}$ and $y_2 = c^2 \frac{\partial u_{FK}}{\partial z}$. By calculating the eigenvalue and eigenvector of $\mathbf{F}$, the general solution of Eq. (17) can be written as

$$\mathbf{y} = \begin{bmatrix} 1 & 1 \\ ic^2\upsilon & -ic^2\upsilon \end{bmatrix}\begin{bmatrix} e^{i\upsilon z} & \\ & e^{-i\upsilon z} \end{bmatrix}\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \tag{18}$$

where C_1 and C_2 correspondto the amplitudes of the down-going and up-going acoustic waves, respectively, and $\upsilon = \sqrt{\frac{\omega^2}{c^2} - k^2}$ is the ray parameter. As shown in Figure 2, the velocity and thickness of the m-th layer are denoted as c_m and h_m, respectively. We define the following matrices for the *m*-th layer:

$$\mathbf{R}_m = \begin{bmatrix} 1 & 1 \\ ic_m^2\upsilon_m & -ic_m^2\upsilon_m \end{bmatrix}, \tag{19}$$

$$\mathbf{H}\left(h_m\right) = \begin{bmatrix} e^{i\upsilon_m h_m} & \\ & e^{-i\upsilon_m h_m} \end{bmatrix}. \tag{20}$$

The relationship between the wavefields at $z = z_{m-1}$ and $z = z_m$ can be written as

$$\mathbf{y}_{m-1} = \mathbf{R}_m\mathbf{H}\left(-h_m\right)\mathbf{R}_m^{-1}\mathbf{y}_m. \tag{21}$$

In Eq. (21), $\mathbf{L}_m = \mathbf{R}_m\mathbf{H}\left(-h_m\right)\mathbf{R}_m^{-1}$ is the propagation matrix. The wavefield at the free surface can be represented by

$$\mathbf{y}_0 = \mathbf{L}_1\cdots\mathbf{L}_n\mathbf{R}_n\mathbf{C}. \tag{22}$$

If we consider that the incident plane wave has only an up-going component and has unit amplitude, then we have

$$C_1 = \pm 1. \tag{23}$$

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

We define

$$\mathbf{A} = \mathbf{L}_1 \cdots \mathbf{L}_n \mathbf{R}_n = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \tag{24}$$

y_2 at the free surface can be written as

$$y_2 = a_{21}C_1 + a_{22}C_2. \tag{24}$$

Based on the free surface boundary condition $y_2 = 0$, we have

$$C_2 = -\frac{a_{21}}{a_{22}} C_1. \tag{25}$$

Substituting Eq. (25) and (23) into Eq. (18), we obtain the wavefield at $z = z_m$. Based on Eq. (21), we can calculate the wavefield in each layer.

3.2 Incident boundary conditions

We assume the infinite space is composed of the hexahedron set

$$B = \left\{ e_1, \ldots, e_{N_1}, e_{N_1+1}, \ldots, e_{N_1+N_2}, \ldots \right\}. \tag{26}$$

The first N_1 elements of B compose the computation domain (the domain bounded by the blue lines in Figure 1); 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

$$B_C = \left\{ e_1, \ldots, e_{N_1} \right\}. \tag{27}$$

A total of N_3 elements in B_C that contact the boundary of PML domain are collected in the set:

$$B_{C,B} = \left\{ e_{j_1}, \ldots, e_{j_{N_2}} \right\}. \tag{28}$$

The set of N_2 elements that compose the PML domain is given by

$$B_P = \left\{ e_{N_1+1}, \ldots, e_{N_1+N_2} \right\}. \tag{29}$$

A total of N_4 elements in B_P that contact the boundary of computation domain are gathered in the set

$$B_{P,B} = \left\{ e_{k_1}, \ldots, e_{k_{N_4}} \right\}. \tag{30}$$

We define an operator that maps the elements of vector A to a target vector A_t according to the numbering set A_n. The index number of each elements of set A_t in set A is gathered in A_n. We have

$$\langle A \rangle_{A_n} = A_t. \tag{31}$$

From Eq. (9), we have

$$\begin{aligned} \langle \ddot{\mathbf{U}}_t \rangle_{A_C} &= \left\langle -\sum_{e \in B} \left(T^e (\mathbf{M})^{-1} \mathbf{K}^e T^e (\mathbf{U}_t) \right) \right\rangle_{A_C} \\ &= \left\langle -\sum_{e \in B_C \cup B_{P,B}} \left(T^e (\mathbf{M})^{-1} \mathbf{K}^e T^e (\mathbf{U}_t) \right) \right\rangle_{A_C}. \end{aligned} \tag{32}$$

