

# 1 Adjoint-state traveltome tomography for azimuthally anisotropic 2 media in spherical coordinates

3 Jing Chen<sup>1</sup>, Guoxu Chen<sup>2</sup>, Masaru Nagaso<sup>1</sup>, Ping Tong<sup>1,3,4\*</sup>

1<sup>1</sup> Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore.

2<sup>2</sup> Department of Mathematical Sciences, Tsinghua University, Beijing, China.

3 Earth Observatory of Singapore, Nanyang Technological University, Singapore.

4 Asian School of the Environment, Nanyang Technological University, Singapore.

4 Received year month day

## 5 SUMMARY

6 Tong (2021) has proposed an adjoint-state traveltome tomography method to determine velocity  
7 heterogeneity and azimuthal anisotropy. This method, however, ignores the Earth's curvature  
8 when deriving the eikonal equation for azimuthally anisotropic media. Thus, further coordinate  
9 transformation or approximation is required to ensure the accuracy of traveltome prediction  
10 in large-scale tomography. To address this problem, we derive the eikonal equation for  
11 azimuthally anisotropic media in spherical coordinates, which naturally considers the Earth's  
12 curvature. Another key ingredient is the forward modeling algorithm, whose accuracy and effi-  
13 ciency dominate the numerical error and computational cost of the inversion. In this study, we  
14 apply a modified fast sweeping method to solve the eikonal equation in spherical coordinates.  
15 Two approaches, including the third-order weighted essentially nonoscillatory approximation  
16 and multiplicative factorization technique, are applied to improve the accuracy. According to  
17 the numerical experiments, this new eikonal solver achieves a second-order accuracy and is  
18 about two orders of magnitude more accurate than the commonly used first-order fast sweeping  
19 method with similar runtime. Taking advantage of the two improvements, we develop a novel  
20 eikonal equation-based adjoint-state traveltome tomography method for azimuthally anisotropic  
21 media in spherical coordinates. This method is applicable for large-scale tomography, and its  
22 performance is verified by a synthetic checkerboard test and a practical seismic tomographic  
23 inversion in central California near Parkfield.

24 **Key words:** Inverse theory; Non-linear differential equations; Numerical solutions; Tomogra-  
25 phy; Seismic anisotropy.

## 26 1 INTRODUCTION

27 Seismic anisotropy is an intrinsic property of elastic materials that describes the dependence of seismic wave speed on the propagation  
28 direction (Long & Becker 2010). This property is widely observed in subsurface structures and has a comparable effect on traveltome  
29 observations to velocity heterogeneity (Anderson 2007; Huang et al. 2015). Thus, it is necessary to incorporate seismic anisotropy into  
30 forward modeling and tomographic inversion. Tong (2021) derives the eikonal equation-based adjoint-state traveltome tomography method  
31 for azimuthally anisotropic media. This method avoids the potential failure of ray-tracing techniques (Vidale 1988; Hole & Zelt 1995;  
32 Rawlinson et al. 2008) and is much computationally cheaper than wave equation-based tomography methods (Luo & Schuster 1991; Tromp  
33 et al. 2005; Tong et al. 2014; Yuan et al. 2016; Chen et al. 2022). Therefore, it is capable of processing large volumes of seismic traveltome data  
34 and producing reliable tomographic results. However, the method of Tong (2021) derives the eikonal equation for azimuthally anisotropic  
35 media in Cartesian coordinates, assuming that the effect of Earth's curvature is negligible. This inaccurate forward modeling may cause  
36 significant numerical errors to predicted traveltimes for large-scale tomographic problems and thus introduce artifacts to the images.

37 To address this challenge, we develop a novel eikonal equation-based adjoint-state traveltome tomography method for azimuthally  
38 anisotropic media in spherical coordinates. It includes two major improvements: First, we derive the eikonal equation for azimuthally

\* Corresponding author. E-mail address: tongping@ntu.edu.sg

## 2 Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media

1 anisotropic media in spherical coordinates. This eikonal equation naturally takes into account the Earth's curvature so that the traveltime  
 2 of the wavefront in the Earth's volume can be accurately described. Second, a modified fast sweeping method is developed to solve the  
 3 eikonal equation with second-order accuracy. As is well known, solving the eikonal equation with a point source condition suffers from the  
 4 point source singularity (Fomel et al. 2009; Luo & Qian 2011; Luo & Qian 2012; Luo et al. 2014). The eikonal solution is not differentiable  
 5 at the source, causing all high-order finite-difference eikonal solvers (Zhang et al. 2006; Xiong et al. 2010) to exhibit only first-order accuracy  
 6 (Luo & Qian 2011). Such loss in order of accuracy often leads to high computational costs for large-scale tomographic problems, because  
 7 the forward grid must be fine enough to ensure numerical accuracy and avoid the contamination of inaccurate traveltimes prediction. Many  
 8 approaches have been proposed to eliminate the point source singularity, such as the solution in celerity domain (Zhang et al. 2005), local  
 9 grid refinement (Kim & Cook 1999; Osher & Symes 2002), and solving factorized eikonal equations (Fomel et al. 2009; Luo & Qian 2012;  
 10 Luo et al. 2014). In this study, we apply the multiplicative factorization technique to the eikonal equation in spherical coordinates and solve  
 11 the factored eikonal equation instead, which avoids the point source singularity. Besides, the third-order weighted essentially nonoscillatory  
 12 (WENO) approximation scheme is used to discretize partial derivatives (Jiang & Shu 1996; Jiang & Peng 2000). These two approaches  
 13 generate a modified fast sweeping method, achieving a second-order accuracy. According to the experimental results, the new fast sweeping  
 14 method is about two orders of magnitude more accurate than the commonly used first-order fast sweeping method with similar runtime. These  
 15 improvements enable us to calculate predicted traveltimes and the sensitivity kernels of the objective function with respect to slowness and  
 16 anisotropic parameters with high accuracy. After obtaining the sensitivity kernels, we use the multiple-grid model parameterization (Tong  
 17 et al. 2019) and step-size controlled gradient descent method (Liu et al. 2019) to update model parameters. The performance of this new  
 18 adjoint-state traveltome method is evaluated by a synthetic checkerboard test and a practical seismic tomographic inversion in  
 19 central California near Parkfield.

20 The rest of the paper is organized as follows. In Section 2, we propose the eikonal equation for azimuthally anisotropic media in spherical  
 21 coordinates. Besides, the sensitivity kernels of the objective function with respect to slowness and azimuthally anisotropic parameters are  
 22 derived by using the adjoint-state method. Section 3 introduces the modified fast sweeping method to solve the eikonal and adjoint equations.  
 23 This eikonal solver achieves a second-order accuracy by employing the third-order WENO approximation and multiplicative factorization  
 24 technique. After that, several numerical experiments are conducted in Section 4 to evaluate the performance of the modified fast sweeping  
 25 method and verify the new adjoint-state traveltome method. In Section 5, we invert traveltome data in central California near  
 26 Parkfield using our method to determine the subsurface velocity and azimuthal anisotropy. The consistency between our tomographic images  
 27 and those of previous studies indicates the effectiveness of our method. Finally, we conclude this paper in Section 6.

## 28 2 ADJOINT-STATE TRAVELTIME TOMOGRAPHY METHOD

### 29 2.1 Eikonal equation for azimuthally anisotropic media

30 We consider the traveltome  $T(\mathbf{x})$  of the wavefront from a point source  $\mathbf{x}_s$  to every position  $\mathbf{x}$  in an azimuthally anisotropic medium, which  
 31 is governed by the eikonal equation with a point source condition

$$\begin{cases} \sqrt{[\nabla T(\mathbf{x})]^t M(\mathbf{x}) \nabla T(\mathbf{x})} = 1, & \mathbf{x} \in \Omega \setminus \{\mathbf{x}_s\}, \\ T(\mathbf{x}_s) = 0. \end{cases} \quad (1)$$

32 Here we consider the spherical coordinates  $\mathbf{x} = (r, \theta, \phi)$ .  $r$  is the distance to the Earth's center.  $\theta$  and  $\phi$  represent latitude and longitude,  
 33 respectively (see Figure 1a). The partial derivatives of  $T(r, \theta, \phi)$  with respect to  $r$ ,  $\theta$ ,  $\phi$  are denoted as  $\partial_r T$ ,  $\partial_\theta T$ ,  $\partial_\phi T$ . The gradient  
 34  $\nabla T(\mathbf{x}) = (\partial_r T, \frac{1}{r} \partial_\theta T, \frac{1}{r \cos \theta} \partial_\phi T)^t$ .  $\Omega$  represents the Earth's volume in the 3-D space  $R \times \Theta \times \Phi = [0, \infty) \times [-\frac{1}{2}\pi, \frac{1}{2}\pi] \times [0, 2\pi]$ .  
 35 Velocity heterogeneity and azimuthal anisotropy are described by the symmetric positive definite matrix  $M(\mathbf{x})$  (Luo & Qian 2012; Tong  
 36 2021).

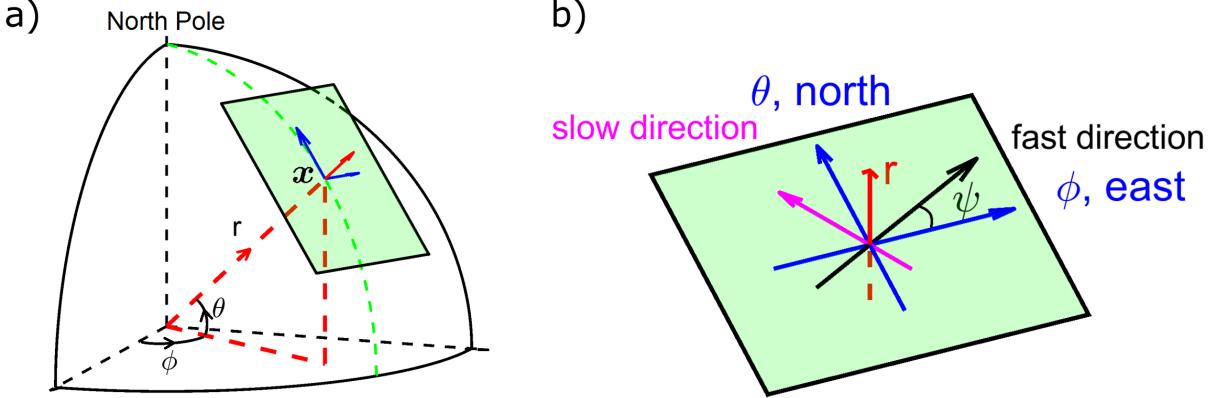
37 Next, we determine  $M(\mathbf{x})$ . It has a decomposition of form

$$M(\mathbf{x}) = c_f^2(\mathbf{x}) \mathbf{p}(\mathbf{x}) \otimes \mathbf{p}(\mathbf{x}) + c_s^2(\mathbf{x}) \mathbf{q}(\mathbf{x}) \otimes \mathbf{q}(\mathbf{x}) + c_r^2(\mathbf{x}) \mathbf{r}(\mathbf{x}) \otimes \mathbf{r}(\mathbf{x}), \quad (2)$$

38 in which  $\mathbf{p}(\mathbf{x})$ ,  $\mathbf{q}(\mathbf{x})$ ,  $\mathbf{r}(\mathbf{x})$  are the unit eigenvectors of  $M(\mathbf{x})$ , and  $c_f^2(\mathbf{x})$ ,  $c_s^2(\mathbf{x})$ ,  $c_r^2(\mathbf{x})$  are corresponding eigenvalues. The symbol  $\otimes$   
 39 indicates the outer product operator. Physically,  $c_f(\mathbf{x})$ ,  $c_s(\mathbf{x})$ , and  $c_r(\mathbf{x})$  indicate the wave velocities along eigenvector directions.

40 In the assumption of azimuthally anisotropic media, both the fast and slow velocity directions are in the tangent plane to the sphere at  $\mathbf{x}$   
 41 (Figure 1a, green plane). Correspondingly, the vertical direction is perpendicular to the tangent plane. Without loss of generality, we assume  
 42 that  $\mathbf{p}(\mathbf{x})$ ,  $\mathbf{q}(\mathbf{x})$  and  $\mathbf{r}(\mathbf{x})$  are the fast velocity direction, the slow velocity direction, and the vertical direction, respectively. It is assumed  
 43 that the fast velocity is greater than the slow velocity, i.e.,  $c_f(\mathbf{x}) \geq c_s(\mathbf{x})$ . And the velocity in the vertical direction,  $c_r(\mathbf{x})$ , is independent of  
 44 the fast velocity and slow velocity. We denote  $\psi$  ( $\psi + \frac{\pi}{2}$ ) as the anti-clockwise angle of the fast velocity direction (slow velocity direction)  
 45 with respect to the  $\phi$ -axis (see Figure 1b) so that the unit eigenvectors are given by

$$\mathbf{p}(\mathbf{x}) = (0, \sin \psi(\mathbf{x}), \cos \psi(\mathbf{x})), \quad \mathbf{q}(\mathbf{x}) = (0, \cos \psi(\mathbf{x}), -\sin \psi(\mathbf{x})), \quad \mathbf{r}(\mathbf{x}) = (1, 0, 0). \quad (3)$$



**Figure 1.** The illustration of the spherical coordinate system. (a) The point  $\mathbf{x} = (r, \theta, \phi)$  in the Earth's volume. The green plane is the tangent plane to the sphere at  $\mathbf{x}$ . (b) The tangent plane in the left figure. The red arrow is the vertical (radial) direction. Blue arrows represent the north ( $\theta$ ) and east ( $\phi$ ) directions. In the assumption of azimuthally anisotropic media, the fast and slow velocity directions lie on the tangent plane, denoted by black and magenta arrows, respectively.

Finally, we obtain the matrix

$$M(\mathbf{x}) = \begin{pmatrix} c_r^2(\mathbf{x}) & 0 & 0 \\ 0 & \frac{c_s^2(\mathbf{x}) + c_f^2(\mathbf{x})}{2} - \frac{c_f^2(\mathbf{x}) - c_s^2(\mathbf{x})}{2} \cos 2\psi(\mathbf{x}) & \frac{c_f^2(\mathbf{x}) - c_s^2(\mathbf{x})}{2} \sin 2\psi(\mathbf{x}) \\ 0 & \frac{c_f^2(\mathbf{x}) - c_s^2(\mathbf{x})}{2} \sin 2\psi(\mathbf{x}) & \frac{c_s^2(\mathbf{x}) + c_f^2(\mathbf{x})}{2} + \frac{c_f^2(\mathbf{x}) - c_s^2(\mathbf{x})}{2} \cos 2\psi(\mathbf{x}) \end{pmatrix}. \quad (4)$$

For simplicity of discussion, we define three functions  $s(\mathbf{x})$ ,  $\epsilon(\mathbf{x})$ , and  $\zeta(\mathbf{x})$  as

$$s(\mathbf{x}) = \sqrt{\frac{2}{c_f^2(\mathbf{x}) + c_s^2(\mathbf{x})}}, \quad \epsilon(\mathbf{x}) = \frac{1}{2} \frac{c_f^2(\mathbf{x}) - c_s^2(\mathbf{x})}{c_f^2(\mathbf{x}) + c_s^2(\mathbf{x})}, \quad \zeta(\mathbf{x}) = \frac{c_r^2(\mathbf{x})}{c_f^2(\mathbf{x}) + c_s^2(\mathbf{x})} - \frac{1}{2}, \quad (5)$$

and introduce two auxiliary parameters

$$\xi(\mathbf{x}) = \epsilon(\mathbf{x}) \cos 2\psi(\mathbf{x}), \quad \eta(\mathbf{x}) = \epsilon(\mathbf{x}) \sin 2\psi(\mathbf{x}). \quad (6)$$

By plugging Equations (4)-(6) into Equation (1), we derive the eikonal equation for azimuthally anisotropic media in spherical coordinates

$$\begin{cases} \sqrt{(1 + 2\zeta(\mathbf{x}))(\partial_r T)^2 + \frac{1 - 2\xi(\mathbf{x})}{r^2}(\partial_\theta T)^2 + \frac{1 + 2\xi(\mathbf{x})}{r^2 \cos^2 \theta}(\partial_\phi T)^2 + \frac{4\eta(\mathbf{x})}{r^2 \cos \theta} \partial_\theta T \partial_\phi T} = s(\mathbf{x}), & \mathbf{x} \in \Omega \setminus \{\mathbf{x}_s\}, \\ T(\mathbf{x}_s) = 0. \end{cases} \quad (7)$$

This equation involves four model parameters  $(\zeta(\mathbf{x}), \xi(\mathbf{x}), \eta(\mathbf{x}), s(\mathbf{x}))$ . Here  $s(\mathbf{x})$  represents the average horizontal slowness.  $\xi(\mathbf{x})$  and  $\eta(\mathbf{x})$  are azimuthally anisotropic parameters that describe the magnitude  $\epsilon(\mathbf{x})$  and the fast velocity direction  $\psi(\mathbf{x})$ . The parameter  $\zeta(\mathbf{x})$  is associated with the wave speed  $c_r(\mathbf{x})$  along the radial direction.

In practice, some studies make the assumption that  $\zeta(\mathbf{x}) = \gamma \sqrt{\xi^2(\mathbf{x}) + \eta^2(\mathbf{x})}$  (Eberhart-Phillips & Mark Henderson 2004; Zhao et al. 2016; Tong 2021). Particularly, if  $\gamma$  equals +1 or -1, then the vertical wave speed  $c_r$  equals the fast velocity  $c_f$  or slow velocity  $c_s$ .  $c_r$  is equal to the azimuthal average of horizontal velocities when choosing  $\gamma$  as 0. Four model parameters  $(\gamma, \xi, \eta, s)$  should be determined. However, in solving inverse problems, it is generally more challenging to determine four parameters than three. As reported by Eberhart-Phillips and Mark Henderson (2004), the selection of  $\gamma$  has negligible influence on the inversion of velocity heterogeneity and azimuthal anisotropy but indeed reduces the number of unknowns. Thus, we choose to fix  $\gamma = 0$  in this study and only invert for  $s(\mathbf{x})$ ,  $\xi(\mathbf{x})$  and  $\eta(\mathbf{x})$ .

Theoretically, singularities of Equation (7) may arise in two cases: (i)  $\cos \theta = 0$  when  $\theta = \pm 90^\circ$ ; and (ii)  $r = 0$  at the Earth's center. Alkhalifah and Fomel (2001) suggest that adding a small constant to  $\cos \theta$  can eliminate numerical singularity in case (i). For the second case (ii), it is nontrivial to remove the singularity at the Earth's center. However, considering that we hardly ever use seismic waves travelling so deep through such a single point, this singular point can be simply excluded by choosing the study region as  $R \times \Theta \times \Phi = [\varepsilon, R_E] \times [-\frac{1}{2}\pi, \frac{1}{2}\pi] \times [0, 2\pi]$ , in which  $\varepsilon$  is a small constant. This trick can avoid the singularity at the Earth's center and has nearly no influence on the traveltimes from source to receiver.

