Geophysical Journal International

Geophys. J. Int. (2026) **244,** 1–12
Advance Access publication 2025 December 22
Research Paper

https://doi.org/10.1093/gji/ggaf524

Triple-difference surface-wave traveltime adjoint tomography

Qicheng Zeng ,[1,†] Fan-Chi Lin[1] and Victor C. Tsai [2]

[1] *Department of Geology and Geophysics, University of Utah, Salt Lake City, UT* 84112, *USA. E-mail: Qicheng.Zeng@utah.edu*
[2] *Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI* 02912, *USA*

Accepted 2025 December 15. Received 2025 November 3; in original form 2025 July 22

SUMMARY

Structural boundaries are often the features of most interest geologically, but imaging them can be difficult due to wavefield scattering and interference caused by the sharp velocity contrasts. One example of this is the apparent Rayleigh-wave anisotropy (1-psi anisotropy) that has been observed near major structural boundaries using seismic arrays. The cause of the apparent anisotropy is the interference between the incident surface wave and waves scattered from velocity discontinuities. In this study, we first investigate the sensitivity of apparent anisotropy measurements to lateral boundary sharpness through 2-D full waveform simulations. We demonstrate that 1-psi anisotropy can vary based on boundary sharpness, station spacing and period of surface waves. We show that a misfit defined using triple-difference traveltimes, that is, the difference in double-difference traveltimes between station pairs with opposite propagation directions, well characterizes the apparent anisotropy. The sensitivity kernel for this triple-difference misfit can be constructed using the adjoint method. We show that triple-difference traveltimes are mainly sensitive to velocity contrasts rather than absolute velocities, in contrast to double-difference traveltimes. With sensitivity kernels constructed, we demonstrate how triple-difference traveltimes can be combined with double-difference traveltimes into a tomography inversion. We show that by including triple-difference traveltimes, seismic inversions converge faster and resolve boundary and average structure better in early iterations, compared to using double-difference traveltimes alone. Recent advancements in dense array experiments could facilitate the application of this method to better delineate tectonic and basin structural boundaries.

Key words: Fourier analysis; Numerical modelling; Seismic anisotropy; Seismic tomography; Surface waves and free oscillations; Wave scattering and diffraction.

1 INTRODUCTION

Seismic tomography is a technique for imaging the interior of the Earth with seismic waves (Q. Liu & Y. J. Gu 2012; A. Fichtner et al. 2024). The development and application of seismic tomography has led to successful discoveries including deep subducting slabs, large low-shear-velocity provinces in the lower mantle and numerous oil-producing reservoirs (Woodhouse & Dziewonski 1984; Fukao *et al.* 2001; Panning & Romanowicz 2006; Cui *et al.* 2024). Images from such techniques, however, are inevitably smoothed due to regularization and inadequate data coverage. The resulting blurry, low-resolution tomographic images make it difficult to make robust geologic interpretations such as the distinguishing between thermal boundaries with gradual change of elastic properties and compositional boundaries with sharp changes (Helffrich *et al.* 1989; Artemieva 2009). It is thus of great interest to develop new tomographic methods to enhance resolution such that structural boundaries can be better resolved.

Rayleigh-wave apparent azimuthal anisotropy (1-psi anisotropy, with 360-degree azimuthal periodicity) has been observed near major structural boundaries (Lin *et al.* 2008; Mauberger *et al.* 2021) and may hold the key to improving tomographic resolution and investigating dynamic processes associated with lithospheric edges (Hutchings *et al.* 2025). Different from the 2-psi and 4-psi azimuthal variations (with 180-degree and 90-degree periodicity, respectively; Smith & Dahlen 1973; Montagner & Nataf 1986; Liu & Ritzwoller 2025), which originate from intrinsic anisotropy in the medium, 1-psi anisotropy cannot arise from true intrinsic anisotropy, and instead results from the interference between the incident wave and waves scattered by velocity perturbations in a heterogeneous yet isotropic medium (Stich & Morelli 2007; A. Fichtner *et al.* 2013; Zeng *et al.* 2024).

[†] *Present Address:* Department of Earth, Environmental and Planetary Sciences, Rice University, Houston, TX 77005, USA.