where A_C is the numbering set of the nodes in the computational domain. Eq. (32) shows only matrix-vector products in the elements of the computational domain and the elements of the PML domain in contact with boundary of the computational domain contribute to the acceleration wavefield on the nodes in the computational domain. From Eq. (32), we further have

$$\begin{aligned} \langle \ddot{\mathbf{U}}_t \rangle_{A_{C,B}} &= \left\langle -\sum_{e \in B_C \cup B_{P,B}} \left(T^e (\mathbf{M})^{-1} \mathbf{K}^e T^e (\mathbf{U}_t) \right) \right\rangle_{A_{C,B}} \\ &= \left\langle -\sum_{e \in B_{C,B} \cup B_{P,B}} \left(T^e (\mathbf{M})^{-1} \mathbf{K}^e T^e (\mathbf{U}_t) \right) \right\rangle_{A_{C,B}}, \\ &= \left\langle -\sum_{e \in B_{C,B}} \left(T^e (\mathbf{M})^{-1} \mathbf{K}^e T^e (\mathbf{U}_t) \right) \right\rangle_{A_{C,B}} + \left\langle -\sum_{e \in B_{P,B}} \left(T^e (\mathbf{M})^{-1} \mathbf{K}^e T^e (\mathbf{U}_t) \right) \right\rangle_{A_{C,B}}, \end{aligned} \tag{33}$$

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. (33) can be written as

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

$$\left\langle -\sum_{e\in B_{P,B}}\left(T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}^{e}T^{e}\left(\mathbf{U}_{t}\right)\right)\right\rangle_{A_{C,B}}=\left\langle -\sum_{e\in B_{P,B}}\left(T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}^{e}T^{e}\left(\mathbf{U}_{FK}+\mathbf{U}_{s}\right)\right)\right\rangle_{A_{C,B}}$$
$$=\left\langle -\sum_{e\in B_{P,B}}\left(T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}^{e}T^{e}\left(\mathbf{U}_{FK}\right)\right)\right\rangle_{A_{C,B}}+\left\langle -\sum_{e\in B_{P,B}}\left(T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}^{e}T^{e}\left(\mathbf{U}_{s}\right)\right)\right\rangle_{A_{C,B}}. \quad (34)$$

Eq. (33) is the plane-wave incident condition of the computational domain. From Eq. (34), we need to store only the boundary wavefield $\left\langle -\sum_{e\in B_{P,B}}\left(T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}^{e}T^{e}\left(\mathbf{U}_{FK}\right)\right)\right\rangle_{A_{C,B}}$ to construct the incident boundary condition. The length of $\left\langle -\sum_{e\in B_{P,B}}\left(T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}^{e}T^{e}\left(\mathbf{U}_{FK}\right)\right)\right\rangle_{A_{C,B}}$ is the same as the number of nodes located on the interface between the computational and PML domains; and it is not necessary to store $\mathbf{U}_{FK}$ 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 linear combination of a boundary wavefield, which was proposed in our previous work (Liu et al., 2015).

To simplify the discussion, the boundary condition of the first equation of Eq. (14) is discussed. For Eq. (14), we have

$$\left\langle \ddot{\mathbf{U}}_{s,1}\right\rangle_{A_P}+\left\langle \sum_{e\in B}2d_{x}^{e}\dot{\mathbf{U}}_{s,1}^{e}\right\rangle_{A_P}+\left\langle \sum_{e\in B}\left(d_{x}^{e}\right)^{2}\mathbf{U}_{s,1}^{e}\right\rangle_{A_P}+\left\langle \sum_{e\in B}T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}_{P,xx}^{e}\mathbf{U}_{s}^{e}\right\rangle_{A_P}=\left\langle \mathbf{P}_{x}\right\rangle_{A_P}, \quad (35)$$

where A_P is the numbering set of nodes in PML domain. The scattered wave on the interface of the computational domain and the PML domain can be obtained by the following equation:

$$\left\langle \mathbf{U}_{s}\right\rangle_{A_{P,B}}=\left\langle \mathbf{U}_{t}\right\rangle_{A_{P,B}}-\left\langle \mathbf{U}_{FK}\right\rangle_{A_{P,B}}, \quad (36)$$

where $A_{P,B}$ is the numbering set of the nodes of the PML domain on the interface. $\left\langle \mathbf{U}_{s}\right\rangle_{A_{P,B}}$ does not need to be calculated by Eq. (14). From Eqs. (35) and (36), we have

$$\left\langle \ddot{\mathbf{U}}_{s,1}\right\rangle_{A_P-A_{P,B}}+\left\langle \sum_{e\in B_P}\left(2d_{x}^{e}\dot{\mathbf{U}}_{s,1}^{e}+\left(d_{x}^{e}\right)^{2}\mathbf{U}_{s,1}^{e}\right)\right\rangle_{A_P-A_{P,B}}+\left\langle \sum_{e\in B_P}T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}_{P,xx}^{e}\mathbf{U}_{s}^{e}\right\rangle_{A_P-A_{P,B}}=\left\langle \mathbf{P}_{x}\right\rangle_{A_P-A_{P,B}}.$$

(37)

In addition to the boundary condition Eq. (36), the Dirichlet boundary condition is added on the outer boundaries of the PML domain, and the natural boundary condition is added on the planes that connecting the free surface of the computation 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 boundary conditions Eq. (33) and (36) involve the plane wavefields, and we call Eqs. (33) and (36) teleseismic wave incident conditions.