## 4 Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media

### 2.2 The Fréchet Gradient

The primary goal of traveltome tomography is to find optimal model parameters  $(s(\mathbf{x}), \xi(\mathbf{x}), \eta(\mathbf{x}))$  that minimize the discrepancy between observed and predicted traveltimes. Mathematically, this inverse problem is formulated as a partial differential equation-constrained optimization problem

$$\min_{s, \xi, \eta} \chi(s, \xi, \eta) \triangleq \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \frac{w_{n,m}}{2} \left[ T_n(\mathbf{x}_{r,m}) - T_{n,m}^{obs} \right]^2, \quad (8)$$

$$\text{subject to } \sqrt{[\nabla T_n(\mathbf{x})]^t M(\mathbf{x}; \xi, \eta) \nabla T_n(\mathbf{x})} = s(\mathbf{x}), \quad \mathbf{x} \in \Omega \setminus \{\mathbf{x}_{s,n}\}, \quad (9)$$

$$T_n(\mathbf{x}_{s,n}) = 0, \quad \forall n = 1, 2, \dots, N_s, \quad (10)$$

$$M(\mathbf{x}; \xi, \eta) = \begin{pmatrix} 1 + 2\gamma\sqrt{\xi^2 + \eta^2} & 0 & 0 \\ 0 & 1 - 2\xi & 2\eta \\ 0 & 2\eta & 1 + 2\xi \end{pmatrix}. \quad (11)$$

Here  $T_{n,m}^{obs}$  is the observed traveltme originating from the  $n$ -th source at  $\mathbf{x}_{s,n} = (r_{s,n}, \theta_{s,n}, \phi_{s,n})$  ( $n = 1, 2, \dots, N_s$ ) and recorded by the  $m$ -th receiver at  $\mathbf{x}_{r,m} = (r_{r,m}, \theta_{r,m}, \phi_{r,m})$  ( $m = 1, 2, \dots, N_r$ ).  $T_n(\mathbf{x})$  is the traveltme field originating from the  $n$ -th source, satisfying the eikonal equations (9)-(11). The weight coefficient  $w_{n,m}$  reflects the reliability of the traveltme data corresponding to the  $n$ -th source and the  $m$ -th receiver.

We assign small perturbations  $(\delta s(\mathbf{x}), \delta \xi(\mathbf{x}), \delta \eta(\mathbf{x}))$  to model parameters, which cause a small perturbation  $\delta T_n(\mathbf{x})$  to the traveltme field. Accordingly, the perturbated eikonal equation writes

$$[\nabla(T_n + \delta T_n)]^t \begin{pmatrix} 1 + 2\gamma\sqrt{(\xi + \delta\xi)^2 + (\eta + \delta\eta)^2} & 0 & 0 \\ 0 & 1 - 2\xi - 2\delta\xi & 2\eta + 2\delta\eta \\ 0 & 2\eta + 2\delta\eta & 1 + 2\xi + 2\delta\xi \end{pmatrix} [\nabla(T_n + \delta T_n)] = (s + \delta s)^2. \quad (12)$$

Subtracting Equation (9) from Equation (12) and ignoring second and higher order terms, we obtain

$$[\nabla T_n]^t M \nabla \delta T_n = [\nabla T_n]^t \begin{pmatrix} \frac{-\gamma(\xi\delta\xi + \eta\delta\eta)}{\sqrt{\xi^2 + \eta^2}} & 0 & 0 \\ 0 & \delta\xi & -\delta\eta \\ 0 & -\delta\eta & -\delta\xi \end{pmatrix} \nabla T_n + s \delta s. \quad (13)$$

We multiply a test function  $P_n(\mathbf{x})$  on both sides of Equation (13) and integrate the Earth's volume  $\Omega$

$$\int_{\Omega} P_n [\nabla T_n]^t M \nabla \delta T_n d\mathbf{x} = \int_{\Omega} P_n [\nabla T_n]^t \begin{pmatrix} \frac{-\gamma(\xi\delta\xi + \eta\delta\eta)}{\sqrt{\xi^2 + \eta^2}} & 0 & 0 \\ 0 & \delta\xi & -\delta\eta \\ 0 & -\delta\eta & -\delta\xi \end{pmatrix} \nabla T_n d\mathbf{x} + \int_{\Omega} P_n s \delta s d\mathbf{x}. \quad (14)$$

The left hand side of Equation (14) is

$$\begin{aligned} L_n &= \int_{\Omega} P_n [\nabla T_n]^t M \nabla \delta T_n d\mathbf{x} = \int_{\Omega} \nabla \cdot (P_n [\nabla T_n]^t M \delta T_n) d\mathbf{x} - \int_{\Omega} \delta T_n \nabla \cdot [P_n [\nabla T_n]^t M] d\mathbf{x} \\ &= \int_{\partial\Omega} \mathbf{n} \cdot (P_n [\nabla T_n]^t M \delta T_n) d\sigma + \int_{\Omega} \delta T_n \nabla \cdot (P_n [-\nabla T_n]^t M) d\mathbf{x}. \end{aligned} \quad (15)$$

The right hand side is splitted into three terms

$$R_n = \int_{\Omega} (P_n s^2) \frac{\delta s}{s} d\mathbf{x} + \int_{\Omega} P_n \left( \frac{-\gamma\xi(\partial_r T_n)^2}{\sqrt{\xi^2 + \eta^2}} + \frac{(\partial_\theta T_n)^2}{r^2} - \frac{(\partial_\phi T_n)^2}{r^2 \cos^2 \theta} \right) \delta \xi d\mathbf{x} + \int_{\Omega} P_n \left( \frac{-\gamma\eta(\partial_r T_n)^2}{\sqrt{\xi^2 + \eta^2}} - \frac{2\partial_\theta T_n \partial_\phi T_n}{r^2 \cos \theta} \right) \delta \eta d\mathbf{x}. \quad (16)$$

If  $P_n(\mathbf{x})$  satisfies the adjoint equation

$$\begin{cases} \nabla \cdot (P_n [-\nabla T_n]^t M) = \sum_{m=1}^{N_r} w_{n,m} (T_n(\mathbf{x}_{r,m}) - T_{n,m}^{obs}) \delta(\mathbf{x} - \mathbf{x}_{r,m}), \\ P_n(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega, \end{cases} \quad (17)$$

we can establish a relationship between Equation (15) and the perturbation  $\delta\chi$  of the objective function by ignorig the second and higher order terms

$$\begin{aligned}\delta\chi &= \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} w_{n,m} (T_n(\mathbf{x}_{r,m}) - T_{n,m}^{obs}) \delta T_n(\mathbf{x}_{r,m}) = \sum_{n=1}^{N_s} \int_{\Omega} \sum_{m=1}^{N_r} w_{n,m} (T_n(\mathbf{x}_{r,m}) - T_{n,m}^{obs}) \delta T_n(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_{r,m}) d\mathbf{x} \\ &= \sum_{n=1}^{N_s} \int_{\Omega} \delta T_n \nabla \cdot (P_n [-\nabla T_n]^t M) d\mathbf{x} = \sum_{n=1}^{N_s} L_n.\end{aligned}\quad (18)$$

1 Eventually, combining Equations (16) and (18) yields an approximate linear relationship between the perturbation  $\delta\chi$  of the objective function  
2 and the perturbations  $(\delta s/s, \delta\theta, \delta\eta)$  of model parameters , given by

$$\delta\chi = \int_{\Omega} K_s \frac{\delta s}{s} d\mathbf{x} + \int_{\Omega} K_{\xi} \delta\xi d\mathbf{x} + \int_{\Omega} K_{\eta} \delta\eta d\mathbf{x}, \quad (19)$$

3 in which

$$K_s(\mathbf{x}) = \sum_{n=1}^{N_s} P_n(\mathbf{x}) s^2(\mathbf{x}), \quad (20)$$

$$K_{\xi}(\mathbf{x}) = \sum_{n=1}^{N_s} P_n(\mathbf{x}) \left( \frac{-\gamma\xi}{\sqrt{\xi^2 + \eta^2}} (\partial_r T_n)^2 + \frac{1}{r^2} (\partial_{\theta} T_n)^2 + \frac{-1}{r^2 \cos^2 \theta} (\partial_{\phi} T_n)^2 \right), \quad (21)$$

$$K_{\eta}(\mathbf{x}) = \sum_{n=1}^{N_s} P_n(\mathbf{x}) \left( \frac{-\gamma\eta}{\sqrt{\xi^2 + \eta^2}} (\partial_r T_n)^2 + \frac{-2}{r^2 \cos \theta} \partial_{\theta} T_n \partial_{\phi} T_n \right). \quad (22)$$

4 Here  $K_s(\mathbf{x}), K_{\xi}(\mathbf{x}), K_{\eta}(\mathbf{x})$  are the Fréchet derivatives of the objective function with respect to the model parameters  $(\ln s(\mathbf{x}), \theta(\mathbf{x}), \eta(\mathbf{x}))$ ,  
5 which are also called sensitivity kernels.

### 6 2.3 Model parameterization and optimization algorithm

7 We solve eikonal and adjoint equations to obtain the traveltime field  $T(\mathbf{x})$  and adjoint field  $P(\mathbf{x})$ . These two fields are further combined to  
8 generate the sensitivity kernels on the forward grid. In general, the forward grid should be dense enough to guarantee the necessary numerical  
9 accuracy. However, it is unsuitable to discretize the perturbed model parameters on the forward grid directly, mainly because seismic data do  
10 not have the resolving ability to determine model perturbations on each node of the dense forward grid. To cope with this problem, we apply  
11 the multiple-grid model parameterization to discretize model perturbations (Tong et al. 2019).

12 First, we design  $H$  coarse grids and denote the nodes of the  $h$ -th grid by  $(r_i^h, \theta_j^h, \phi_k^h)$  ( $1 \leq i \leq N_I^h, 1 \leq j \leq N_J^h, 1 \leq k \leq N_K^h, 1 \leq$   
13  $h \leq H$ ). Each node is associated with a basis function

$$B_l^h(\mathbf{x}) = B_l^h(r, \theta, \phi) = w_i^h(r) w_j^h(\theta) w_k^h(\phi), \quad l = (k-1)N_I^h N_J^h + (j-1)N_I^h + i, \quad 1 \leq l \leq L_h = N_I^h N_J^h N_K^h, \quad (23)$$

14 where

$$w_i^h(r) = \begin{cases} \frac{r - r_{i-1}^h}{r_i^h - r_{i-1}^h}, & \text{if } r_1^h \leq r_{i-1}^h \leq r \leq r_i^h, \\ \frac{r_{i+1}^h - r}{r_{i+1}^h - r_i^h}, & \text{if } r_i^h \leq r \leq r_{i+1}^h \leq r_{N_I}^h, \\ 0, & \text{otherwise,} \end{cases} \quad (24)$$

15 and similarly for  $w_j^h(\theta)$  and  $w_k^h(\phi)$ . Then, we assume that the perturbations of model parameters are relatively smooth so that they can be  
16 formulated as linear combinations of basis functions. By averaging the contributions from all coarse grids, the model perturbations are given  
17 by

$$\frac{\delta s(\mathbf{x})}{s(\mathbf{x})} = \frac{1}{H} \sum_{h=1}^H \sum_{l=1}^{L_h} \delta C_{l,s}^h B_l^h(\mathbf{x}), \quad \delta\xi(\mathbf{x}) = \frac{1}{H} \sum_{h=1}^H \sum_{l=1}^{L_h} \delta C_{l,\xi}^h B_l^h(\mathbf{x}), \quad \delta\eta(\mathbf{x}) = \frac{1}{H} \sum_{h=1}^H \sum_{l=1}^{L_h} \delta C_{l,\eta}^h B_l^h(\mathbf{x}). \quad (25)$$

18 Plugging Equation (25) into Equation (19) yields

$$\delta\chi = \sum_{h=1}^H \sum_{l=1}^{L_h} \left( \frac{1}{H} \int_{\Omega} K_s(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x} \right) \delta C_{l,s}^h + \sum_{h=1}^H \sum_{l=1}^{L_h} \left( \frac{1}{H} \int_{\Omega} K_{\xi}(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x} \right) \delta C_{l,\xi}^h + \sum_{h=1}^H \sum_{l=1}^{L_h} \left( \frac{1}{H} \int_{\Omega} K_{\eta}(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x} \right) \delta C_{l,\eta}^h. \quad (26)$$

## 6 Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media

1 It derives the Fréchet derivatives  $\nabla_C \chi$  of the objective function  $\chi$  with respect to the auxiliary parameters  $\mathbf{C} = (C_{1,s}^1, C_{1,\xi}^1, C_{1,\eta}^1, \dots, C_{L_1,s}^1, C_{L_1,\xi}^1, C_{L_1,\eta}^1, \dots, C_{L_H,s}^H, C_{L_H,\xi}^H, C_{L_H,\eta}^H)$ , which write

$$\frac{\partial \chi(s, \xi, \eta)}{\partial C_{l,s}^h} = \frac{1}{H} \int_{\Omega} K_s(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x}, \quad \frac{\partial \chi(s, \xi, \eta)}{\partial C_{l,\xi}^h} = \frac{1}{H} \int_{\Omega} K_\xi(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x}, \quad \frac{\partial \chi(s, \xi, \eta)}{\partial C_{l,\eta}^h} = \frac{1}{H} \int_{\Omega} K_\eta(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x}. \quad (27)$$

3 Equation (27) gives new sensitivity kernels on coarse grids, which take a weighted average of original sensitivity kernels on the forward grid.  
4 Thus, it has a smooth effect on sensitivity kernels and decreases the number of unknowns as compared to the number of nodes of the forward  
5 grid. More importantly, instead of using only one grid, we design multiple grids to simultaneously update model parameters. It mitigates  
6 potential errors caused by subjective grid selection. As a result, the stability of the inversion could be enhanced using the multiple-grid model  
7 parameterization.

8 After calculating the Fréchet derivatives in Equation (27), we apply the step-size-controlled gradient descent method to update model  
9 parameters (Tong 2021). It consists of two steps: First, auxiliary parameters on coarse grids are updated along the negative gradient direction  
10 with a controlled step length  $\alpha$ , that is,  $\delta \mathbf{C} = -\alpha \nabla_C \chi$ . Second, the model parameters  $(s(\mathbf{x}), \xi(\mathbf{x}), \eta(\mathbf{x}))$  on the forward grid are  
11 updated by assigning perturbations  $(\delta s(\mathbf{x}), \delta \xi(\mathbf{x}), \delta \eta(\mathbf{x}))$  calculated by Equation (25). In this study, the step length  $\alpha$  is chosen to make  
12  $\|\delta s(\mathbf{x})/s(\mathbf{x}), \delta \xi, \delta \eta\|_\infty = 0.01$  at the beginning, which means the maximum model update at the first iteration is 1%. The maximum model  
13 update will be reduced by ten percent when the value of the objective function increases in the immediately following iteration. This strategy  
14 can be regarded as choosing a damping parameter to solve the optimization problem in an iterative fashion (Tong 2021).

### 15 3 THE FAST SWEEPING METHOD

16 One key ingredient of adjoint-state traveltome is to calculate the traveltome field  $T(\mathbf{x})$  and adjoint field  $P(\mathbf{x})$  with high accuracy  
17 and efficiency. These two fields are repeatedly calculated to obtain predicted traveltimes and sensitivity kernels. Numerical algorithms with  
18 low order of accuracy require sufficiently fine grids to ensure numerical accuracy, which may lead to high computational costs especially for  
19 large-scale tomographic problems.

20 In this study, we calculate predicted traveltimes using the fast sweeping method. This grid-based finite-difference method origins from  
21 the pioneering work of Vidale (1988), which solves the eikonal equation along an expanding square. Later, the post sweeping is applied  
22 to address the causality issue caused by the expanding square formalism (Hole & Zelt 1995), and the accuracy and stability are enhanced  
23 by using the essentially non-oscillatory (ENO) scheme (Shu & Osher 1988; Kim & Cook 1999). More recently, the fast sweeping method  
24 uses monotone numerical Hamiltonian for numerical discretization and solves the resulting system using the Gauss-Seidel iterations with  
25 alternative sweeping order, which becomes one of the most effective algorithms to solve eikonal equations (Tsai et al. 2003; Kao et al.  
26 2004; Zhao 2005; Qian et al. 2007; Luo & Qian 2012; Waheed et al. 2015; Lan & Chen 2018; Zhang et al. 2021; Cui et al. 2021; Guo et  
27 al. 2022). Importantly, this method is highly feasible for parallel computation (Zhao 2007; Detrixhe et al. 2013; Detrixhe & Gibou 2016).  
28 However, the solution of the eikonal equation with a point source condition is not differentiable at the source. Directly solving the eikonal  
29 equation always suffers from the point source singularity. Accordingly, at most first-order accuracy can be achieved even though high-order  
30 numerical schemes are applied (Luo & Qian 2012; Luo et al. 2014). To address this problem, we employ a modified fast sweeping method to  
31 solve eikonal equations, in which the multiplicative factorization technique (Fomel et al. 2009; Luo & Qian 2012; Luo et al. 2014) and the  
32 third-order WENO approximation (Jiang & Peng 2000; Zhang et al. 2006) are applied. This new method achieves a second-order accuracy  
33 so that the traveltome field  $T(\mathbf{x})$  can be accurately calculated. Apart from that, the adjoint equation (17) is also solved numerically using the  
34 fast sweeping method designed by Leung and Qian (2006).

#### 35 3.1 Multiplicative factorization technique

36 The traveltome field  $T(\mathbf{x})$  originating from the point source  $\mathbf{x}_s = (r_s, \theta_s, \phi_s)$  satisfies the eikonal equation (7). Because  $T(\mathbf{x})$  is not  
37 differentiable at the source, directly solving the eikonal equation will cause relatively large numerical errors (Luo & Qian 2011; Luo & Qian  
38 2012; Luo et al. 2014). Instead, we follow the idea proposed by Luo et al. (2014) and decompose  $T(\mathbf{x})$  into two multiplicative factors

$$T(\mathbf{x}) = U(\mathbf{x}) \tau(\mathbf{x}). \quad (28)$$

39 Here  $U(\mathbf{x})$  is a known function that captures the source singularity. For instance, we can choose  $U(\mathbf{x})$  that converges to  $T(\mathbf{x})$  when the  
40 location  $\mathbf{x}$  tends to the point source  $\mathbf{x}_s$ , given by

$$U(\mathbf{x}) = U(r, \theta, \phi) = s(\mathbf{x}_s) \left[ \frac{(r - r_s)^2}{1 + 2\zeta(\mathbf{x}_s)} + \frac{(1 + 2\xi(\mathbf{x}_s))r_s^2}{1 - 4\xi^2(\mathbf{x}_s) - 4\eta^2(\mathbf{x}_s)} (\theta - \theta_s)^2 \right. \\ \left. + \frac{(1 - 2\xi(\mathbf{x}_s))r_s^2 \cos^2 \theta_s}{1 - 4\xi^2(\mathbf{x}_s) - 4\eta^2(\mathbf{x}_s)} (\phi - \phi_s)^2 - \frac{4\eta(\mathbf{x}_s)r_s^2 \cos \theta_s}{1 - 4\xi^2(\mathbf{x}_s) - 4\eta^2(\mathbf{x}_s)} (\theta - \theta_s)(\phi - \phi_s) \right]^{1/2}. \quad (29)$$

41 Accordingly, the other factor  $\tau(\mathbf{x})$  is smooth near the source, satisfying the factored eikonal equation

$$\begin{cases} H_f(\tau, \tau_r, \tau_\theta, \tau_\eta) \triangleq \sqrt{(1+2\zeta)(U_r\tau + \tau_r U)^2 + \frac{1-2\xi}{r^2}(U_\theta\tau + \tau_\theta U)^2 + \frac{1+2\xi}{r^2 \cos^2 \theta}(U_\phi\tau + \tau_\phi U)^2} \\ + \frac{4\eta}{r^2 \cos \theta}(U_\theta\tau + \tau_\theta U)(U_\phi\tau + \tau_\phi U) \\ \tau(\mathbf{x}_s) = 1, \end{cases} = s(\mathbf{x}), \quad (30)$$

in which  $(\tau_r, \tau_\theta, \tau_\eta)$  and  $(U_r, U_\theta, U_\eta)$  are partial derivatives of  $\tau(\mathbf{x})$  and  $U(\mathbf{x})$  with respect to  $(r, \theta, \eta)$ , respectively. Rather than solving the original eikonal equation, we choose to solve the factored eikonal equation (30) so that the point source singularity is eliminated. According to the experimental results presented in Section 4, the numerical solver has a second-order accuracy when calculating traveltime fields. In fact, the eikonal solver can achieve third- or even higher-order accuracy if the factor  $U(\mathbf{x})$  is precisely designed to generate a high-order smooth function  $\tau(\mathbf{x})$  near the source (Luo et al. 2014). Analysis and discussion in detail are presented in Appendix A.