4 Analysis of the computational costs

We use the model in Figure 1 to quantitatively discuss the computational cost of EBE-SEM for teleseismic wave modeling. Because the thickness of the PML domain is only three elements wide in our numerical examples, the number of float point operatorations is trivial compared with the computational domain, and the main computational cost of the PML domain is mainly from the storage requirement of the boundary condition (Eq. (36)).

The model is decomposed into a total of 75000 cubic elements with a size of $2\ km \times 2\ km \times 2\ km$. The fifth-order interpolation polynomial is used in the space. The time interval is 0.01 s, and the total time length for numerical modeling is 120 s.

The float point operation in the computation domain is mainly from the element matrix-vector product, and a total of 6.9822×10^9 float point operations are required in each time step. If a global stiffness matrix is assembled, then the product of the global stiffness matrix with the global solution vector requires a total of 6.435220951×10^9 float point operations. The computational burden of EBE-SEM is larger than that of the conventional SEM because the acceleration on common nodes shared by the adjacent elements is calculated

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

more than once. However, the number of operations EBE-SEM is increased by only 8.5%. This increased compuational amount may be compensated for the great 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 13.64 GB is required to store the boundary wavefield to construct the teleseismic incident conditions. If the classical compressed sparse row (CSR) storage format (Greathouse and Daga, 2014) is used for the storage of spare stiffness matrix of the conventional SEM, then the storage requirement is 24.04 GB, which is nearly four times as large as the storage requirement of EBE-SEM to store the element stiffness matrices. 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 non-zero elements. Second, the element stiffness matrix of EBE-SEM is symmetric, i.e., only the upper triangle of element stiffness matrix must be stored.

5 Numerical examples

Three numerical examples are provided to validate the efficiency of EBE-SEM for teleseismic wave modeling. In all the examples, the Gaussian source-time function with a cut-off frequency of 1 Hz is used (Tong et al., 2014a).

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

Except for parameters presented in Section 4, the upper velocity and the lower layers are 3000 m/s and 4500 m/s, respectively; the incidence angle is 15^o, and the azimuth angle is 170^o. The cross section of the plane-wave front with Moho is (-300 km, 0, 30 km) at the initial time t=0. A seismic station is located at (70 km, 50 km, 0 km). The computed results are shown in Figures 3-5.

Figure 3 shows snapshots at three time instants. When t=10 s, the plane waves are still outside of the domain (Figure 3(a-c)). When t=30 s, the incident plane wave is reflected by and transmitted through the Moho. Figure 3(d) clearly illustrates that the wave length in the upper mantle is larger than that in the crust, because the velocity of the upper mantle is 1.5 times as large as the velocity of the crust. When t=35, the transmitted wave is reflected by the free surface (Figure 3(h)) and the Moho reflected wave is propagating outside of the model.

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 FK. The result is shown in Figure 4. As depicted in Figure 4(a), excellent agreement is achieved between the synthetic and reference waveforms. The direct wave with largest amplitude is followed by crust multiples with relatively small amplitude. Although the amplitude of the crust multiples is small, the phases, amplitudes, and travel times of these multiples are correctly modeled. The error curve is shown in Figure 4(b). The maximum difference between the numerical solution and the reference solution is approximately 7×10^{-5}, which is extremely small compared with the amplitudes of the direct wave and crust multiples.

The computation was performed on a workstation with 2 Intel Xeon CPUs (E5-2680 v3) and 128 GB of RAM, and the total number of CPU cores is 24. 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 Figure 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 Figure 5. When the number of CPU cores is 2, the communication amount between the CPU cores is negligible

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

compared with the computation amount. If we denote the CPU time is as T_2 when 2 CPU cores are used in parallel computation, the CPU time of n CPU cores can be estimated by $T_n = \frac{2}{n} T_2$. The red line in Figure 5 is 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 arrow in Figure 5 indicates the abnormal CPU time. This phenomenon is attributed to the excessive communication amount when the number of CPU cores is 9. 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 speed is added at the center of model in Figure 1. The size of the abnormal structure is $20\ km \times 20\ km \times 20\ km$. Except for velocity structure, the other computational parameters are the same as those of the first numerical example. The simulation results are shown in Figure 6.

The red arrows in the upper plots of Figure 6 indicate the distortion of the wave front because of the velocity anomaly of the media. Because of the plus velocity anomaly, the distorted wave front travels faster than the undistorted wave. From the lower plots of Figure 6, we can observe strong scatted waves. As the yellow arrows indicates, the strong scattered waves do not reflected into the interior of the model because of the efficiency of the PML ABC constructed in the paper.

5.3 Teleseismic waveform misfit kernel

One key advantage of the EBE-SEM is the great convenience of constructing the misfit kernel, because the element stiffness matrix is kept in computer memory and no extra effort is require to recalculate the element stiffness matrix. To illustrate this advantage, we first define the misfit function:

$$E(m)=\frac{1}{2}\sum_{s}\sum_{r}\frac{\int_0^T\left[d\left(t,\mathbf{x}_r;\mathbf{x}_s\right)-s\left(t,\mathbf{x}_r;\mathbf{x}_s\right)\right]^2 dt}{\int_0^T\left|d\left(t,\mathbf{x}_r;\mathbf{x}_s\right)\right|^2 dt}, \tag{38}$$

where $\mathbf{x}_r$ is the station location, $\mathbf{x}_s$ is the teleseismic incident parameters, d is the observed dataset, and 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, only one source and one station are considered in this example. Based on the continuous adjoint method discussed by Fichtner (2011), the adjoint equation is

$$\ddot{w}_t-\nabla\cdot\left[c^2\nabla w_t\right]=\frac{d\left(T-t,\mathbf{x}_r;\mathbf{x}_s\right)-s\left(T-t,\mathbf{x}_r;\mathbf{x}_s\right)}{\int_0^T\left|d\left(t,\mathbf{x}_r;\mathbf{x}_s\right)\right|^2 dt}, \tag{39}$$

where w_t is the adjoint total wavefield. The misfit kernel $\nabla_m E$ obeys the following equation:

$$\nabla_m E\delta\mathbf{m}=\int_0^T dt\int_{\Omega_C}\left[-2c\nabla w_t\left(T-t\right)\cdot\nabla u_t\right]\delta\mathbf{m}d\Omega. \tag{40}$$

where Ω_C denotes the computational domain. The *i*-th element of the misfit kernel can be calculated by the following equation:

$$\nabla_{m_i}E=\sum_{k=1}^{NT}\left[-2c_i\mathbf{W}_t^i\left(T-k\Delta t\right)\mathbf{K}^i\mathbf{U}_t^i\left(k\Delta t\right)\right]\Delta t, \tag{41}$$

where NT is the total time step. Eq. (41) shows that element stiffness matrices are used for the construction of the misfit kernel.

We consider the model in the second numerical example to be the real model, and the 1-D layered model in Figure 1 is the initial domain. The observed and synthetic waveforms at the station (red triangle in Figure 1) are shown in Figure 7(a). As Figure 7(b) shows, the time-reversed and $1/\int_0^T \left|d\left(t,\mathbf{x}_r;\mathbf{x}_s\right)\right|^2 dt$ scaled waveform difference between the observed data and the theoretical data acts as the source term in Eq. (39).

The constructed misfit kernels in the y=50 km plane at four time slices are shown in Figure 8. Because of the singularity of the source, large amplitudes are distributed in the vicinity of the adjoint source. The misfit kernel is no longer banana-doughnut shaped (Tromp et al., 2005) but, rather, similar to the Greek letter Λ (Figure 8(b)) because of the plane-wave source.

6 Discussion and conclusions

Teleseismic wave adjoint tomography has the ability to image the deep structure of the lithosphere. Thus, a highly efficient method for teleseismic wave forward modeling and misfit calculation is important. In this work, the EBE-SEM was specially tailored for teleseismic wave modeling and misfit calculation. In this approach, the PML ABC is discretized by EBE-SEM, and the method can efficiently absorb scatted teleseismic waves. Teleseismic wave incident conditions are constructed for computation and PML domains. An economic technique for boundary wavefield 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 EBE-SEM in modeling teleseismic waves but also the validity of the constructed teleseismic wave incident boundary condition. As shown in the third numerical example, no extra effort is required to construct the misfit kernel. The EBE-SEM has advantages over the traditional SEM in three respects: first, the reduction of the computer memory requirement; second, the easy calculation of the misfit kernel; and third, the great efficiency of parallelization.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

Acknowledgments

The MPI library was downloaded from http://www.mpich.org/downloads. This study was supported by the National Natural Science Foundation of China (grant nos. 41230210 and 41604034). S. Liu was financially supported by the China Postdoctoral Science Foundation (grant no. 2015M580085).