### 3.2 The third-order WENO-based Lax-Friedrichs fast sweeping method

The numerical eikonal solver consists of two key components: the accurate and stable local discretization scheme for the eikonal equation, and the efficient method to solve this discretization system following the causality (Zhao 2005). Here we first discuss the local discretization scheme. In general, there may exist multiple continuous solutions that satisfy the eikonal equations but are not differentiable everywhere (Crandall & Lions 1983). Thus, we need to carefully design the local discretization scheme to obtain the unique viscosity solution consistent with the physical observation (Crandall & Lions 1983). Rouy and Tourin (1992) prove that a stable, consistent and monotone numerical scheme, such as the numerical Hamiltonian, could converge to the viscosity solution. Furthermore, the WENO scheme is proposed for approximating Hamiltonian, which yields high order of accuracy and resolves the discontinuities of derivatives (Osher & Shu 1991; Jiang & Peng 2000). Thus, we use the third-order WENO-based Lax-Friedrichs Hamiltonian as the local discretization scheme for the factored eikonal equation (30) (Zhang et al. 2006; Luo & Qian 2011; Luo et al. 2014).

We denote  $u = \tau_r$ ,  $v = \tau_\theta$ , and  $w = \tau_\phi$ . The Lax-Friedrichs Hamiltonian (Crandall & Lions 1984) is

$$H_{LF}(\tau, u^+, u^-, v^+, v^-, w^+, w^-) = H_f(\tau, \frac{u^+ + u^-}{2}, \frac{v^+ + v^-}{2}, \frac{w^+ + w^-}{2}) - (\sigma_r \frac{u^+ - u^-}{2} + \sigma_\theta \frac{v^+ - v^-}{2} + \sigma_\phi \frac{w^+ - w^-}{2}), \quad (31)$$

where  $u^\pm, v^\pm, w^\pm$  are the forward and backward difference approximations of  $\tau_r, \tau_\theta$ , and  $\tau_\phi$ .  $H_f(\tau, u, v, w)$  is already clarified in Equation (30).  $\sigma_r, \sigma_\theta, \sigma_\phi$  are artificial viscosities to ensure monotonicity property that the numerical Hamiltonian  $H_{LF}$  is nonincreasing in  $u^+, v^+, w^+$  and nondecreasing in  $u^-, v^-, w^-$  (Osher & Shu 1991; Kao et al. 2004). These parameters should satisfy

$$\sigma_r \geq \max_{\substack{u \in [A, B], v \in [C, D] \\ w \in [E, F]}} \left| \frac{\partial H_f}{\partial u}(\tau, u, v, w) \right|, \quad \sigma_\theta \geq \max_{\substack{u \in [A, B], v \in [C, D] \\ w \in [E, F]}} \left| \frac{\partial H_f}{\partial v}(\tau, u, v, w) \right|, \quad \sigma_\phi \geq \max_{\substack{u \in [A, B], v \in [C, D] \\ w \in [E, F]}} \left| \frac{\partial H_f}{\partial w}(\tau, u, v, w) \right|, \quad (32)$$

in which  $[A, B]$ ,  $[C, D]$ , and  $[E, F]$  are the respective value ranges for  $u^\pm, v^\pm$ , and  $w^\pm$  (Zhang et al. 2006). Particularly, for moderate azimuthal anisotropies where  $1 - 4\xi^2 - 4\eta^2 \geq 0$ , we can simply choose

$$\sigma_r = U(x)\sqrt{1+2\zeta}, \quad \sigma_\theta = U(x)\sqrt{\frac{1-2\xi}{r^2}}, \quad \sigma_\phi = U(x)\sqrt{\frac{1+2\xi}{r^2 \cos^2 \theta}}, \quad (33)$$

to satisfy the monotonicity property (see analysis in Appendix B).

Now we consider the numerical Hamiltonian on a uniform grid with grid spacing of  $\Delta r, \Delta\theta, \Delta\phi$ . At the grid node  $(r_i, \theta_j, \phi_k)$ , approximating  $H_f$  by  $H_{LF}$  yields

$$s_{i,j,k} = [H_{LF}]_{i,j,k} = [H_f]_{i,j,k} - (\sigma_r \frac{D_{i,j,k}^{r+} + D_{i,j,k}^{r-}}{2\Delta r} + \sigma_\theta \frac{D_{i,j,k}^{\theta+} + D_{i,j,k}^{\theta-}}{2\Delta\theta} + \sigma_\phi \frac{D_{i,j,k}^{\phi+} + D_{i,j,k}^{\phi-}}{2\Delta\phi}) + (\frac{\sigma_r}{\Delta r} + \frac{\sigma_\theta}{\Delta\theta} + \frac{\sigma_\phi}{\Delta\phi})\tau_{i,j,k}. \quad (34)$$

in which

$$\begin{aligned} D_{i,j,k}^{r+} &= \tau_{i,j,k} + \Delta r \cdot u_{i,j,k}^+, & D_{i,j,k}^{r-} &= \tau_{i,j,k} - \Delta r \cdot u_{i,j,k}^-, \\ D_{i,j,k}^{\theta+} &= \tau_{i,j,k} + \Delta\theta \cdot v_{i,j,k}^+, & D_{i,j,k}^{\theta-} &= \tau_{i,j,k} - \Delta\theta \cdot v_{i,j,k}^-, \\ D_{i,j,k}^{\phi+} &= \tau_{i,j,k} + \Delta\phi \cdot w_{i,j,k}^+, & D_{i,j,k}^{\phi-} &= \tau_{i,j,k} - \Delta\phi \cdot w_{i,j,k}^-, \\ [H_f]_{i,j,k} &= H_f(\tau_{i,j,k}, \frac{u_{i,j,k}^+ + u_{i,j,k}^-}{2}, \frac{v_{i,j,k}^+ + v_{i,j,k}^-}{2}, \frac{w_{i,j,k}^+ + w_{i,j,k}^-}{2}). \end{aligned} \quad (35)$$

Replacing  $\tau_{i,j,k}$  on the right hand side by  $\tau_{i,j,k}^{new}$  leads to the update formula

## 8 Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media

$$\tau_{i,j,k}^{new} = \left( \frac{1}{\frac{\sigma_r}{\Delta r} + \frac{\sigma_\theta}{\Delta \theta} + \frac{\sigma_\phi}{\Delta \phi}} \right) \left( s_{i,j,k} - [H_f]_{i,j,k} + \sigma_r \frac{D_{i,j,k}^{r+} + D_{i,j,k}^{r-}}{2\Delta r} + \sigma_\theta \frac{D_{i,j,k}^{\theta+} + D_{i,j,k}^{\theta-}}{2\Delta \theta} + \sigma_\phi \frac{D_{i,j,k}^{\phi+} + D_{i,j,k}^{\phi-}}{2\Delta \phi} \right). \quad (36)$$

1 The forward and backward partial derivatives  $u_{i,j,k}^\pm$  in Equation (35) are approximated using the third-order WENO scheme (Jiang & Peng  
2 2000; Zhang et al. 2006), formulated by

$$u_{i,j,k}^+ = (1 - W_+) \left( \frac{\tau_{i+1,j,k} - \tau_{i-1,j,k}}{2\Delta r} \right) + W_+ \left( \frac{-3\tau_{i,j,k} + 4\tau_{i+1,j,k} - \tau_{i+2,j,k}}{2\Delta r} \right), \quad (37)$$

$$u_{i,j,k}^- = (1 - W_-) \left( \frac{\tau_{i+1,j,k} - \tau_{i-1,j,k}}{2\Delta r} \right) + W_- \left( \frac{3\tau_{i,j,k} - 4\tau_{i-1,j,k} + \tau_{i-2,j,k}}{2\Delta r} \right), \quad (38)$$

3 in which

$$W_+ = \frac{1}{1 + 2g_+^2}, \quad g_+ = \frac{\epsilon + (\tau_{i,j,k} - 2\tau_{i+1,j,k} + \tau_{i+2,j,k})^2}{\epsilon + (\tau_{i-1,j,k} - 2\tau_{i,j,k} + \tau_{i+1,j,k})^2}, \quad (39)$$

$$W_- = \frac{1}{1 + 2g_-^2}, \quad g_- = \frac{\epsilon + (\tau_{i,j,k} - 2\tau_{i-1,j,k} + \tau_{i-2,j,k})^2}{\epsilon + (\tau_{i+1,j,k} - 2\tau_{i,j,k} + \tau_{i-1,j,k})^2}. \quad (40)$$

4 Here  $\epsilon$  is chosen as  $10^{-12}$  to avoid division by zero. Similar approximations are also made for  $v_{i,j,k}^\pm$  and  $w_{i,j,k}^\pm$ .

5 The second key component is to solve the discretized system, which is required to follow the causality along characteristics (Vidale  
6 1988). In the fast sweeping method, the traveltimes field is updated using the Gauss-Seidel iterations with alternating sweeping orders:

$$\begin{array}{ll} (1) i = 1 : N_I, j = 1 : N_J, k = 1 : N_K; & (2) i = 1 : N_I, j = 1 : N_J, k = N_K : 1; \\ (3) i = 1 : N_I, j = N_J : 1, k = 1 : N_K; & (4) i = 1 : N_I, j = N_J : 1, k = N_K : 1; \\ (5) i = N_I : 1, j = 1 : N_J, k = 1 : N_K; & (6) i = N_I : 1, j = 1 : N_J, k = N_K : 1; \\ (7) i = N_I : 1, j = N_J : 1, k = 1 : N_K; & (8) i = N_I : 1, j = N_J : 1, k = N_K : 1. \end{array} \quad (41)$$

7 These eight sweeping orders represent eight groups of propagation directions in 3D space, which cover all directions of characteristics. Each  
8 characteristic is composed of several pieces that belong to one of the above groups. Thus, the traveltime information propagating alone  
9 characteristics will be piecewise and simultaneously recovered by iteratively sweeping with the alternating order (Zhao 2005; Qian et al.  
10 2007). The numerical algorithm of the fast sweeping method to solve the eikonal equation (7) is presented below.

11 (i) **Initialization:** The multiplicative factor  $U(\mathbf{x})$  is given by Equation (29). We set the iteration step index  $k = 0$ , and the initial guess  
12  $\tau^{(0)}(\mathbf{x}) = 1$ . The convergence threshold  $\sigma$  is selected as  $10^{-5}$ .

13 (ii) **Iteration:** Let  $\tau(\mathbf{x}) = \tau^{(k)}(\mathbf{x})$ . We sweep the whole computational domain by the Gauss-Seidel iterations with alternating sweeping  
14 orders as in Equation (41).  $\tau_{i,j,k}^{new}$  is calculated at each grid node  $\mathbf{x}_{i,j,k}$  in the following four cases:

15 **Case 1.** If  $\mathbf{x}_{i,j,k}$  is a neighbor of the point source  $\mathbf{x}_s$ , that is,  $\frac{|r_i - r_s|}{\Delta r}, \frac{|\theta_j - \theta_s|}{\Delta \theta}, \frac{|\phi_k - \phi_s|}{\Delta \phi} \leq 2$ , let  $\tau_{i,j,k}^{new} = 1$ ;

16 **Case 2.** If  $\mathbf{x}_{i,j,k}$  is on the computational boundary, we calculate  $\tau_{i,j,k}^{new}$  with linear extrapolation (Kao et al. 2004), given by

$$\begin{aligned} \tau_{1,j,k}^{new} &= \min(\max(2\tau_{2,j,k} - \tau_{3,j,k}, \tau_{3,j,k}), \tau_{1,j,k}), & \tau_{N_I,j,k}^{new} &= \min(\max(2\tau_{N_I-1,j,k} - \tau_{N_I-2,j,k}, \tau_{N_I-2,j,k}), \tau_{N_I,j,k}), \\ \tau_{i,1,k}^{new} &= \min(\max(2\tau_{i,2,k} - \tau_{i,3,k}, \tau_{i,3,k}), \tau_{i,1,k}), & \tau_{i,N_J,k}^{new} &= \min(\max(2\tau_{i,N_J-1,k} - \tau_{i,N_J-2,k}, \tau_{i,N_J-2,k}), \tau_{i,N_J,k}), \\ \tau_{i,j,1}^{new} &= \min(\max(2\tau_{i,j,2} - \tau_{i,j,3}, \tau_{i,j,3}), \tau_{i,j,1}), & \tau_{i,j,N_K}^{new} &= \min(\max(2\tau_{i,j,N_K-1} - \tau_{i,j,N_K-2}, \tau_{i,j,N_K-2}), \tau_{i,j,N_K}). \end{aligned} \quad (42)$$

17 **Case 3.** If  $\mathbf{x}_{i,j,k}$  is adjacent to the computational boundary, we replace the third-order WENO approximation with the first-order scheme:

$$\begin{aligned} u_{2,j,k}^- &= \frac{1}{\Delta r} (\tau_{2,j,k} - \tau_{1,j,k}), & u_{N_I-1,j,k}^+ &= \frac{1}{\Delta r} (\tau_{N_I,j,k} - \tau_{N_I-1,j,k}), \\ v_{i,2,k}^- &= \frac{1}{\Delta \theta} (\tau_{i,2,k} - \tau_{i,1,k}), & v_{i,N_J-1,k}^+ &= \frac{1}{\Delta \theta} (\tau_{i,N_J,k} - \tau_{i,N_J-1,k}), \\ w_{i,j,2}^- &= \frac{1}{\Delta \phi} (\tau_{i,j,2} - \tau_{i,j,1}), & w_{i,j,N_K-1}^+ &= \frac{1}{\Delta \phi} (\tau_{i,j,N_K} - \tau_{i,j,N_K-1}). \end{aligned} \quad (43)$$

18 Then, calculate  $\tau_{i,j,k}^{new}$  according to Equation (36);

19 **Case 4.** Otherwise, calculate  $\tau_{i,j,k}^{new}$  according to Equation (36).

20 (iii) **Convergence:** If the change of  $\tau(\mathbf{x})$  between two iteration steps is smaller than the threshold  $\sigma$ , that is,  $\|\tau^{new}(\mathbf{x}) - \tau(\mathbf{x})\|_1 < \sigma$ ,  
21 we terminate the iteration and output the traveltimes field  $T(\mathbf{x}) = U(\mathbf{x})\tau^{new}(\mathbf{x})$ . Otherwise, let  $k = k + 1$ ,  $\tau^{(k)}(\mathbf{x}) = \tau^{new}(\mathbf{x})$  and return  
22 to Step ‘Iteration’.

23 As summarized above, we calculate the traveltimes field near the computational boundary using the first-order approximation. As a result,  
24 only moderate accuracy can be achieved near the computational boundary (Zhang et al. 2006). In practical application, we usually extend the  
25 computational domain to make sources and receivers far from the computational boundary. The traveltimes from sources to receivers, thus,  
26 can be obtained with high accuracy.

### 3.3 The fast sweeping method for the adjoint equation

We follow the same idea proposed by Leung and Qian (2006) to discretize the adjoint equation in the conservation form. The coefficient matrix of this resulting linear system is irreducibly diagonally dominant (Leung & Qian 2006). Thus, it will converge to the solution by using the Gauss-Seidel iteration. Detailed discrete formulations are given below.