Appendix A: Elements of element matrices

The elements for the element matrices in the computational and PML domains are presented below:

$$\mathbf{M}_{ii}^{e} = \int_{\Omega_e} \phi_i \phi_i d\Omega, \tag{A1}$$

$$\mathbf{K}_{ij}^{e} = \int_{\Omega_e} c^2 \nabla\phi_i \cdot \nabla\phi_j d\Omega, \tag{A2}$$

$$\mathbf{K}_{x,ij}^{e} = \int_{\Omega_e} \phi_i d_x'^e \frac{\partial \phi_j}{\partial x} d\Omega, \tag{A3}$$

$$\mathbf{K}_{y,ij}^{e} = \int_{\Omega_e} \phi_i d_y'^e \frac{\partial \phi_j}{\partial y} d\Omega, \tag{A4}$$

$$\mathbf{K}_{z,ij}^{e} = \int_{\Omega_e} \phi_i d_z'^e \frac{\partial \phi_j}{\partial z} d\Omega, \tag{A5}$$

$$\mathbf{K}_{xx,ij}^{e} = \int_{\Omega_e} c^2 \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} d\Omega, \tag{A6}$$

$$\mathbf{K}_{yy,ij}^{e} = \int_{\Omega_e} c^2 \frac{\partial \phi_i}{\partial y} \frac{\partial \phi_j}{\partial y} d\Omega, \tag{A7}$$

$$\mathbf{K}_{zz,ij}^{e} = \int_{\Omega_e} c^2 \frac{\partial \phi_i}{\partial z} \frac{\partial \phi_j}{\partial z} d\Omega, \tag{A8}$$

$$\mathbf{D}_{x,ii}^{e} = 2\int_{\Omega_e} d_x^e \phi_i \phi_i d\Omega, \tag{A9}$$

$$\mathbf{D}_{y,ii}^{e} = 2\int_{\Omega_e} d_y^e \phi_i \phi_i d\Omega, \tag{A10}$$

$$\mathbf{D}_{z,ii}^{e} = 2\int_{\Omega_e} d_z^e \phi_i \phi_i d\Omega, \tag{A11}$$

$$\mathbf{D}_{xx,ii}^{e} = \int_{\Omega_e} \left(d_x^e\right)^2 \phi_i \phi_i d\Omega, \tag{A12}$$

$$\mathbf{D}^e_{yy,ii} = \int_{\Omega_e} \left(d^e_y\right)^2 \phi_i \phi_i d\Omega, \tag{A13}$$

$$\mathbf{D}^e_{zz,ii} = \int_{\Omega_e} \left(d^e_z\right)^2 \phi_i \phi_i d\Omega, \tag{A14}$$

where d^e represents the average value of the damping coefficient on the element *e*.

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.

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, Bull. Seism. Soc. Am., 67, 1529-1540, 1977.

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.

Cohen, G.: High-order numerical methods for transient wave equations, Scientific Computation: Springer-Verlag, 2002.

Dablain, M.: The application of high-order differencing to the scalar wave equation, Geophysics, 51, 54-66, 1986.

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. Bull. Seism. Soc. Am., 43, 17-34, 1953.

Kind, K., Yuan, X., and Kumar, P.: Seismic receiver functions and the lithosphere-astheosphere boundary, Technophysics, 536-537, 25-43, 2012.

Komatitsch, D., and Vilotte, J.: The spectral element method: an efficient tool to simulation the seismic response of 2D and 3D geological structures, Bull. Seism. Soc. Am., 88, 368-392, 1998.

Komatitsch, D., and Tromp, J.: Spectral-element simulations of global seismic wave propagation-Ⅰ. Validation, 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.

Langston, C.: Corvallis, Oregon, crustal and upper mantel receiver structure from teleseismic P and S waves, B. Seism. Soc. Am., 67, 713-724, 1977.

Liu, Q., and Liu Y.: Seismic imaging: From classical to adjoint tomography, Tectonophysics, 566-567, 31-66, 2012.

Liu, Q., and Tromp, J.: Finite-frequency kernels based upon adjoint method, Bull. Seism. 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, 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, 2017b.

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.: Three-dimensional full waveform inversion of short-period teleseismic wavefields based upon the SEM-DSM hybrid method, Geophys. J. Int., 202, 811-827, 2015.

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, 377-405, 2009.

Shragge, J., Artman, B., and Wilson, C.: Teleseismic shot-profile migration, Geophysics, 71, S1221-S1229, 2006.

Seriani, G.: A parallel spectral element method for acoustic wave modeling, J. Comp. Acoust., 5, 53-69, 1997.

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.: Hihg-resoultion 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_w 7.3) area, Solid Earth, 5, 1169-1188, 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, 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., Operto, S.: An overview of full-waveform inversion in exploration geophysics, Geophysics, 74, WCC1-WCC26, 2009.

Yang, D., Wang, S., Zhang, Z., Teng, J.; n-times absorbing boundary conditions for compact finite-difference modeling of acoustic ad elastic wave propagation in 2D TI medium, Bull. Seism. Soc. Am., 93, 2389-2401, 2002.

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. (in Chinese), 54, 1233-1242, 2011.

Zhu, L. and Rivera, A.: A note on the dynamic and static displacement from a point source in

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

multilayered media, Geophys. J. Int., 148, 619-627, 2002.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

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 giga byte.

Method	Stiffness matrix	$\left\langle -\sum_{e\in B_{P,B}} \left(T^{e}\left(\mathbf{M}\right)^{-1}\mathbf{K}^{e}T^{e}\left(\mathbf{U}_{t}\right)\right)\right\rangle_{A_{C,B}}$	$\left\langle \mathbf{U}_{FK}\right\rangle_{A_{P,B}}$
EBE-SEM	6.55 GB	6.82 GB	6.82 GB
general SEM	24.04 GB	6.82 GB	6.82 GB

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

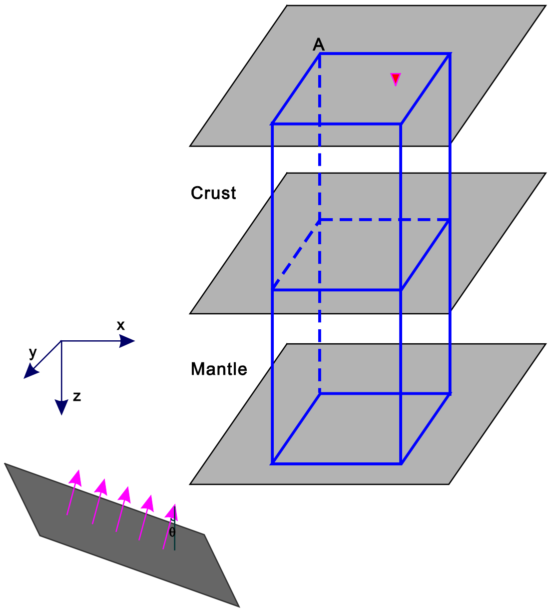

Figure 1. Schematic of a teleseismic plane wave entering a localized model, which is a two-layered crust-upper mantle model. The area of study is delineated by the blue lines. The origin of the Cartesian coordinate is located on the surface, in accordance with the corner A of the model. The positive x and y directions are pointing the east and south, respectively, and the positive z is pointing downward. The inverted red triangle located at (70 km, 50 km, 0 km) represents a seismic station. The plane in the lower left of the figure denotes wave front of the incident teleseismic plane wave. The purple arrows indicates the normal vector of the plane wave. The incident angle is θ, which is the angle between the normal vector and –z. The azimuth angle is φ, which is the angle between the projection of normal vector into x-y plane and –y. The azimuth angle is not explicitly shown in this figure.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

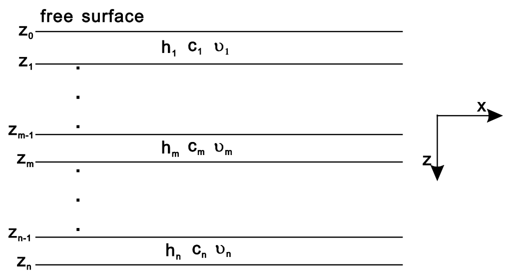

Figure 2. 1-D layered model. The thickness, velocity, and ray parameter of each layer are shown in the plot.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