The adjoint equation (17) is specified as

$$\partial_r(aP_n) + \partial_\theta(bP_n) + \partial_\phi(cP_n) = \sum_{m=1}^{N_r} w_{n,m} \left( T_n(\mathbf{x}_{r,m}) - T_{n,m}^{obs} \right) \delta(\mathbf{x} - \mathbf{x}_{r,m}), \quad (44)$$

in which

$$a = -(1 + 2\zeta)\partial_r T_n(\mathbf{x}), \quad b = -\frac{1 - 2\xi}{r^2} \partial_\theta T_n - \frac{2\eta}{r^2 \cos \theta} \partial_\phi T_n, \quad c = -\frac{1 + 2\xi}{r^2 \cos^2 \theta} \partial_\phi T_n - \frac{2\eta}{r^2 \cos \theta} \partial_\theta T_n. \quad (45)$$

This equation can be numerically solved by using a modified fast sweeping method (Leung & Qian 2006) whose update formula is

$$\begin{aligned} [P_n]_{i,j,k}^{new} = & \frac{1}{\frac{a_{i+1/2,j,k}^+ - a_{i-1/2,j,k}^-}{\Delta r} + \frac{b_{i,j+1/2,k}^+ - b_{i,j-1/2,k}^-}{\Delta \theta} + \frac{c_{i,j,k+1/2}^+ - c_{i,j,k-1/2}^-}{\Delta \phi}} \\ & \left( \frac{a_{i-1/2,j,k}^+ [P_n]_{i-1,j,k} - a_{i+1/2,j,k}^- [P_n]_{i+1,j,k}}{\Delta r} + \frac{b_{i,j-1/2,k}^+ [P_n]_{i,j-1,k} - b_{i,j+1/2,k}^- [P_n]_{i,j+1,k}}{\Delta \theta} \right. \\ & \left. + \frac{c_{i,j,k-1/2}^+ [P_n]_{i,j,k-1} - c_{i,j,k+1/2}^- [P_n]_{i,j,k+1}}{\Delta \phi} \right) + \sum_{m=1}^{N_r} w_{n,m} \left( T_n(\mathbf{x}_{r,m}) - T_{n,m}^{obs} \right) \delta_{i,j,k}(\mathbf{x}_{r,m}). \end{aligned} \quad (46)$$

Here we have

$$a_{i-1/2,j,k}^\pm = \frac{a_{i-1/2,j,k} \pm |a_{i-1/2,j,k}|}{2}, \quad b_{i,j-1/2,k}^\pm = \frac{b_{i,j-1/2,k} \pm |b_{i,j-1/2,k}|}{2}, \quad c_{i,j,k-1/2}^\pm = \frac{c_{i,j,k-1/2} \pm |c_{i,j,k-1/2}|}{2}, \quad (47)$$

$$a_{i+1/2,j,k}^\pm = \frac{a_{i+1/2,j,k} \pm |a_{i+1/2,j,k}|}{2}, \quad b_{i,j+1/2,k}^\pm = \frac{b_{i,j+1/2,k} \pm |b_{i,j+1/2,k}|}{2}, \quad c_{i,j,k+1/2}^\pm = \frac{c_{i,j,k+1/2} \pm |c_{i,j,k+1/2}|}{2}, \quad (48)$$

in which the central difference scheme is employed to approximate the coefficients  $a_{i-1/2,j,k}$ ,  $b_{i,j-1/2,k}$ ,  $c_{i,j,k-1/2}$  as

$$a_{i-1/2,j,k} = -(1 + \zeta_{i-1,j,k} + \zeta_{i,j,k}) \frac{[T_n]_{i,j,k} - [T_n]_{i-1,j,k}}{\Delta r}, \quad (49)$$

$$\begin{aligned} b_{i,j-1/2,k} = & -\frac{1 - \xi_{i,j-1,k} - \xi_{i,j,k}}{r_i^2} \left( \frac{[T_n]_{i,j,k} - [T_n]_{i,j-1,k}}{\Delta \theta} \right. \\ & \left. - \frac{\eta_{i,j-1,k} + \eta_{i,j,k}}{r_i^2 \cos \theta_{j-1/2}} \left( \frac{[T_n]_{i,j,k+1} - [T_n]_{i,j,k-1} + [T_n]_{i,j-1,k+1} - [T_n]_{i,j-1,k-1}}{4 \Delta \phi} \right) \right), \end{aligned} \quad (50)$$

$$\begin{aligned} c_{i,j,k-1/2} = & -\frac{1 + \xi_{i,j,k-1} + \xi_{i,j,k}}{r_i^2 \cos^2 \theta_j} \left( \frac{[T_n]_{i,j,k} - [T_n]_{i,j,k-1}}{\Delta \phi} \right) \\ & - \frac{\eta_{i,j,k-1} + \eta_{i,j,k}}{r_i^2 \cos \theta_j} \left( \frac{[T_n]_{i,j+1,k-1} - [T_n]_{i,j-1,k-1} + [T_n]_{i,j+1,k} - [T_n]_{i,j-1,k}}{4 \Delta \theta} \right). \end{aligned} \quad (51)$$

$a_{i+1/2,j,k}$  can be calculated by setting  $i$  as  $i + 1$  in Equation (49), and similarly for  $b_{i,j+1/2,k}$  and  $c_{i,j,k+1/2}$ . In addition, the delta function is discretized by

$$\delta_{i,j,k}(\mathbf{x}_{r,m}) = \frac{w_i(r_{r,m}) w_j(\theta_{r,m}) w_k(\phi_{r,m})}{\Delta r \Delta \theta \Delta \phi}, \quad (52)$$

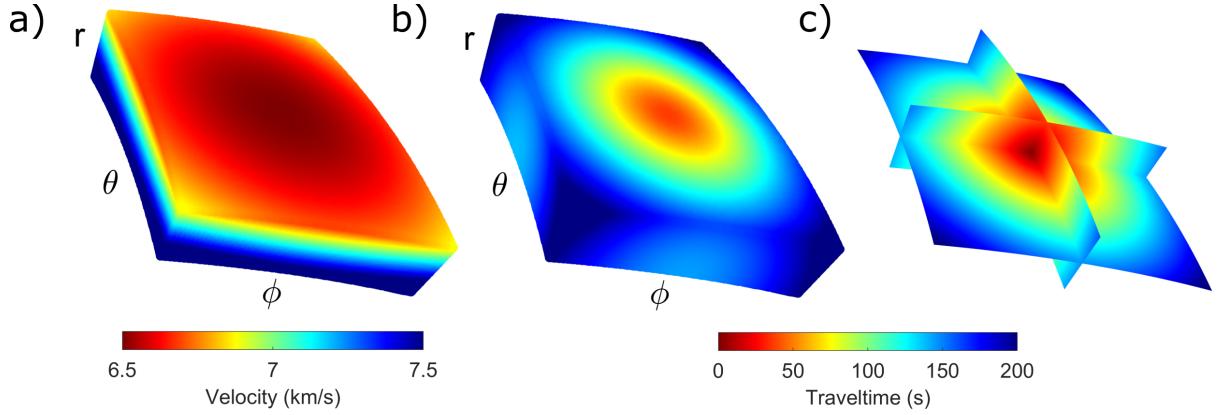
in which  $w_i(r)$  satisfies

$$w_i(r) = \begin{cases} \frac{r - r_{i-1}}{r_i - r_{i-1}}, & \text{if } r_1 \leq r_{i-1} \leq r \leq r_i, \\ \frac{r_{i+1} - r}{r_{i+1} - r_i}, & \text{if } r_i \leq r \leq r_{i+1} \leq r_{N_I}, \\ 0, & \text{otherwise,} \end{cases} \quad (53)$$

and similarly for  $w_j(\theta)$  and  $w_k(\phi)$ .

Here, we summarize the numerical algorithm of the fast sweeping method to solve the adjoint equation (17).

10    *Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media*



**Figure 2.** Illustration of the velocity model and traveltime field in Example 1. (a) The 3-D velocity model  $c(\mathbf{x}) = \frac{1}{s(\mathbf{x})}$ . (b) The 3-D true traveltime field  $T_{ana}(\mathbf{x})$ . (c) The horizontal and vertical sections of  $T_{ana}(\mathbf{x})$  crossing the point source  $\mathbf{x}_s$ .

- 1       (i) **Initialization:** We set the iteration step index  $k = 0$ , and the initial guess  $P_n^{(0)}(\mathbf{x}) = 0$ . The convergence threshold  $\sigma$  is selected as  $10^{-5}$ .
- 2       (ii) **Iteration:** Let  $P_n(\mathbf{x}) = P_n^{(k)}(\mathbf{x})$ . We sweep the whole computational domain by the Gauss-Seidel iterations with alternating sweeping
- 3       orders as in Equation (41).  $[P_n]_{i,j,k}^{new}$  is calculated at each grid node  $\mathbf{x}_{i,j,k}$  in the following two cases:
- 4       **Case 1.** If  $\mathbf{x}_{i,j,k}$  is on the computational boundary, let  $[P_n]_{i,j,k}^{new} = 0$  to satisfy the boundary condition in Equation (17).
- 5       **Case 2.** Otherwise, calculate  $[P_n]_{i,j,k}^{new}$  according to Equation (46).
- 6       (iii) **Convergence:** If the change of  $P_n(\mathbf{x})$  between two iteration steps is smaller than the threshold  $\sigma$ , that is,  $\|P_n^{new}(\mathbf{x}) - P_n(\mathbf{x})\|_1 < \sigma$ , we terminate the iteration and output the adjoint field  $P_n^{new}(\mathbf{x})$ . Otherwise, let  $k = k + 1$ ,  $P_n^{(k)}(\mathbf{x}) = P_n^{new}(\mathbf{x})$  and return to Step ‘Iteration’.

9    **4 NUMERICAL EXPERIMENTS**

10    **4.1 Forward modeling**

11    We conduct two numerical experiments to evaluate the new fast sweeping method (FSM). The parameter setting is designed for the purpose of performance evaluation, which may differs from the actual earth model. The calculated traveltime fields are compared with analytic solutions in isotropic and anisotropic media. The results generated by the first-order fast sweeping method (first-order FSM) are also presented for comparison.

15    **Example 1. 3-D isotropic media.** Consider a bounded domain  $R \times \Theta \times \Phi = [5900 \text{ km}, 6400 \text{ km}] \times [30^\circ, 50^\circ] \times [15^\circ, 40^\circ]$ . The point source  $\mathbf{x}_s = (r_s, \theta_s, \phi_s) = (6150 \text{ km}, 40^\circ, 27.5^\circ)$  is located at the center of the domain. The slowness  $s(\mathbf{x})$  is given by

$$\frac{1}{s(\mathbf{x})} = 7.0 + \mathbf{g} \cdot \mathbf{d}(\mathbf{x}) \quad (54)$$

17    where

$$\mathbf{g} = (-1.36 \times 10^{-3}, -7.08 \times 10^{-4}, -1.29 \times 10^{-3}), \quad (55)$$

$$\mathbf{d}(\mathbf{x}) = (r \cos \theta \cos \phi - r_s \cos \theta_s \cos \phi_s, r \cos \theta \sin \phi - r_s \cos \theta_s \sin \phi_s, r \sin \theta - r_s \sin \theta_s). \quad (56)$$

18    Figure 2a depicts the velocity model  $c(\mathbf{x}) = 1/s(\mathbf{x})$ , which resembles the wave velocity in the Earth’s volume that increases with depth. Besides, the anisotropic parameters  $\xi(\mathbf{x})$  and  $\eta(\mathbf{x})$  are zero in this isotropic model. Based on this parameter setting, the analytic solution of the eikonal equation writes

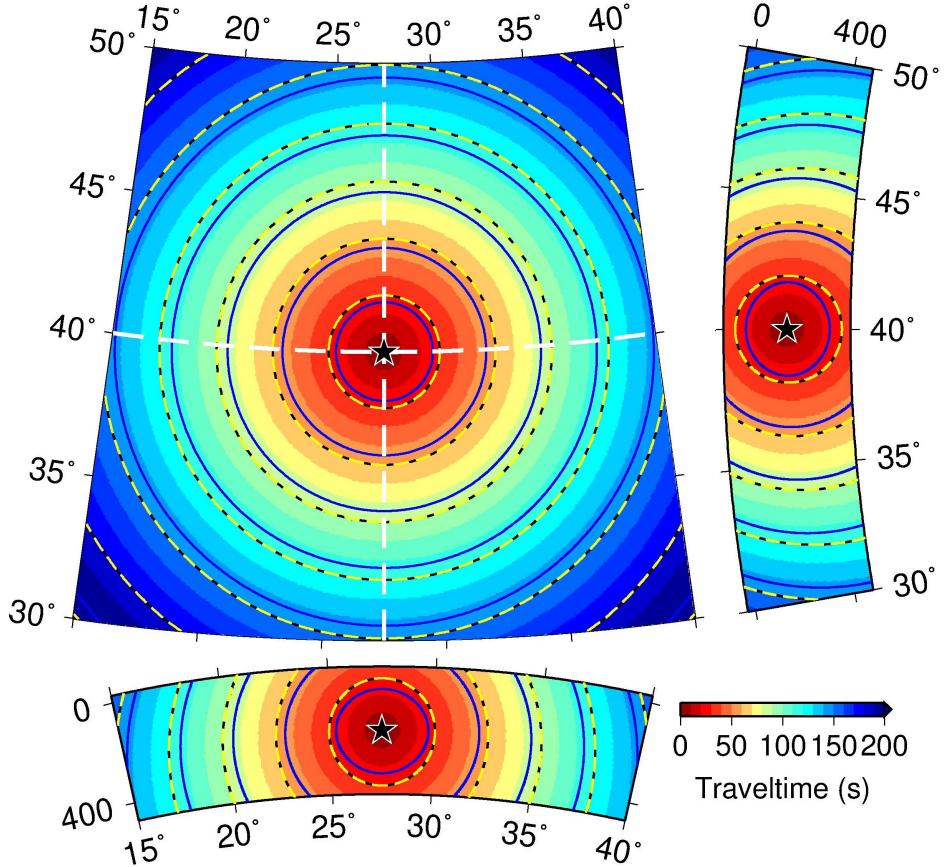
$$T_{ana}(\mathbf{x}) = \frac{1}{\|\mathbf{g}\|_2} \operatorname{arccosh}(1 + \frac{1}{2} s(\mathbf{x}) s(\mathbf{x}_s) \|\mathbf{g}\|_2^2 \|\mathbf{d}(\mathbf{x})\|_2^2), \quad (57)$$

21    in which

$$\operatorname{arccosh}(z) = \ln(z + \sqrt{z^2 - 1}), \quad z \geq 1. \quad (58)$$

22    Figure 2b shows the traveltime field in the computational domain, and Figure 2c plots the 2-D profiles crossing the point source.

23    We illustrate the solutions of the eikonal equation in horizontal and vertical sections crossing the point source in Figure 3 and also plot



**Figure 3.** The horizontal and vertical sections crossing the point source (black star) of the traveltime field calculated on a mesh of  $120 \times 120 \times 120$  grid points for Example 1, which correspond to the profiles in Figure 2c. Black and yellow dashed lines are the contours of the analytic solution  $T_{ana}(\mathbf{x})$  and the result generated by the new FSM, respectively. Solid blue lines represent the contours of the result obtained using the first-order FSM. Left top: the horizontal section at  $r = 6150$  km. The white dashed lines are the location of vertical sections; Left bottom: the vertical section at  $\theta = 40^\circ$ ; Right top: the vertical section at  $\phi = 27.5^\circ$ , "0" and "400" in the vertical axis correspond to  $r = 6371$  km and  $r = 5971$  km.

the traveltime errors in Figure S1. The traveltime field generated by the new FSM (dashed yellow lines) is almost the same as the analytic solution (dashed black lines). Since the point source singularity is eliminated and the third-order WENO scheme is applied, traveltime errors inside the computational domain are negligible and smaller than the moderate errors near the boundary, where only first-order approximation is adopted (Figure S1a). Fortunately, the greater errors near the boundary do not affect the accuracy in the interior of the computational domain (see Remark 8 in Zhang et al. (2006)). By contrast, a prominent gap exists between the analytic solution and the result obtained using the first-order FSM (solid blue lines). The traveltime error accumulates dramatically near the source resulting from the point source singularity and keeps increasing with the distance away from the source due to numerical approximation (Figure S1b). Table 1 gives the  $L^1$  traveltime error, the order of accuracy, and computing time of each of the first-order and new FSM. Here the  $L^1$  error is the mean value of the traveltime differences between numerical solution and analytic solution in the domain  $[5915 \text{ km}, 6385 \text{ km}] \times [30.5^\circ, 49.5^\circ] \times [15.5^\circ, 39.5^\circ]$ . The order of accuracy at the  $i$ -th row is calculated by

$$\text{order} = \frac{\ln(\text{error}_i / \text{error}_{i-1})}{\ln(h_i / h_{i-1})}, \quad (59)$$

in which  $\text{error}_i$  and  $h_i$  are the  $L^1$  traveltime error and grid spacing at the  $i$ -th row, respectively. According to this table, the first-order FSM only has a polluted first-order accuracy due to the point source singularity. In comparison, the new FSM achieves a second-order accuracy since the third-order WENO approximation and multiplicative factorization technique are employed. The  $L^1$  traveltime errors presented in Table 1 suggest that the new FSM is about three orders of magnitude more accurate than the first-order FSM with similar runtime. As a result, we can conclude that using the new FSM to calculate traveltime fields could be a better choice.

**Example 2. 3-D anisotropic media.** We set a bounded domain  $R \times \Theta \times \Phi = [5900 \text{ km}, 6400 \text{ km}] \times [30^\circ, 50^\circ] \times [15^\circ, 40^\circ]$ . The point source  $\mathbf{x}_s = (r_s, \theta_s, \phi_s) = (6150 \text{ km}, 40^\circ, 27.5^\circ)$  is located at the center of the domain. Let us consider the traveltime field in the medium with strong azimuthal anisotropy. The eikonal equation writes

## 12 Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media

**Table 1.** The numerical error and runtime of the first-order FSM and the new FSM for Example 1.

Mesh	First-order FSM			New FSM		
	$L^1$ error (s)	Order of accuracy	Computing time (s)	$L^1$ error (s)	Order of accuracy	Computing time (s)
$40 \times 40 \times 40$	1.34 E+1	-	0.92	5.08 E-2	-	2.33
$60 \times 60 \times 60$	1.02 E+1	0.6613	6.98	2.02 E-2	2.2770	15.34
$80 \times 80 \times 80$	8.44 E+0	0.6699	24.12	1.22 E-2	1.7606	70.02
$120 \times 120 \times 120$	6.29 E+0	0.7222	131.11	5.37 E-3	2.0165	417.75
$160 \times 160 \times 160$	5.08 E+0	0.7482	551.58	3.02 E-3	2.0000	1599.39

$$\begin{cases} \sqrt{e^{2W}(\partial_r T)^2 + \frac{e^{2W}}{r_s^2}(\partial_\theta T)^2 + \frac{2e^{2W}}{r_s^2}(\partial_\phi T)^2 - \frac{2e^{2W}}{r_s^2}(\partial_\theta T)(\partial_\phi T)} = s, & \mathbf{x} \in \Omega \setminus \{\mathbf{x}_s\}, \\ T(\mathbf{x}_s) = 0, \end{cases} \quad (60)$$

1 in which

$$W = \frac{1}{1000} \sqrt{(r - r_s)^2 + 2r_s^2(\theta - \theta_s)^2 + r_s^2(\phi - \phi_s)^2 + 2r_s^2(\theta - \theta_s)(\phi - \phi_s)}, \quad s = 0.2. \quad (61)$$

2 Then, the analytic solution can be given by

$$T_{ana}(r, \theta, \phi) = 200 \times (1 - \exp(-W)). \quad (62)$$

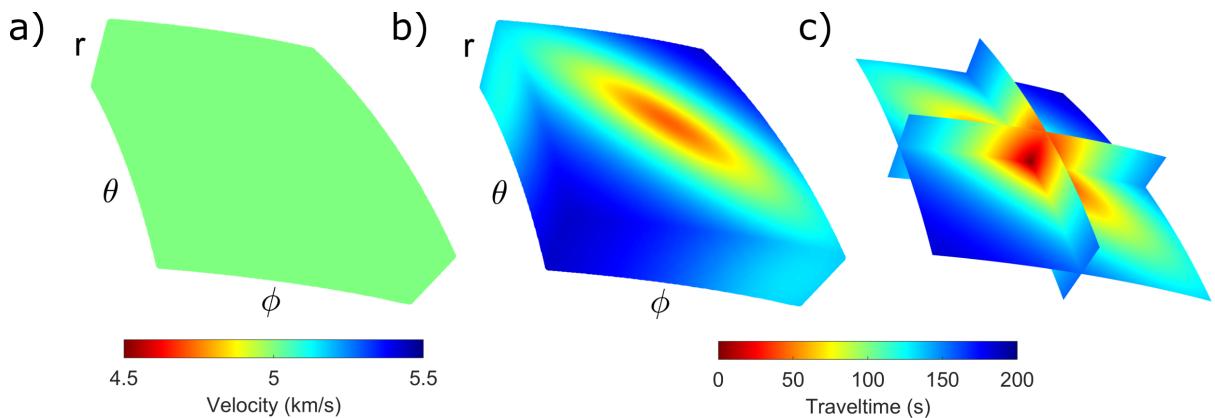
3 Though the wave speed is a constant (Figure 4a), the wave propagation shows a preferred direction due to the azimuthal anisotropy (see  
4 Figures 4b and 4c).