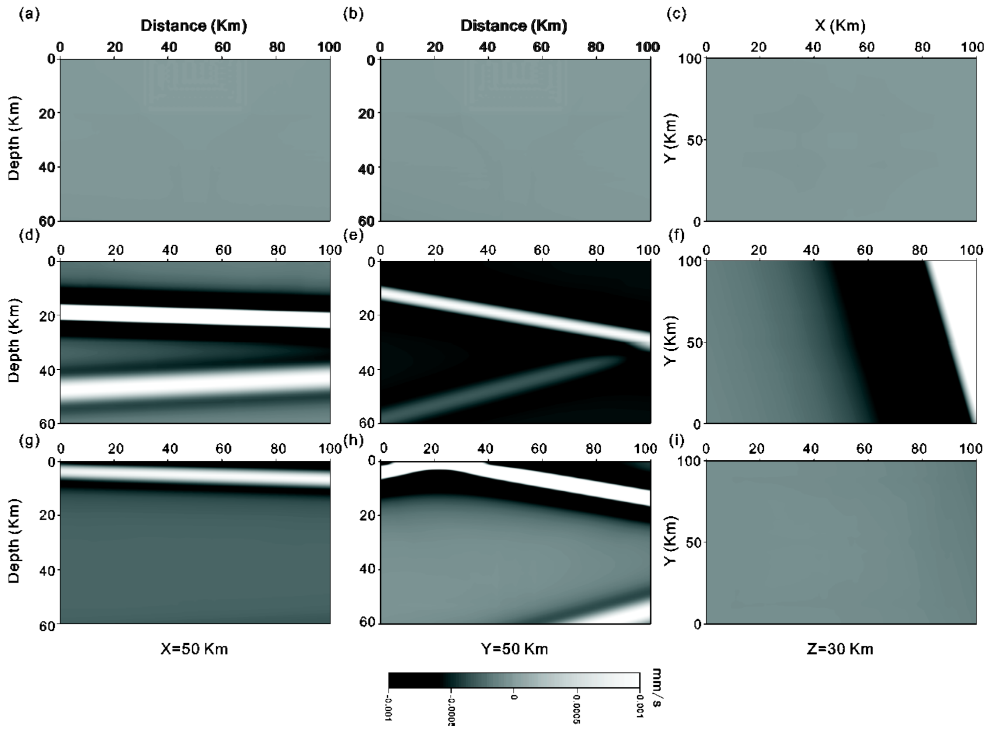

Figure 3. Plane wave incident on a 1-D crust-upper mantle model. The snapshots at the upper, middle, and lower panels are taken at t=10 s, 30 s, and 35s, respectively. The left, middle, and right panels are corresponding to snapshots in the three planes x=50 km, y=50 km, and z= 30 km, respectively. The grayscale is shown at the bottom of the Figure.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

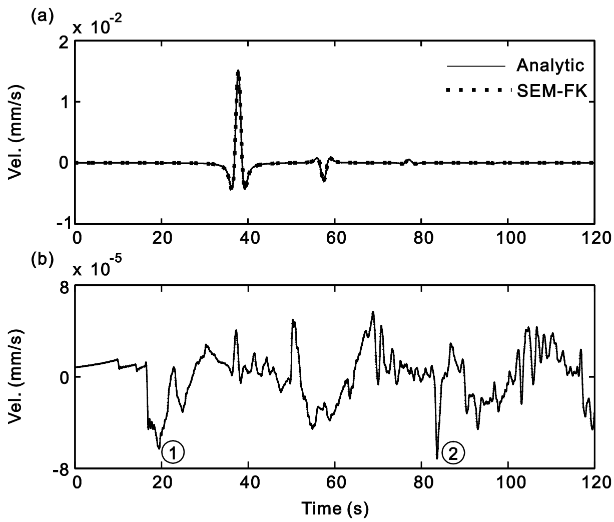

Figure 4. (a) Comparison of the waveforms generated by EBE-SEM (solid line) and FK (dotted line). The waveform computed by FK is regarded as the reference solution. (b) The error curve is obtained by subtraction of the numerical solution from the reference solution. The circle numbers in (b) denote the two error peaks.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

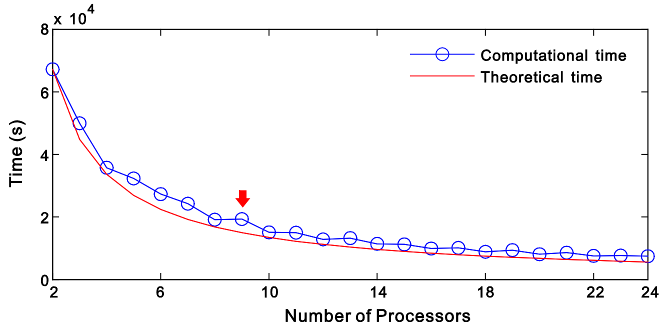

Figure 5. Parallel efficiency of the EBE-SEM. The red line denotes the theoretical CPU time; the blue line with circles is the practical CPU time. The red arrow designates the abnormal CPU time compared with the neighboring CPU times.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

Solid Earth
Discussions

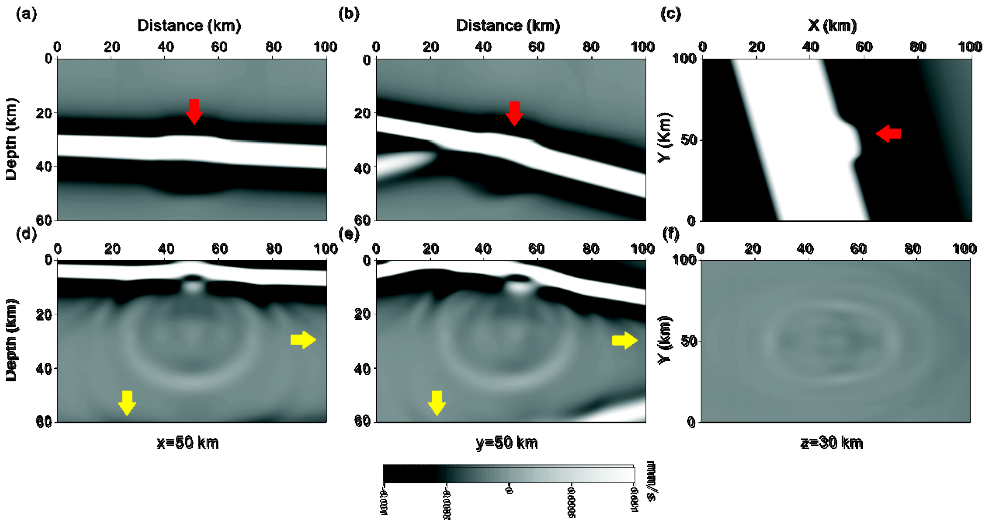

Figure 6. Snapshots in a heterogeneous model. Snapshots in the upper and lower panels are taken at t=26 s and 35 s, respectively. The left, middle, and right panels are snapshots in the x=50 km, y=50 km, and z=30 km planes, respectively. The red arrows in the upper plots indicate the distortion of the wave front because of the velocity anomaly of the media. The yellow arrows in the lower plots denote locations where the scatter waves are efficiently absorbed by the constructed PML ABC. The grayscale is show at the bottom of the figure.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

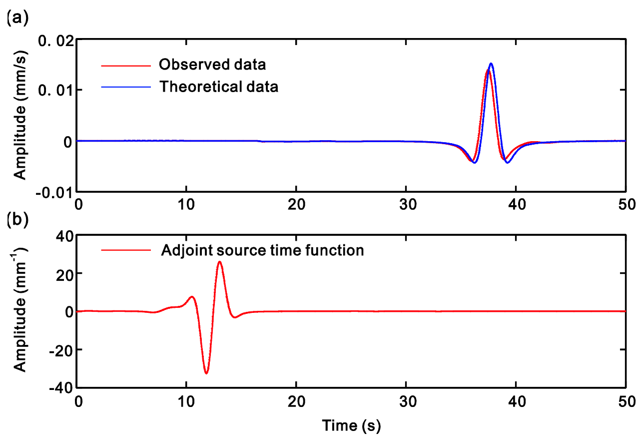

Figure 7. The construction adjoint source. (a) The synthetic waveform from the second example is regarded as the observed data (red line); the synthetic waveform from the first example is treated as the theoretical data (blue line). The adjoint-source time function is the scaled time-reversed waveform difference between the observed data and the theoretical data.

Solid Earth Discuss., doi:10.5194/se-2017-38, 2017
Manuscript under review for journal Solid Earth
Discussion started: 8 May 2017

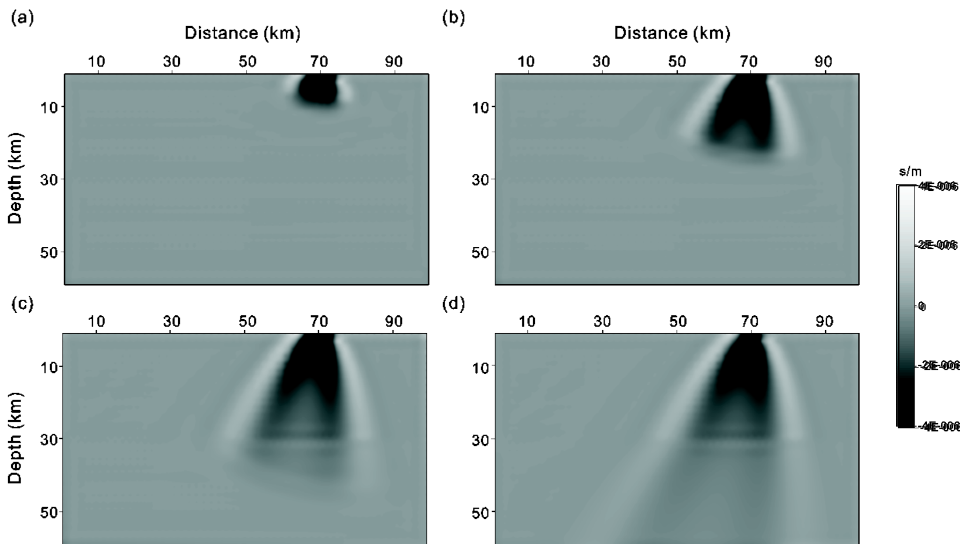

Figure 8. The misfit kernels in the y=50 km plane at four time slices. Plots (a-d) correspond to four time slices t=15 s, 20 s, 25 s, and 50 s, respectively. The grayscale is shown on the right side of the figure.