5 Similar to Example 1, we plot the solutions of the eikonal equation in horizontal and vertical sections crossing the point source in Figure  
6 5. The analytic solution (dashed black lines) coincides with the traveltimes field generated by the new FSM (dashed yellow lines) but is far  
7 from the result obtained by using the first-order FSM (solid blue lines). According to Table 2, we can also find that the new FSM captures the  
8 strong anisotropy and achieves a second-order accuracy. However, the first-order FSM only has a polluted first-order accuracy. Furthermore,  
9 the new FSM is about two orders of magnitude more accurate than the first-order FSM with similar runtime. Thus, we can draw the same  
10 conclusion as Example 1.

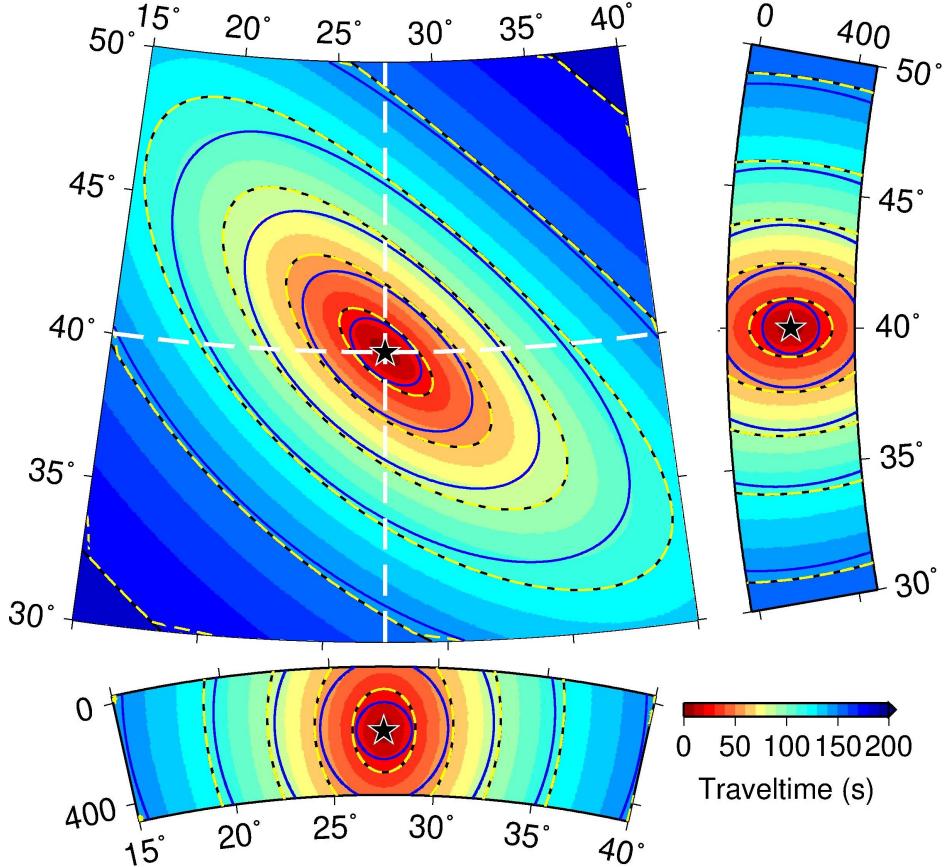
### 11 4.2 Synthetic checkerboard test

12 We perform a synthetic checkerboard test to evaluate the performance of the adjoint-state traveltomey method for azimuthally  
13 anisotropic media in spherical coordinates.

14 **Example 3. 3-D synthetic checkerboard test.** Consider the study region ranging from  $30^\circ\text{N}$  to  $50^\circ\text{N}$  in latitude and from  $15^\circ\text{E}$  to  $40^\circ\text{E}$  in  
15 longitude. The surface is at  $r = 6371$  km. The initial velocity model  $c(\mathbf{x}) = \frac{1}{s(\mathbf{x})}$  is formulated by



**Figure 4.** Illustration of the velocity model and traveltimes field in Example 2. (a) The 3-D velocity model  $c(\mathbf{x}) = \frac{1}{s(\mathbf{x})}$ . (b) The 3-D true traveltimes field  $T_{ana}(\mathbf{x})$ . (c) The horizontal and vertical sections of  $T_{ana}(\mathbf{x})$  crossing the point source  $\mathbf{x}_s$ .



**Figure 5.** The horizontal and vertical sections crossing the point source (black star) of the traveltimes calculated on a mesh of  $120 \times 120 \times 120$  grid points for Example 2, which correspond to the profiles in Figure 4c. Black and yellow dashed lines are the contours of the analytic solution  $T_{ana}(\mathbf{x})$  and the result generated by the new FSM, respectively. Solid blue lines represent the contours of the result obtained using the first-order FSM. Left top: the horizontal section at  $r = 6150$  km. The white dashed lines are the location of vertical sections; Left bottom: the vertical section at  $\theta = 40^\circ$ ; Right top: the vertical section at  $\phi = 27.5^\circ$ , "0" and "400" in the vertical axis correspond to  $r = 6371$  km and  $r = 5971$  km.

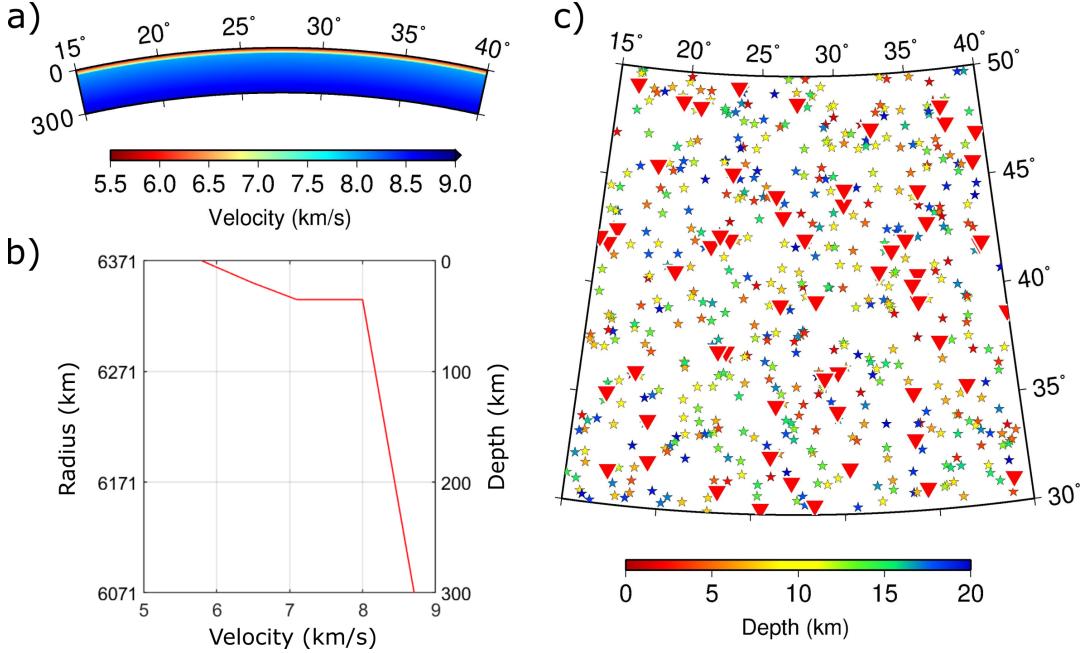
$$c(\mathbf{x}) = \frac{1}{s(r, \theta, \phi)} = \begin{cases} 5.8 + (6371 - r)/20 * 0.7 \text{ km/s} & \text{if } 6351 \text{ km} \leq r \\ 6.5 + (6351 - r)/15 * 0.6 \text{ km/s} & \text{if } 6336 \text{ km} \leq r < 6351 \text{ km}, \\ 8.0 + (6336 - r)/375 * 0.1 \text{ km/s} & \text{if } 5961 \text{ km} \leq r < 6336 \text{ km}, \\ 9.0 \text{ km/s} & \text{if } r < 5961 \text{ km}. \end{cases} \quad (63)$$

<sup>1</sup> This velocity model is modified from the AK135 global reference model (Kennett et al. 1995), including two crustal layers and a mantle layer. The Moho discontinuity is located at 35 km depth with a moderate velocity jump from 7.1 km/s to 8.0 km/s, (see Figures 6a and 6b)  
<sup>2</sup> The anisotropic parameters  $\xi(\mathbf{x})$  and  $\eta(\mathbf{x})$  are set 0 in the initial model. Furthermore, the checkerboard model is constructed by assigning  
<sup>3</sup> velocity and anisotropic perturbations to the initial model, that is,  $(s(\mathbf{x}) + \Delta s(\mathbf{x}), \Delta \eta(\mathbf{x}), \Delta \xi(\mathbf{x}))$ . The perturbations are  
<sup>4</sup>

$$\frac{\Delta s(\mathbf{x})}{s(\mathbf{x})} = -\frac{4\% \cdot \sigma(\mathbf{x})}{1 + 4\% \cdot \sigma(\mathbf{x})}, \quad \Delta \xi(\mathbf{x}) = 3\% \cdot |\sigma(\mathbf{x})| \cdot \cos(2\psi(\mathbf{x})), \quad \Delta \eta(\mathbf{x}) = 3\% \cdot |\sigma(\mathbf{x})| \cdot \sin(2\psi(\mathbf{x})), \quad (64)$$

**Table 2.** The numerical error and runtime of the first-order FSM and the new FSM for azimuthally anisotropic media in Example 2

Mesh	First-order FSM			New FSM		
	$L^1$ error (s)	Order of accuracy	Computing time (s)	$L^1$ error (s)	Order of accuracy	Computing time (s)
$40 \times 40 \times 40$	9.39 E+0	-	2.14	5.68 E-1	-	3.73
$60 \times 60 \times 60$	8.22 E+0	0.3282	14.98	2.64 E-1	1.8849	21.75
$80 \times 80 \times 80$	7.33 E+0	0.3976	49.70	1.58 E-1	1.8011	98.75
$120 \times 120 \times 120$	6.04 E+0	0.4753	295.36	7.28 E-2	1.9043	609.89
$160 \times 160 \times 160$	5.16 E+0	0.5466	1160.28	4.09 E-2	2.0039	2283.38



**Figure 6.** Illustration of the parameter setting in Example 3. (a) The vertical profile of the initial model. (b) The velocity with respect to the depth of the initial model. (c) The distribution of earthquakes and stations. Stars represent earthquakes, color-coded by depth. Stations are denoted by red inverted triangles.

in which

$$\sigma(\mathbf{x}) = \begin{cases} \sin\left(\frac{\theta-30^\circ}{50^\circ-30^\circ} \cdot 4\pi\right) \sin\left(\frac{\phi-15^\circ}{40^\circ-15^\circ} \cdot 4\pi\right) \sin\left(\frac{r-6211}{6371-6211} \cdot 2\pi\right), & \text{if } 6211 \text{ km} < r < 6371 \text{ km}, \\ 0, & \text{otherwise,} \end{cases} \quad (65)$$

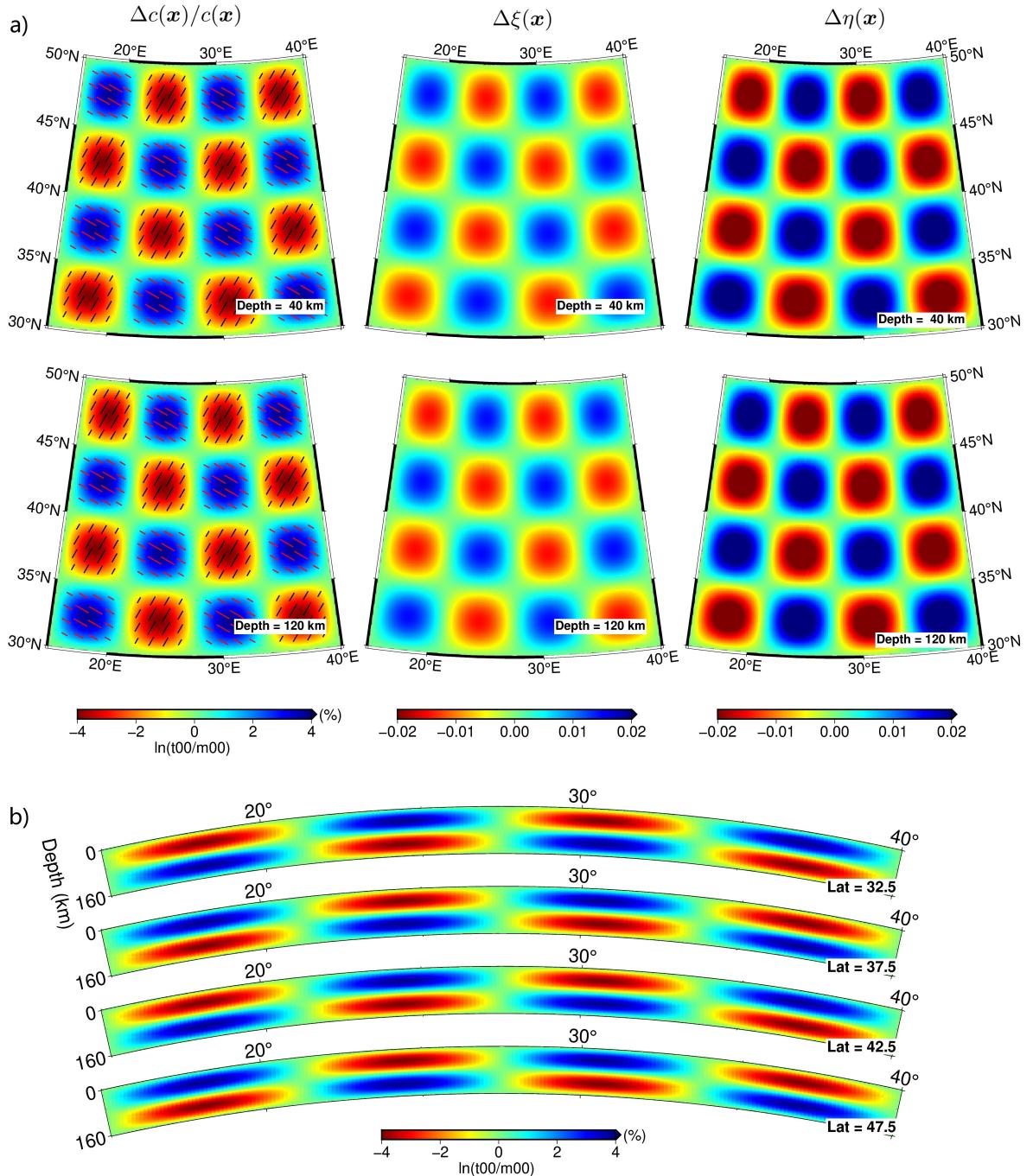
$$\psi(\mathbf{x}) = \begin{cases} 150^\circ, & \text{if } \sigma(\mathbf{x}) > 0, \\ 60^\circ, & \text{if } \sigma(\mathbf{x}) \leq 0. \end{cases} \quad (66)$$

Figure 7 illustrates the checkerboard model in horizontal and vertical sections. We aim to recover it starting with the initial model.

There are 60 seismic stations randomly distributed on the surface. Meanwhile, 500 earthquakes are set in the upper and middle crust, randomly located in the domain of  $[1 \text{ km}, 20 \text{ km}] \times [30^\circ, 50^\circ] \times [15^\circ, 40^\circ]$ . Figure 6c shows the distribution of seismic stations and earthquakes. In the synthetic checkerboard test, the observed traveltimes are calculated using the new FSM method in the checkerboard model. It is worth noting that we can regard stations as sources and earthquakes as receivers in the inversion by using the reciprocity principle (Aki & Richards 2002). This trick makes the computational cost of our method nearly independent of the number of earthquakes, which can lighten the computational burden especially for regional or large-scale tomography using hundreds of stations and tens of thousands of earthquakes.

Figure 8 plots the individual kernels corresponding to one station and four earthquakes, which have finite widths due to the grid smoothing when solving the adjoint equation on a mesh with finite grid intervals. Theoretically, the individual sensitivity kernel should be a curve with infinitesimal width, representing the transportation of traveltime misfit  $T_n(\mathbf{x}_{r,m}) - T_{n,m}^{obs}$  backward from the receiver to the source along the negative gradient direction of the traveltime field. Five inversion grids, placed in a staggered way along the diagonal direction, are designed to discretize model parameters (see Figure S2). We perform 100 iterations and plot the misfit reduction curve and misfit distribution in Figure S3. The objective function that measures the traveltime misfits of 30,000 source-receiver pairs sharply decreases at the beginning and experiences oscillation until the 40th iteration. The misfit reduction indicates the model convergence to a minimum, and the oscillation possibly results from the model update with moderate step lengths near the minimum. After that, with the dynamic step length decreasing, the inversion and objective function become steady, generating the final output model shown in Figure 9. During the 100 iterations, the objective function decreases from  $34,002 \text{ s}^2$  to  $72 \text{ s}^2$ , which suggests a perfect data fit. The velocity and anisotropy perturbations from 0 km to 80 km depth are well recovered (Figure 9). It suggests that this adjoint-state traveltimes tomography method is capable of illuminating the azimuthally anisotropic media in the crust and uppermost mantle. However, from 80 km to 160 km depth, only the velocity and azimuthal anisotropy associated with high-velocity blocks are satisfactorily recovered, while the low-velocity anomalies and corresponding anisotropies are not fully resolved.

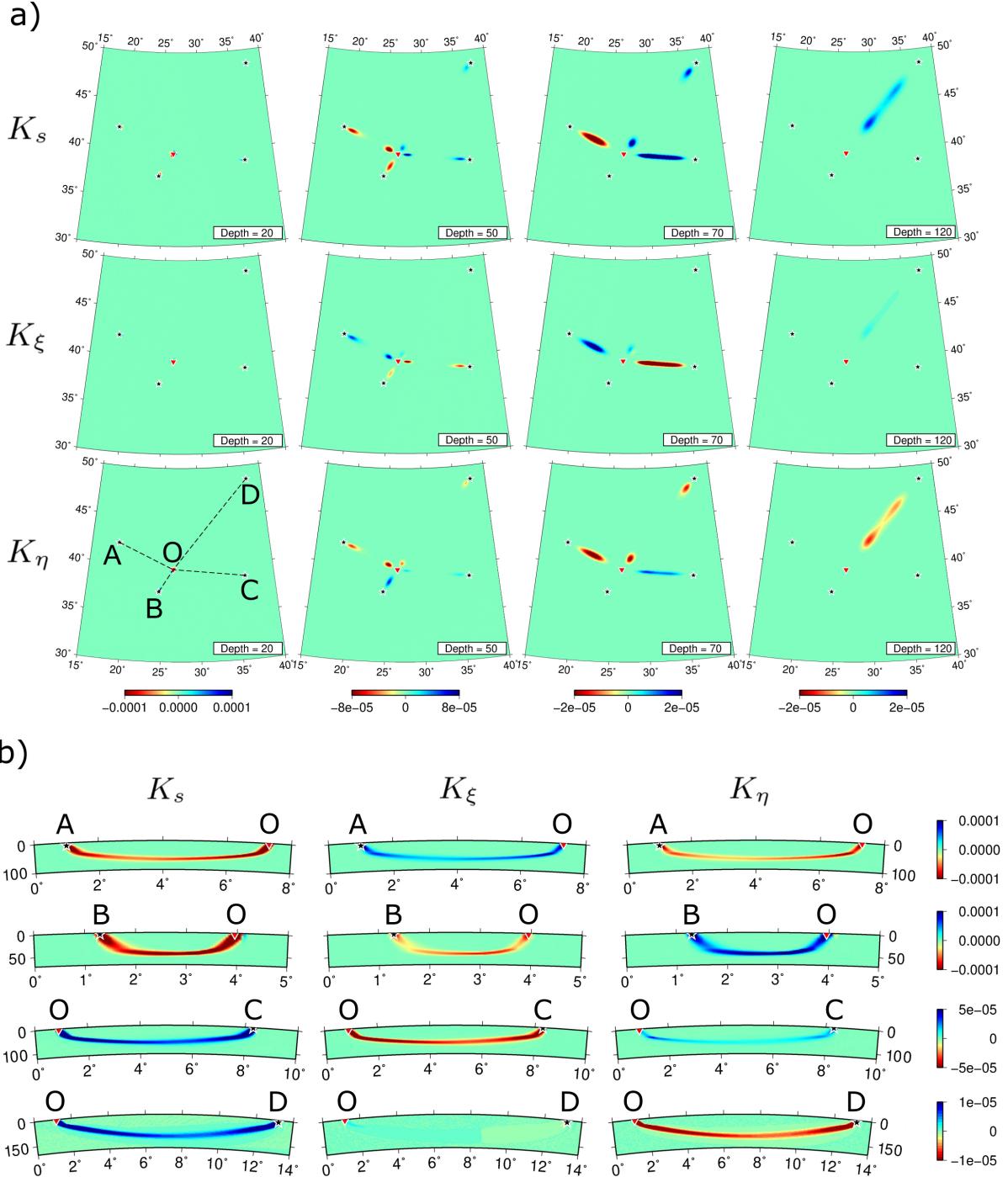
The poor resolution for the low-velocity anomalies below 80 km depth may result from the less sensitivity of traveltimes to these regions because seismic waves prefer to travel through high-velocity regions. With the synthetic model converging to the checkerboard, more waves concentrate on high-velocity regions (Figure S4). Thus, the high-velocity perturbations are well recovered. By contrast, fewer and fewer first



**Figure 7.** Horizontal and vertical sections of the checkerboard model. (a) The horizontal sections of the perturbation model of velocity  $\Delta c/c$  and anisotropic parameters  $\Delta \xi$  and  $\Delta \eta$ . Small solid lines in the first column align with the fast velocity directions, but are only shown when the magnitude of azimuthal anisotropy is greater than 0.005. (b) The vertical sections of the velocity perturbation model at different latitudes.

1 seismic arrivals cross the low-velocity zones (Figure S4). It implies that the low-velocity anomalies below 80 km depth has less effect on the  
2 objective function. As a result, low-velocity bodies are not adequately constrained and exhibit as low amplitude perturbations.

3 To evaluate the influence of picking errors on the imaging result, we perform the synthetic checkerboard test again with the same pa-  
4 rameter setting but assign random deviations obeying the Gaussian distribution with mean 0 s and standard deviation 0.05 s to the ‘observed’  
5 traveltimes (Figure S5a). Adding the data noise mimics the realistic traveltime picking errors. The inversion result is almost the same as the  
6 result of the checkboard test without noise (Figure S6). The main difference is the traveltimes misfits computed in the final output model  
7 (Figures S5b and S5c), whose standard deviation (0.068 s) is slightly greater than the standard deviation (0.049 s) of the misfits without  
8 noise. This synthetic test may suggest that the large volume of data can efficiently mitigate the influence caused by traveltime picking errors.

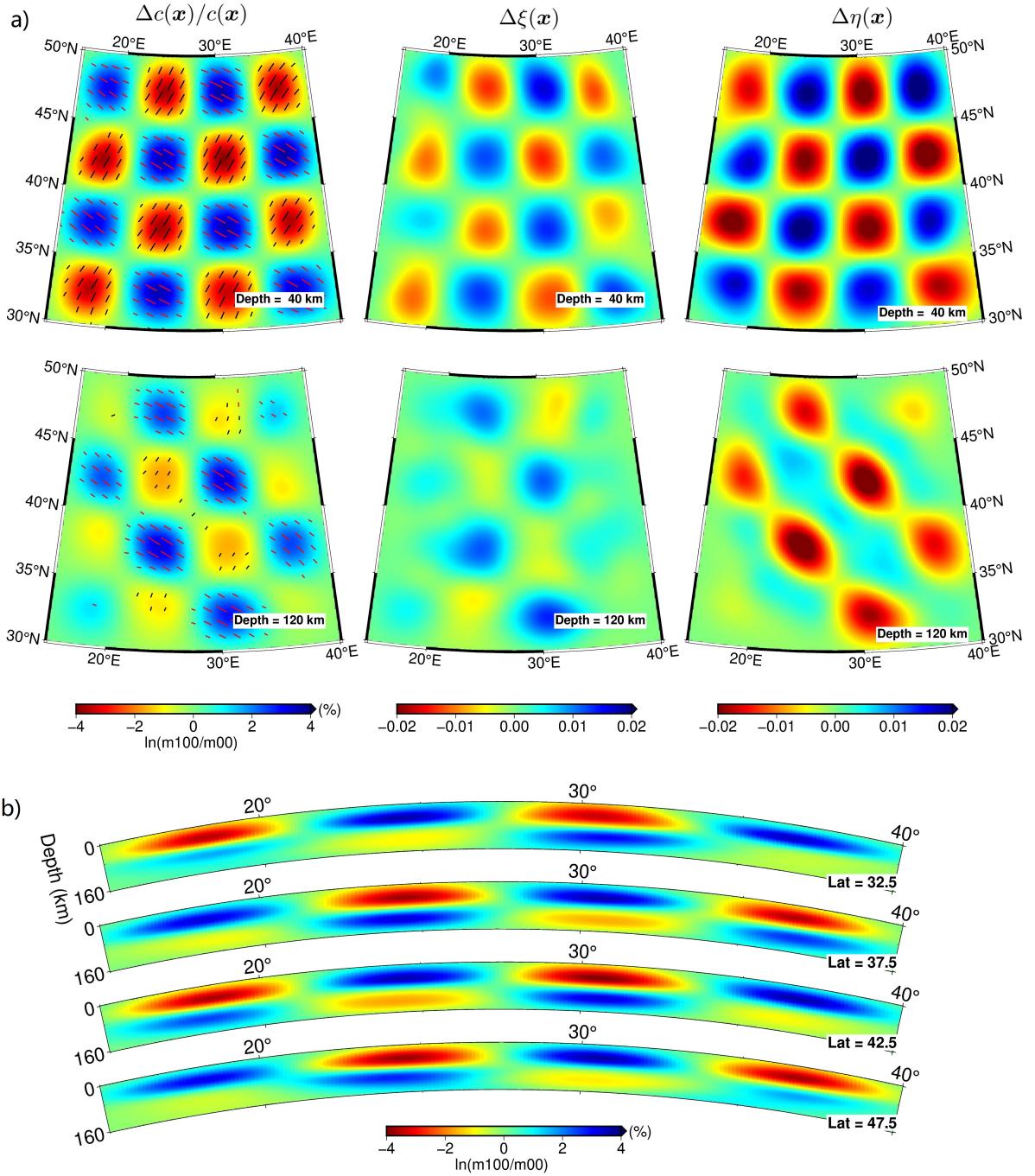


**Figure 8.** Horizontal and vertical sections of the sensitivity kernel of the objective function with respect to slowness and azimuthally anisotropic parameters. (a) The horizontal section of the sensitivity kernel corresponding to one station (red inverted triangle) and four earthquakes (black stars). (b) The vertical sections of individual kernels from the earthquake to the station. The locations of profiles are denoted by dashed lines in Figure 8a. In the horizontal axis, one great-circle degree equals 111.12 km.

## 5 APPLICATION IN CENTRAL CALIFORNIA NEAR PARKFIELD

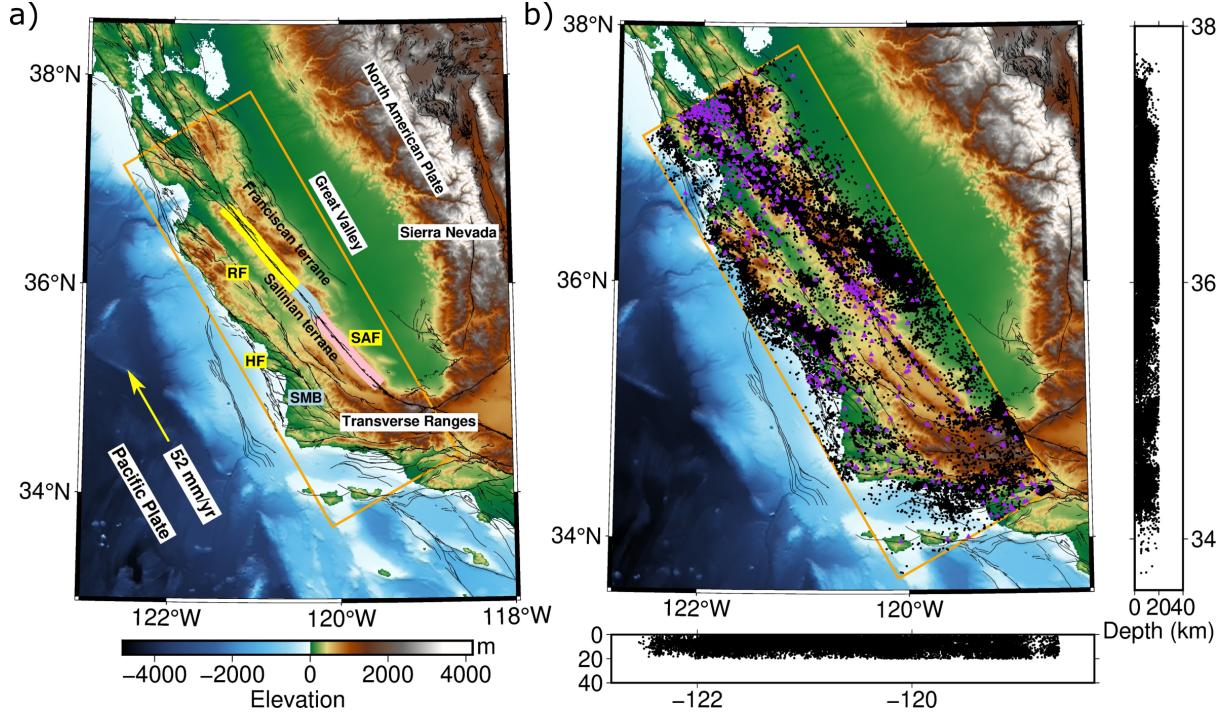
### 5.1 Background and motivation

The main purpose of the present practical tomographic inversion is to benchmark the performance of our tomography method in spherical coordinates against its counterpart in Cartesian coordinates. Same as Tong (2021), we choose the region of central California near Parkfield as the test field, where pervasive crustal azimuthal anisotropy and strong high- and low-velocity anomalies are widely observed. As an important feature in this region, the San Andreas Fault (SAF) separates the northeastern Franciscan Complex from the southeastern



**Figure 9.** Horizontal and vertical sections of the inversion result of checkerboard test in Example 3 without data noise. All symbols are the same as Figure 7.

1 Salinian terrane (Audet 2015), and is also regarded as the boundary between the oceanic Pacific Plate and the North American Plate (Porter  
2 et al. 2011). According to the surface creeping rate, the SAF is divided into three segments: the creeping part in the northwest, the locked  
3 section in the southeast, and the Parkfield segment as the transition in between (Titus et al. 2006; Ozacar & Zandt 2009). During the past  
4 forty years, quantities of earthquakes are recorded by the seismic stations in this region. These massive seismic data enable the illumination  
5 of crustal azimuthal anisotropy near the SAF (Audet 2015; Ozacar & Zandt 2009; Porter et al. 2011; Boness & Zoback 2006; Yang et al.  
6 2011) and the strong velocity contrast laterally across the SAF (Eberhart-Phillips & Michael 1993; Thurber et al. 2006; Zeng et al. 2016;  
7 Lippoldt et al. 2017). In this study, we invert first P-wave traveltimes to determine the anisotropic velocity structure in central California near  
8 Parkfield, which can verify our method and also provide constraints on crustal azimuthal anisotropy from P-waves.



**Figure 10.** (a) Tectonic setting and topography in central California. Our study region is within the origin box. Active faults are shown as black curves, some of which are labeled as HF (the offshore Hosgri Fault), RF (the Rinconada Fault), and SAF (San Andreas Fault). According to the surface creeping rates, the San Andreas Fault is divided into three segments: the creeping segment in yellow, the Parkfield segment as the transition zone in blue, and the locked segment in pink (Titus et al. 2006; Lippoldt et al. 2017). SMB is short for Santa Maria Basin. The yellow arrow indicates the direction of relative movement between the Pacific Plate and the North American Plate. (b) Distributions of the earthquakes (black dots) and seismic stations (purple triangles) used in the inversion.

## 5.2 Data and Initial model

We select the oblique study region within the orange lines in Figure 10 and rotate it to a regular domain of  $[0^\circ, 4.0^\circ] \times [0^\circ, 1.4^\circ]$  for the purpose of reducing computational cost. The first P-wave traveltimes and associated seismic stations and earthquakes with magnitude from 1.5 to 7.0 that occurred from January 1980 to July 2022 are downloaded from the Northern California Earthquake Data Center (NCEDC 2014) and Southern California Earthquake Data Center (SCEDC 2013). To ensure the reliability of the generated tomographic results, we apply strict data filtering criteria to select high-quality traveltimes for inversion. First, we only retain earthquakes with focal depths from 1 km to 20 km to mitigate the influence of Moho topography uncertainty on traveltimes. Then, earthquakes and stations with records less than 4 are discarded. Moreover, to avoid event clustering, we divide the study region into many subdomains with sizes of  $0.01^\circ \times 0.01^\circ \times 0.5$  km. Only one earthquake with the maximum records is selected in each subdomain. As a consequence, a total of 1,222,946 first P-wave traveltimes originating from 33,051 earthquakes and recorded by 555 stations are selected to reveal the P-wave velocity heterogeneity and azimuthal anisotropies in our study region (see Figure 10b).

We begin the inversion with the isotropic layered model suggested by Tong (2021), which has layers separated by two horizontal discontinuities at 4 km and 16 km and an undulating Moho discontinuity extracted from the Crust1.0 model (Laske et al. 2013). These four layers are namely sedimentary layer, upper crust, lower crust, and uppermost mantle which have constant velocities of 4.17, 6.03, 6.59, and 7.94 km/s. The Gaussian smoother with mean 0 km and standard deviation 0.4 km is further applied at the discontinuities. Besides, the initial anisotropic parameters  $\xi$  and  $\eta$  are set to be 0. The mean value of traveltimes misfits in this initial model is -0.101 s, which is close to 0, suggesting that the initial  $V_p$  is not systematically deviated upward or downward.

## 5.3 Inversion procedures and checkerboard resolution test

We first conduct the checkerboard resolution test to estimate the resolving ability of the selected data. The checkerboard model (Figure S7) is built by assigning staggered positive and negative  $V_p$  perturbations and varying azimuthal anisotropies to the initial model, given by

$$\frac{\Delta v(\mathbf{x})}{v(\mathbf{x})} = 8\% \cdot \sin(6\pi \frac{\theta}{4.0^\circ}) \sin(2\pi \frac{\phi}{1.4^\circ}) \sin(2\pi \frac{\sqrt{25+8z}-5}{8}), \quad (67)$$

$$\Delta\xi(\mathbf{x}) = 8\% \cdot |\sigma(\mathbf{x})| \cdot \cos(2\psi(\mathbf{x})), \quad \Delta\eta(\mathbf{x}) = 8\% \cdot |\sigma(\mathbf{x})| \cdot \sin(2\psi(\mathbf{x})), \quad (68)$$

in which

$$\sigma(\mathbf{x}) = \sin(2\pi \frac{\theta}{4.0^\circ}) \sin(4\pi \frac{\phi}{1.4^\circ}) \sin(2\pi \frac{\sqrt{25+8z}-5}{8}), \quad \psi(\mathbf{x}) = \begin{cases} 150^\circ, & \text{if } \sigma(\mathbf{x}) > 0, \\ 60^\circ, & \text{if } \sigma(\mathbf{x}) \leq 0. \end{cases} \quad (69)$$

To mimic the traveltime picking errors, we assign random deviations (Figure S8a) obeying the Gaussian distribution with mean 0 s and standard deviation 0.05 s to the predicted traveltimes computed in the checkerboard model to obtain the ‘observed traveltimes data’. Five sets of inversion grids are placed in a staggered way along the diagonal direction to update velocity and azimuthal anisotropy (Figure S9). We ensure that one wavelength anomaly is sampled by about five nodes of each inversion grid to avoid spatial aliasing (Tong 2021).

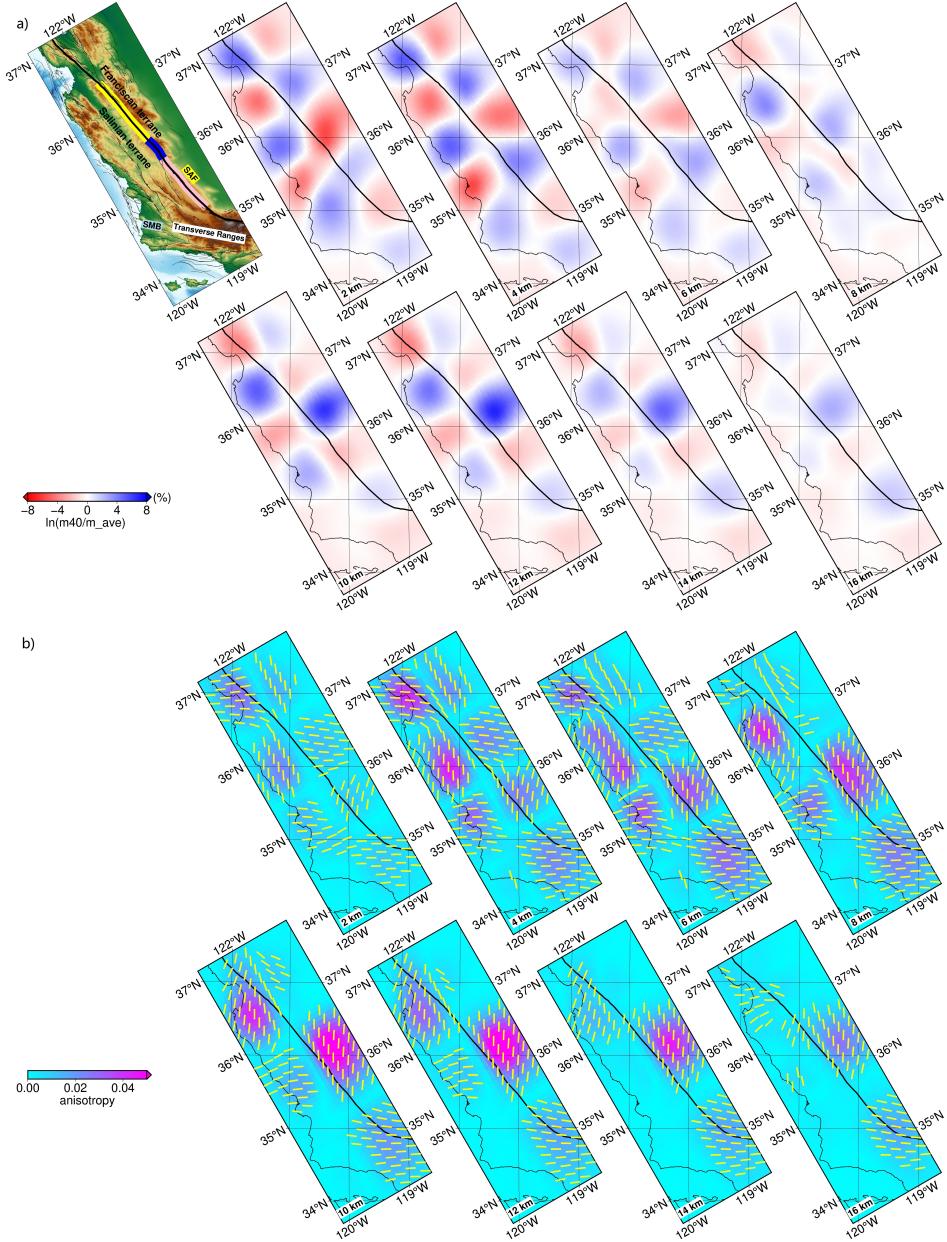
We execute the inversion procedure described in Section 2.3 and stop at the 40th iteration. The standard deviation of traveltime misfits decreases from 0.168 s to 0.075 s, suggesting a better fit for the traveltime data (see Figure S8b for the distributions of misfits). The horizontal sections of velocity perturbation with respect to the horizontal average are presented in Figure 11a. The checkerboard velocity anomalies are well recovered in the central region above 14 km depth. However, as a result of few seismic data in the south end of the study region, only the velocity anomaly in the southern termination above 8 km is recovered. The revealed azimuthal anisotropy shown in Figure 11b is similar. The magnitude of azimuthal anisotropy and the fast velocity directions are well recovered in the center. But the anisotropy is only partially recovered in the north end above 8 km depth and is totally not constrained in the southern termination due to the poor data coverage in those regions. Both the revealed velocity perturbation and azimuthal anisotropy have smaller magnitudes than those of the target model, possibly due to the smoothing effect imposed by the inversion grids. Overall, the geometry of staggered anomalies and the fast velocity directions are satisfactorily recovered, suggesting that the selected traveltimes data are capable of revealing velocity heterogeneity and azimuthal anisotropy in central California near Parkfield.

#### 5.4 Results and discussion

We perform the same inversion procedures as the checkerboard resolution test using the real traveltimes data in central California near Parkfield. The objective function reduces from 436,220 s<sup>2</sup> to 262,184 s<sup>2</sup> during 40 iterations (Figure S10). Figure 12a illustrates the velocity perturbation with respect to the horizontal average at different depths, revealing a strong velocity contrast laterally across the creeping segment and the Parkfield transition zone of the SAF in the upper crust down to 12 km depth. This velocity contrast perfectly matches the location of the SAF that separates the Franciscan terrane shown as a low-velocity zone from the Salinian terrane exhibited as a high-velocity body, which is consistent with the imaging results of previous studies (Eberhart-Phillips & Michael 1993; Thurber et al. 2006; Zeng et al. 2016; Lippoldt et al. 2017). In the northern termination at 37.2°, a high-velocity anomaly is revealed above 6 km depth in the east of the SAF. This high Vp body is also observed in the tomographic images of Liu and Tong (2021), which may indicate the relatively high-velocity Franciscan terrane compared with the low-velocity Great Valley Sequence (Liu & Tong 2021; Hole et al. 2000). In the middle west of the study region at 35°N, a low-velocity anomaly is revealed beneath the Santa Maria Basin, slightly extending northward from the surface to 12 km depth. Its shallow part correlates with the sediments of the Santa Maria Basin and offshore faults (Tong 2021). Its root is connected to a broad low-velocity anomaly beneath the creeping segment of the SAF at 14 km, which is believed as the continuation of the low-velocity zone in the upper crust beneath the Franciscan terrane (Lippoldt et al. 2017). Finally, a high-velocity zone is imaged in the southeast termination beneath the Transverse Range, showing a high agreement with the Vp model obtained by using both absolute and differential traveltimes (Lin et al. 2010).

The imaging results of the azimuthal anisotropy provide information on the mechanisms of pervasive crustal seismic anisotropy in central California near Parkfield. As illustrated in Figure 12b, strong azimuthal anisotropy is observed near the SAF in the upper and middle crust above 16 km depth, where the fast P-wave velocity direction aligns with the SAF. It suggests that the mechanism of seismic anisotropy near the SAF is mainly structure-induced. This observation is consistent with previous imaging results using teleseismic receiver functions analysis (Audet 2015; Ozacar & Zandt 2009). Apart from that, in the west and east of the creeping and the transition segments of the SAF at about 36°N, the fast velocity directions have high angles or even are perpendicular to the SAF. Rather than parallel to active faults, these fast velocity directions coincide well with the maximum horizontal compression ( $S_{H\max}$ ) directions obtained from borehole breakouts (Townend & Zoback 2004). Furthermore, the fast P-wave velocity directions above 8 km beneath the Transverse Ranges are also consistent with the  $S_{H\max}$  directions obtained from borehole breakouts and earthquake focal mechanism inversions (Townend & Zoback 2004). Thus, the crustal seismic anisotropies in those off-fault regions are mainly stress-induced.

Compared with Tong (2021), our study region is slightly enlarged on the north and south sides, with more earthquakes and seismic stations included in the inversion. The horizontal and vertical forward grid intervals are set to be 2 km and 0.5 km to ensure satisfactory numerical accuracy for traveltime prediction. The inversion takes a total of 1378 CPU hours for 40 iterations involving 555 receivers and 33051 sources. It generates almost the same Vp perturbation and azimuthal anisotropy as the results of Tong (2021) in the common region (Figures S11 and S12). The main difference in Vp perturbation is that our tomographic image at 2 km depth shows a clear separation of the low and high velocity zones by the SAF at 36.8°N. However, in the model of Tong (2021), a small portion of the eastern low-velocity zone crosses the SAF to the west side (red circle in Figure S11). The main reason is that the low-velocity zone is close to the boundary of the study region of Tong (2021) and may be distorted by the insufficient data coverage at that position. Besides, discrepancies in azimuthal anisotropy are observable on deep profiles at 14 km and 16 km (red circles in Figure S12), where azimuthal anisotropy is poorly constrained according to the checkerboard resolution test. Overall, our tomographic images show a high-level of similarity with the result of Tong (2021), indicating

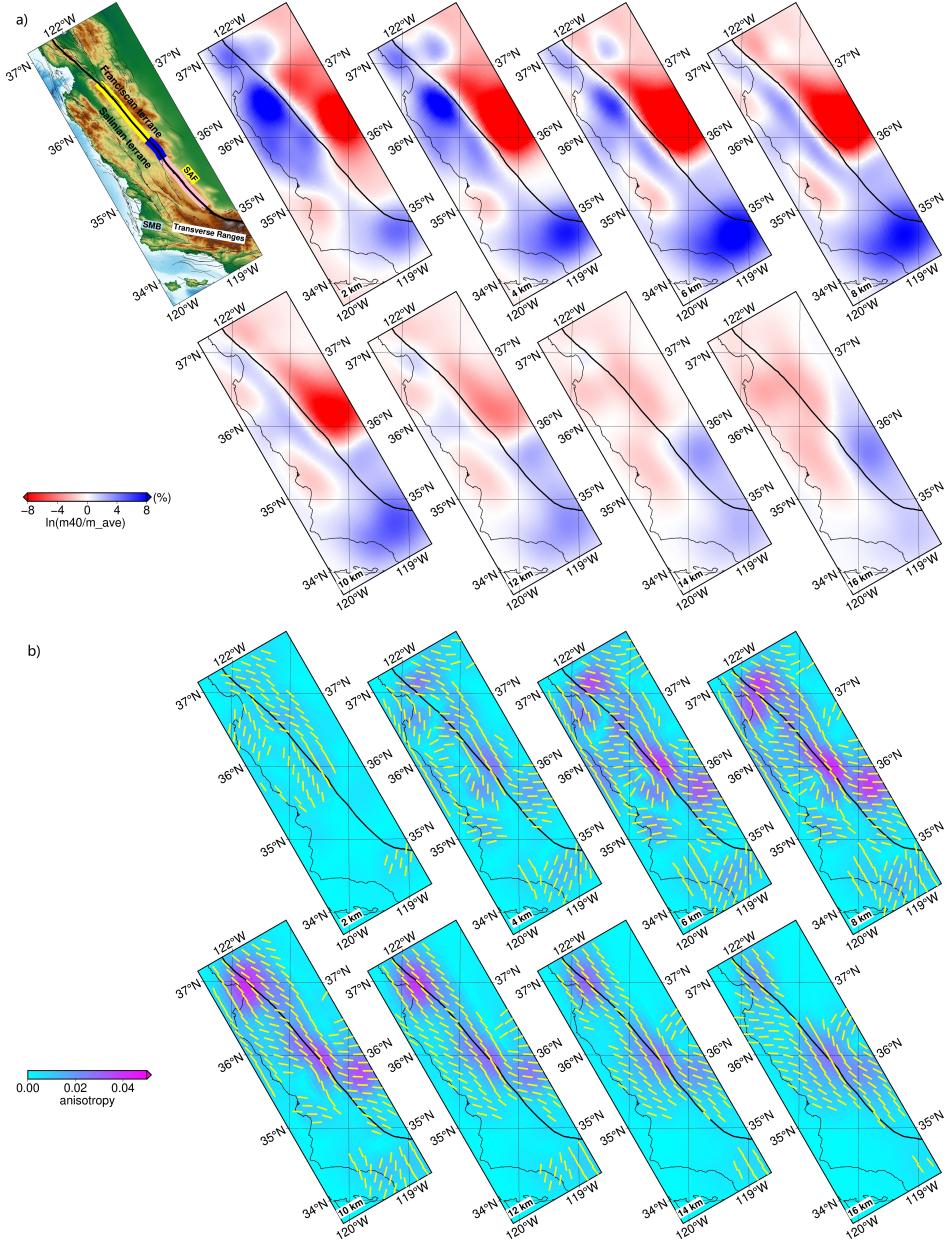


**Figure 11.** Horizontal sections of velocity perturbation and azimuthal anisotropy of the result in the checkerboard resolution test. The black solid curve is the San Andreas Fault. Many small yellow lines align with the fast velocity directions, but are only shown when the magnitude of azimuthal anisotropy is greater than 0.005.

that our method is effective to invert traveltimes for velocity heterogeneity and azimuthal anisotropy. Importantly, the new method is suitable for large-scale tomographic inversions.

## 6 CONCLUSIONS

We have developed the eikonal equation-based adjoint-state traveltime tomography method for azimuthal anisotropic media in spherical coordinates. Compared to the method of Tong (2021), this new method has two advantages. First, we consider the Earth's curvature and derive the eikonal equation for azimuthally anisotropic media in spherical coordinates. It enables us to describe the traveltime field in the Earth's volume accurately. Second, we use a modified fast sweeping method to solve the anisotropic eikonal equation. The multiplicative factorization technique and the third-order WENO approximation are employed to overcome the source singularity and raise the numerical accuracy. According to numerical experiments, this new fast sweeping method achieves a second-order accuracy. Its solution is about two orders of magnitude more accurate than the first-order fast sweeping method. These improvements render this new adjoint-state traveltime tomography method applicable to imaging velocity heterogeneity and azimuthal anisotropy of large-scale regions.



**Figure 12.** Horizontal sections of velocity perturbation and azimuthal anisotropy of the final output model. The black solid curve is the San Andreas Fault. Many small yellow lines align with the fast velocity directions, but are only shown when the magnitude of azimuthal anisotropy is greater than 0.005.

## 7 ACKNOWLEDGEMENTS

This work is funded by MOE AcRF Tier-2 Grant MOE2019-T2-2-112. PT is also partly supported by the National Research Foundation Singapore and the Singapore Ministry of Education under the Research Centers of Excellence Initiative (Project Code Number: 04MNS001953A620). The authors thank Hao Wu for valuable discussion. The authors would also appreciate Michal Malinowski, Tianshi Liu, and an anonymous reviewer for their constructive suggestions and remarkable comments, which are helpful to improve this article.

## 8 DATA AVAILABILITY STATEMENT

Traveltime data for this study were accessed through the Northern California Earthquake Data Center (NCEDC 2014) and the Southern California Earthquake Data Center (SCEDC 2013). Most figures are made with the Generic Mapping Tool (GMT) (Wessel & Smith 1991).

## 22 Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media

### 1 REFERENCES

- 2 Aki, K., & Richards, P. G., 2002. Quantitative seismology: Theory and methods (2nd ed.), University Science Books.
- 3 Alkhalifah, T., & Fomel, S., 2001. Implementing the fast marching eikonal solver: spherical versus Cartesian coordinates. *Geophys. Prospect.*, 49(2), 165-178.
- 4 Anderson, D. L., 2007. New theory of the Earth. Cambridge University Press.
- 5 Audet, P., 2015. Layered crustal anisotropy around the San Andreas fault near Parkfield, California. *J. Geophys. Res. Solid Earth*, 120(5), 3527-3543.
- 6 Boness, N. L., & Zoback, M. D., 2006. Mapping stress and structurally controlled crustal shear velocity anisotropy in California. *Geology*, 34(10), 825-828.
- 7 Chen, J., Chen, G., Wu, H., Yao, J., & Tong, P., 2022. Adjoint Tomography of Northeast Japan Revealed by Common-Source Double-Difference Travel-Time Data. *Seismol. Res. Lett.*, 93(3), 1835-1851.
- 8 Crandall, M. G., & Lions, P. L., 1983. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Am. Math. Soc.*, 277(1), 1-42.
- 9 Crandall, M. G., & Lions, P. L., 1984. Two approximations of solutions of Hamilton-Jacobi equations. *Mathematics of computation*, 43(167), 1-19.
- 10 Cui, N., Huang, G., Luo, S., Li, H., & Zhang, H., 2021. A hybrid fast sweeping method for the isotropic eikonal equation. *Explor. Geophys.*, 53(1), 104-115.
- 11 Detrixhe, M., & Gibou, F., 2016. Hybrid massively parallel fast sweeping method for static Hamilton-Jacobi equations. *J. Comput. Phys.*, 322, 199-223.
- 12 Detrixhe, M., Gibou, F., & Min, C., 2013. A parallel fast sweeping method for the Eikonal equation. *J. Comput. Phys.*, 237, 46-55.
- 13 Eberhart-Phillips, D., & Mark Henderson, C., 2004. Including anisotropy in 3-D velocity inversion and application to Marlborough, New Zealand. *Geophys. J. Int.*, 156(2), 237-254.
- 14 Eberhart-Phillips, D., & Michael, A. J., 1993. Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California. *J. of Geophys. Res. Solid Earth*, 98(B9), 15737-15758.
- 15 Fomel, S., Luo, S., & Zhao, H., 2009. Fast sweeping method for the factored eikonal equation. *J. Comput. Phys.*, 228(17), 6440-6455.
- 16 Guo, G., Lan, H., Zhou, X., Liu, Y., Bin Waheed, U., & Chen, J., 2022. Topography-dependent eikonal tomography based on the fast-sweeping scheme and the adjoint-state technique. *Geophysics*, 87(2), U29-U41.
- 17 Hole, J.A., Brocher, T.M., Klemperer, S.L., Parsons, T., Benz, H.M. & Furlong, K.P., 2000. Three-dimensional seismic velocity structure of the San Francisco Bay area. *J. geophys Res.*, 105(B6), 13859-13874.
- 18 Hole, J. A., & Zelt, B. C., 1995. 3-D finite-difference reflection traveltimes. *Geophys. J. Int.*, 121(2), 427-434.
- 19 Huang, Z., Zhao, D., & Liu, X., 2015. On the trade-off between seismic anisotropy and heterogeneity: Numerical simulations and application to Northeast Japan. *J. Geophys. Res. Solid Earth*, 120(5), 3255-3277.
- 20 Jiang, G. S., & Peng, D., 2000. Weighted ENO schemes for Hamilton-Jacobi equations. *SIAM J. Sci. Comput.*, 21(6), 2126-2143.
- 21 Jiang, G. S., & Shu, C. W., 1996. Efficient implementation of weighted ENO schemes. *textit{J. Comput. Phys.}*, 126(1), 202-228.
- 22 Kao, C. Y., Osher, S., & Qian, J., 2004. Lax-Friedrichs sweeping scheme for static Hamilton-Jacobi equations. *J. Comput. Phys.*, 196(1), 367-391.
- 23 Kennett, B. L., Engdahl, E. R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes. *Geophys. J. Int.*, 122(1), 108-124.
- 24 Kim, S., & Cook, R., 1999. 3-D travelttime computation using second-order ENO scheme. *Geophysics*, 64(6), 1867-1876.
- 25 Lan, H., & Chen, L. (2018). An upwind fast sweeping scheme for calculating seismic wave first-arrival travel times for models with an irregular free surface. *Geophys. Prospect.*, 66(2), 327-341.
- 26 Laske, G., Masters, G., Ma, Z., & Pasyanos, M., 2013. Update on CRUST1.0—A 1-degree global model of Earth's crust. *Geophysical research abstracts*, 15.
- 27 Leung, S., & Qian, J., 2006. An adjoint state method for three-dimensional transmission travelttime tomography using first-arrivals. *Comm. Math. Sci.*, 4(1), 249-266.
- 28 Lin, G., Thurber, C. H., Zhang, H., Hauksson, E., Shearer, P. M., Waldhauser, F., ... & Hardebeck, J., 2010. A California statewide three-dimensional seismic velocity model from both absolute and differential times. *Bulletin of the Seismological Society of America*, 100(1), 225-240.
- 29 Lippoldt, R., Porritt, R. W., & Sammis, C. G. (2017). Relating seismicity to the velocity structure of the San Andreas Fault near Parkfield, CA. *Geophysical Journal International*, 209(3), 1740-1745.
- 30 Liu, S., Suardi, I., Zheng, M., Yang, D., Huang, X., & Tong, P., 2019. Slab morphology beneath northern Sumatra revealed by regional and teleseismic travelttime tomography. *J. Geophys. Res. Solid Earth*, 124(10), 10544-10564.
- 31 Liu, Y., & Tong, P., 2021. Eikonal equation-based P-wave seismic azimuthal anisotropy tomography of the crustal structure beneath northern California. *Geophysical Journal International*, 226(1), 287-301.
- 32 Long, M. D., & Becker, T. W., 2010. Mantle dynamics and seismic anisotropy. *Earth Planet. Sci. Lett.*, 297(3-4), 341-354.
- 33 Luo, S., & Qian, J., 2011. Factored singularities and high-order Lax-Friedrichs sweeping schemes for point-source traveltimes and amplitudes. *textit{J. Comput. Phys.}*, 230(12), 4742-4755.
- 34 Luo, S., & Qian, J., 2012. Fast sweeping methods for factored anisotropic eikonal equations: multiplicative and additive factors. *J. Sci. Comput.*, 52(2), 360-382.
- 35 Luo, S., Qian, J., & Burridge, R., 2014. High-order factorization based high-order hybrid fast sweeping methods for point-source eikonal equations. *SIAM J. Numer. Anal.*, 52(1), 23-44.
- 36 Luo, Y., & Schuster, G. T., 1991. Wave-equation traveltime inversion. *Geophysics*, 56(5), 645-653.
- 37 NCEDC, 2014, *Northern California Earthquake Data Center*[Dataset]. UC Berkeley Seismological Laboratory, <https://doi.org/10.7932/NCEDC>
- 38 Osher, S., & Shu, C. W., 1991. High-order essentially nonoscillatory schemes for Hamilton-Jacobi equations. *SIAM J. Numer. Anal.*, 28(4), 907-922.
- 39 Ozacar, A. A., & Zandt, G., 2009. Crustal structure and seismic anisotropy near the San Andreas Fault at Parkfield, California. *Geophys. J. Int.*, 178(2), 1098-1104.
- 40 Porter, R., Zandt, G., & McQuarrie, N., 2011. Pervasive lower-crustal seismic anisotropy in Southern California: Evidence for underplated schists and active tectonics. *Lithosphere*, 3(3), 201-220.
- 41 Qian, J., & Symes, W. W., 2002. An adaptive finite-difference method for traveltimes and amplitudes. *Geophysics*, 67(1), 167-176.
- 42 Qian, J., Zhang, Y. T., & Zhao, H. K., 2007. A fast sweeping method for static convex Hamilton-Jacobi equations. *J. Sci. Comput.*, 31(1), 237-271.
- 43 Rawlinson, N., Hauser, J., & Cambridge, M., 2008. Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Adv. Geophys.*, 49, 203-273.
- 44 Rouy, E., & Tourin, A., 1992. A viscosity solutions approach to shape-from-shading. *SIAM J. Numer. Anal.*, 29(3), 867-884.
- 45 SCEDC, 2013, *Southern California Earthquake Data Center*[Dataset]. Caltech. <https://doi.org/10.7909/C3WD3xH1>
- 46 Shu, C. W., & Osher, S., 1988. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.*, 77(2), 439-471.
- 47 Thurber, C., Zhang, H., Waldhauser, F., Hardebeck, J., Michael, A., & Eberhart-Phillips, D., 2006. Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region. *Bulletin of the Seismological Society of America*, 96(4B), S38-S49.
- 48 Titus, S. J., DeMets, C., & Tikoff, B., 2006. Thirty-five-year creep rates for the creeping segment of the San Andreas fault and the effects of the 2004 Parkfield earthquake: Constraints from alignment arrays, continuous global positioning system, and creepmeters. *Bull. seism. Soc. Am.*, 96(4B), S250-S268.

- 1 Tong, P., 2021. Adjoint-State Traveltime Tomography for Azimuthally Anisotropic Media and Insight Into the Crustal Structure of Central California Near  
2 Parkfield. *J. Geophys. Res. Solid Earth*, 126(10), e2021JB022365.
- 3 Tong, P., Yang, D., & Huang, X., 2019. Multiple-grid model parametrization for seismic tomography with application to the San Jacinto fault zone. *Geophys.  
4 J. Int.*, 218(1), 200-223.
- 5 Tong, P., Chen, C. W., Komatitsch, D., Basini, P., & Liu, Q., 2014. High-resolution seismic array imaging based on an SEM-FK hybrid method. *Geophys. J.  
6 Int.*, 197(1), 369-395.
- 7 Townend, J., & Zoback, M. D., 2004. Regional tectonic stress near the San Andreas fault in central and southern California. *Geophysical Research Letters*,  
8 31(15), L15S11.
- 9 Tromp, J., Tape, C., & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys. J. Int.*, 160(1), 195-216.
- 10 Tsai, Y. H. R., Cheng, L. T., Osher, S., & Zhao, H. K., 2003. Fast sweeping algorithms for a class of Hamilton-Jacobi equations. *SIAM J. Numer. Anal.*, 41(2),  
11 673-694.
- 12 Vidale, J., 1988. Finite-difference calculation of travel times. *Bull. Seismol. Soc. Am.*, 78(6), 2062-2076.
- 13 Waheed, U. B., Yarman, C. E., & Flagg, G., 2015. An iterative, fast-sweeping-based eikonal solver for 3D tilted anisotropic media. *Geophysics*, 80(3), C49-  
14 C58.
- 15 Wessel, P., & Smith, W. H. F., 1991. Free software helps map and display data. *Eos, Transactions American Geophysical Union*, 72(1), 441-448.
- 16 Xiong, T., Zhang, M., Zhang, Y. T., & Shu, C. W., 2010. Fast sweeping fifth order WENO scheme for static Hamilton-Jacobi equations with accurate boundary  
17 treatment. *J. Sci. Comput.*, 45(1), 514-536.
- 18 Yang, Z., Sheehan, A., & Shearer, P. (2011). Stress-induced upper crustal anisotropy in southern California. *Journal of Geophysical Research: Solid Earth*,  
19 116(B2).
- 20 Yuan, Y. O., Simons, F. J., & Tromp, J., 2016. Double-difference adjoint seismic tomography. *Geophys. J. Int.*, 206(3), 1599-1618.
- 21 Zeng, X., Thurber, C. H., Shelly, D. R., Harrington, R. M., Cochran, E. S., Bennington, N. L., et al., 2016. 3-D P-and S-wave velocity structure and low-  
22 frequency earthquake locations in the Parkfield, California region. *Geophys. J. Int.*, 206(3), 1574-1585.
- 23 Zhang, L., Rector III, J. W., & Hoversten, G. M., 2005. Eikonal solver in the celerity domain. *Geophys. J. Int.*, 162(1), 1-8.
- 24 Zhang, Q., Ma, X., & Nie, Y., 2021. An iterative fast sweeping method for the eikonal equation in 2D anisotropic media on unstructured triangular meshes.  
25 *Geophysics*, 86(3), U49-U61.
- 26 Zhang, Y. T., Zhao, H. K., & Qian, J., 2006. High order fast sweeping methods for static Hamilton-Jacobi equations. *J. Sci. Comput.*, 29(1), 25-56.
- 27 Zhao, D., Yu, S., & Liu, X., 2016. Seismic anisotropy tomography: New insight into subduction dynamics. *Gondwana Research*, 33, 24-43.
- 28 Zhao, H., 2005. A fast sweeping method for eikonal equations. *Math. Comput.*, 74(250), 603-627.
- 29 Zhao, H., 2007. Parallel implementations of the fast sweeping method. *J. Comput. Math.*, 421-429.

## 30 APPENDIX A: MULTIPLICATIVE FACTORIZATION

We follow the idea proposed by Luo et al. (2014) and give a brief introduction to the multiplicative factorization technique for the anisotropic eikonal equation. We assume

$$t(\mathbf{x}) = T^2(\mathbf{x}), \quad S(\mathbf{x}) = s^2(\mathbf{x}), \quad \partial t(\mathbf{x}) = (\partial_r t, \partial_\theta t, \partial_\phi t)^t, \quad (\text{A.1})$$

Then, the eikonal equation in terms of  $t$  writes

$$[\partial t(\mathbf{x})]^t Q(\mathbf{x}) \partial t(\mathbf{x}) = [\nabla t(\mathbf{x})]^t M(\mathbf{x}) \nabla t(\mathbf{x}) = 4T(\mathbf{x})^2 [\nabla T(\mathbf{x})]^t M(\mathbf{x}) [\nabla T(\mathbf{x})] = 4t(\mathbf{x}) S(\mathbf{x}), \quad (\text{A.2})$$

in which

$$Q(\mathbf{x}) = \begin{pmatrix} 1 + 2\gamma\sqrt{\xi^2 + \eta^2} & 0 & 0 \\ 0 & \frac{1-2\xi}{r^2} & \frac{2\eta}{r^2 \cos \theta} \\ 0 & \frac{2\eta}{r^2 \cos \theta} & \frac{1+2\xi}{r^2 \cos^2 \theta} \end{pmatrix}. \quad (\text{A.3})$$

By using the Tayler expansion, we expand  $t$ ,  $S$ ,  $Q$  at the source  $\mathbf{x}_s = (r_s, \theta_s, \tau_s)$  into power series

$$t(\mathbf{x}) = \sum_{u=0}^{\infty} t_u(\mathbf{x}), \quad S(\mathbf{x}) = \sum_{v=0}^{\infty} S_v(\mathbf{x}), \quad Q(\mathbf{x}) = \sum_{w=0}^{\infty} Q_w(\mathbf{x}), \quad (\text{A.4})$$

in which  $t_u(\mathbf{x})$ ,  $S_v(\mathbf{x})$ , and  $Q_w(\mathbf{x})$  have the following forms

$$t_u(\mathbf{x}) = t_u(r, \theta, \phi) = \sum_{i+j+k=u} \alpha_{i,j,k} (r - r_s)^i (\theta - \theta_s)^j (\phi - \phi_s)^k, \quad (\text{A.5})$$

$$S_v(\mathbf{x}) = S_v(r, \theta, \phi) = \sum_{i+j+k=v} \beta_{i,j,k} (r - r_s)^i (\theta - \theta_s)^j (\phi - \phi_s)^k, \quad (\text{A.6})$$

$$Q_w(\mathbf{x}) = Q_w(r, \theta, \phi) = \sum_{i+j+k=w} \gamma_{i,j,k} (r - r_s)^i (\theta - \theta_s)^j (\phi - \phi_s)^k. \quad (\text{A.7})$$

Combining Equations (A.2) and (A.4) yields

$$\left( \sum_{u=0}^{\infty} t_u(\mathbf{x}) \right) \left( \sum_{v=0}^{\infty} S_v(\mathbf{x}) \right) = \frac{1}{4} \left[ \sum_{u=0}^{\infty} \partial t_u(\mathbf{x}) \right]^t \left( \sum_{w=0}^{\infty} Q_w(\mathbf{x}) \right) \left( \sum_{u=0}^{\infty} \partial t_u(\mathbf{x}) \right). \quad (\text{A.8})$$

Now we calculate  $t_i$  ( $i = 0, 1, 2, \dots$ ) sequentially. According to Equation (A.5) and the point source condition  $T(\mathbf{x}_s) = 0$ , we have

$$t_0(\mathbf{x}) = \alpha_{0,0,0} = t(\mathbf{x}_s) = T^2(\mathbf{x}_s) = 0. \quad (\text{A.9})$$

## 24 Adjoint-state Traveltime Tomography for Azimuthal Anisotropic Media

We also have  $S_0(\mathbf{x}) = s^2(\mathbf{x}_s)$  and  $Q_0(\mathbf{x}) = Q(\mathbf{x}_s)$ , and Equation (A.8) rewrites

$$\left( \sum_{u=1}^{\infty} t_u(\mathbf{x}) \right) \left( s^2(\mathbf{x}_s) + \sum_{v=1}^{\infty} S_v(\mathbf{x}) \right) = \frac{1}{4} \left[ \sum_{u=1}^{\infty} \partial t_u(\mathbf{x}) \right]^t \left( Q(\mathbf{x}_s) + \sum_{w=1}^{\infty} Q_w(\mathbf{x}) \right) \left( \sum_{u=1}^{\infty} \partial t_u(\mathbf{x}) \right). \quad (\text{A.10})$$

Next, we compare the constant terms of both sides in Equation (A.10). Obviously, the constant term of the left side is 0 since  $t_0 = 0$ . Thus, the constant term of the right side should equal 0, yielding

$$0 = [\partial t_1]^t Q(\mathbf{x}_s) \partial t_1, \quad (\text{A.11})$$

which implies

$$t_1(\mathbf{x}) = 0. \quad (\text{A.12})$$

Again we can rewrite Equation (A.10) as

$$\left( \sum_{u=2}^{\infty} t_u(\mathbf{x}) \right) \left( s^2(\mathbf{x}_s) + \sum_{v=1}^{\infty} S_v(\mathbf{x}) \right) = \frac{1}{4} \left[ \sum_{u=2}^{\infty} \partial t_u(\mathbf{x}) \right]^t \left( Q(\mathbf{x}_s) + \sum_{w=1}^{\infty} Q_w(\mathbf{x}) \right) \left( \sum_{u=2}^{\infty} \partial t_u(\mathbf{x}) \right). \quad (\text{A.13})$$

Similarly, we consider the quadratic terms of both sides and obtain

$$t_2(\mathbf{x}) s^2(\mathbf{x}_s) = \frac{1}{4} [\partial t_2(\mathbf{x})]^t Q(\mathbf{x}_s) \partial t_2(\mathbf{x}). \quad (\text{A.14})$$

$t_2(\mathbf{x})$  has a quadratic form so that we assume

$$t_2(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_s)^t A(\mathbf{x} - \mathbf{x}_s), \quad \partial t_2(\mathbf{x}) = 2A(\mathbf{x} - \mathbf{x}_s), \quad (\text{A.15})$$

in which  $A$  is a symmetric matrix. Then, Equation (A.14) is

$$s^2(\mathbf{x}_s) (\mathbf{x} - \mathbf{x}_s)^t A (\mathbf{x} - \mathbf{x}_s) = (\mathbf{x} - \mathbf{x}_s)^t A^t Q(\mathbf{x}_s) A (\mathbf{x} - \mathbf{x}_s), \quad \forall \mathbf{x} \in \Omega. \quad (\text{A.16})$$

It means  $s^2(\mathbf{x}_s) A = A Q(\mathbf{x}_s) A$ . If we assume that  $A$  is invertible, we can get

$$A = s^2(\mathbf{x}_s) Q(\mathbf{x}_s)^{-1}. \quad (\text{A.17})$$

Eventually, we obtain

$$t_2(\mathbf{x}) = s^2(\mathbf{x}_s) (\mathbf{x} - \mathbf{x}_s)^t Q(\mathbf{x}_s)^{-1} (\mathbf{x} - \mathbf{x}_s). \quad (\text{A.18})$$

<sup>1</sup> Following the same manner, we can sequentially calculate  $t_i(\mathbf{x})$ ,  $i = 0, 1, 2, \dots$ .

The most important observation is that the truncated summation  $U_N(\mathbf{x}) = \sqrt{\sum_{u=0}^N t_u}$  could approximate  $T(\mathbf{x})$  with  $N$ -th order accuracy near the source. It means if we choose  $U_N(\mathbf{x})$  as the multiplicative factor  $U(\mathbf{x})$  in Equation (28), the other factor  $\tau(\mathbf{x})$  will be  $N$ -th order smooth near the source because

$$\begin{aligned} \tau &= \frac{T}{U_N} = 1 + \frac{T - U_N}{U_N} = 1 + \frac{T^2 - U_N^2}{U_N(T + U_N)} = 1 + \frac{t - \sum_{u=0}^N t_u}{U_N(T + U_N)} = 1 + \frac{\sum_{u=N+1}^{\infty} t_u}{U_N(T + U_N)} = \\ &1 + \frac{\sum_{u=N+1}^{\infty} t_u}{U_N(1 + U_N/T) \sqrt{\sum_{u=2}^{\infty} t_u}} = 1 + \frac{O(|\mathbf{x} - \mathbf{x}_s|^{N+1})}{O(|\mathbf{x} - \mathbf{x}_s|)} = 1 + O(|\mathbf{x} - \mathbf{x}_s|^N). \end{aligned} \quad (\text{A.19})$$

Thus, the source singularity can be eliminated by solving the factored eikonal equation, providing the condition to achieve the  $N$ -th order accuracy. In our study, we choose  $N = 2$ , and the factor  $U(\mathbf{x})$  is given by

$$U(\mathbf{x}) = U_2(\mathbf{x}) = \sqrt{t_0 + t_1 + t_2} = s(\mathbf{x}_s) \sqrt{(\mathbf{x} - \mathbf{x}_s)^t Q(\mathbf{x}_s)^{-1} (\mathbf{x} - \mathbf{x}_s)}. \quad (\text{A.20})$$

<sup>2</sup> A more specific formulation is presented in Equation (29).

### 3 APPENDIX B: SELECTION OF VISCOSITY PARAMETERS

For simplicity of notation, we write Equation (30) as

$$H_f(\tau, u, v, w) = \sqrt{aX^2 + bY^2 + cZ^2 - 2fYZ}, \quad (\text{B.1})$$

where

$$a = 1 + 2\xi, \quad b = \frac{1 - 2\xi}{r^2}, \quad c = \frac{1 + 2\xi}{r^2 \cos^2 \theta}, \quad f = \frac{-2\eta}{r^2 \cos \theta}, \quad (\text{B.2})$$

$$X = U_r \tau + uU, \quad Y = U_\theta \tau + vU, \quad Z = U_\theta \tau + wU. \quad (\text{B.3})$$

In the assumption that the anisotropy is moderate ( $1 > 4\xi^2 + 4\eta^2$ ), we have

$$bc - f^2 = \frac{1 - 4\xi^2 - 4\eta^2}{r^4 \cos^2 \theta} > 0. \quad (\text{B.4})$$

The absolute values of the partial derivatives of  $H_f(u, v, w)$  with respect to  $u, v, w$  satisfy

$$\left| \frac{\partial H_f}{\partial u} \right| = \frac{|aUX|}{\sqrt{aX^2 + bY^2 + cZ^2 - 2fYZ}} = \sqrt{a}U \frac{|\sqrt{a}X|}{\sqrt{aX^2 + (\sqrt{b}Y - \frac{f}{\sqrt{b}}Z)^2 + \frac{bc-f^2}{b}Z^2}} \leq \sqrt{a}U, \quad (\text{B.5})$$

$$\left| \frac{\partial H_f}{\partial v} \right| = \frac{|bYU - fZU|}{\sqrt{aX^2 + bY^2 + cZ^2 - 2fYZ}} = \sqrt{b}U \frac{|\sqrt{b}Y - \frac{f}{\sqrt{b}}Z|}{\sqrt{(\sqrt{b}Y - \frac{f}{\sqrt{b}}Z)^2 + \frac{bc-f^2}{b}Z^2 + aX^2}} \leq \sqrt{b}U, \quad (\text{B.6})$$

$$\left| \frac{\partial H_f}{\partial w} \right| = \frac{|cZU - fYU|}{\sqrt{aX^2 + bY^2 + cZ^2 - 2fYZ}} = \sqrt{c}U \frac{|\sqrt{c}Y - \frac{f}{\sqrt{c}}Z|}{\sqrt{(\sqrt{c}Z - \frac{f}{\sqrt{c}}Y)^2 + \frac{bc-f^2}{c}Y^2 + aX^2}} \leq \sqrt{c}U. \quad (\text{B.7})$$

Thus, we can simply select

$$\sigma_r = U(x)\sqrt{1+2\xi}, \quad \sigma_\theta = U(x)\sqrt{\frac{1-2\xi}{r^2}}, \quad \sigma_\phi = U(x)\sqrt{\frac{1+2\xi}{r^2 \cos^2 \theta}}, \quad (\text{B.8})$$

to satisfy Equation (32).