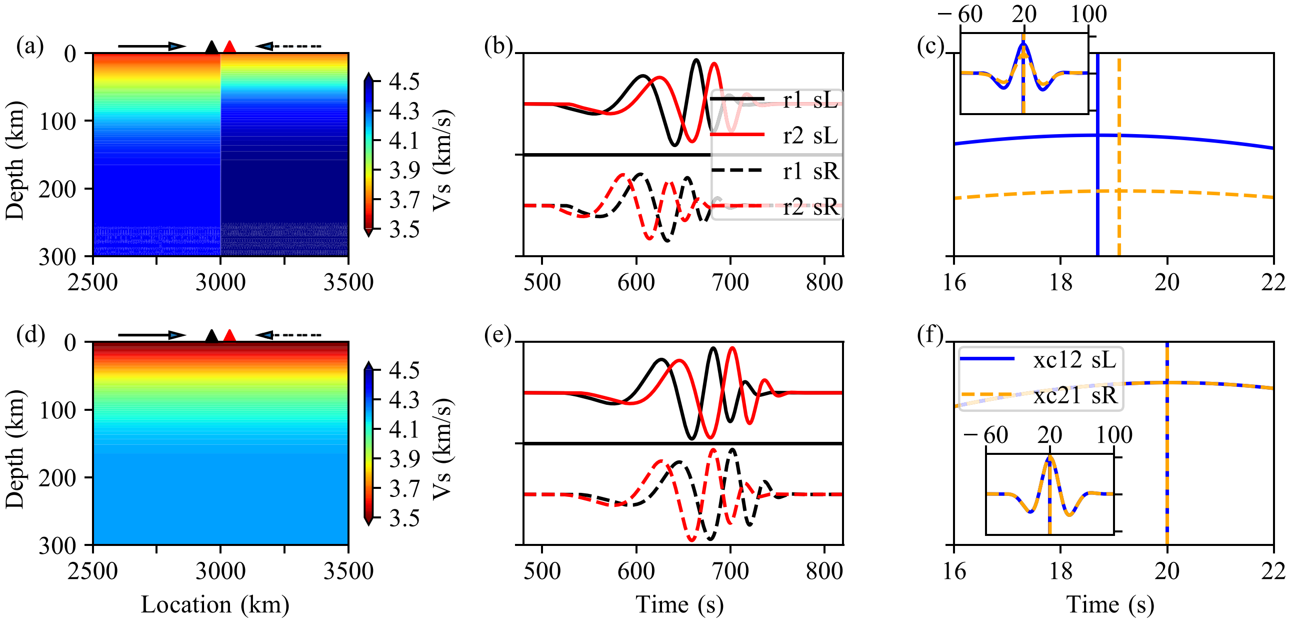

Figure 1. Demonstration of 1-psi anisotropy or triple-difference traveltime observations with respect to different boundaries. (a) Shear velocity model with a sharp boundary. Receivers are denoted by triangles. Sources from opposite direction are denoted by arrows. (b) Waveforms at two receivers from opposite source directions. Upper panel is from the left source (sL, solid lines), lower panel is from the right source (sR, dashed lines). Waveforms in black (r1) and red (r2) correspond to receivers in the same colour in (a). (c) Cross-correlations of waveforms from a common source. Vertical lines denote the time of the cross-correlation maxima. Inset plot is a zoom-out version. Time axes for cross-correlation from left source are reversed for comparison and visualization. (d) laterally homogeneous model. 1-D V_s is 95 per cent of the average of two 1-D V_s in (a). (c) Same as (b) but from model (d). (f) Same as (c) but from waveforms in (e).

The objective of this study is twofold: first, to investigate the precise sensitivity of the 1-psi anisotropy measurement; and secondly, to develop a method for incorporating this novel measurement into tomographic inversions, with the goal of enhancing resolution and more accurately delineating subsurface structural boundaries. To investigate the sensitivity, we first perform a series of 2-D full waveform simulations and determine how 1-psi anisotropy could vary based on boundary sharpness, station spacing and wave period. We then introduce a new triple-difference traveltime measurement, using the difference between two double-difference traveltimes for waves propagating in opposite directions, and construct finite frequency sensitivity kernel for this triple-difference misfit using the adjoint method (Tromp *et al.* 2005) by modifying a recent double difference adjoint framework (Yuan *et al.* 2016; Liu 2020; Pan *et al.* 2020; Tong *et al.* 2024) to incorporate the triple-difference traveltime misfit function. We demonstrate that the new double plus triple difference adjoint tomography has the advantage of resolving structural boundaries better and converging faster when compared to using double difference measurements alone.

2 FORWARD CALCULATIONS

Previous studies have suggested that the observation of 1 psi anisotropy could be affected by multiple factors including boundary sharpness, velocity contrast across the boundary, wave period, station spacing, and measurement method (Lin & Ritzwoller 2011a;b; Mauerberger *et al.* 2021; Zeng *et al.* 2024). To investigate the precise effects of these factors on 1 psi anisotropy, we perform 2-D full waveform simulations using different velocity models and measurement schemes.

2.1 Numerical setting

We perform 2-D forward wavefield simulations using the SPECFEM2D package (Tromp *et al.* 2008). The model domain is 6000 km × 300 km, with a uniform element size of 10 km × 10 km. We impose free surface boundary conditions on top and absorbing boundary conditions on the two sides and bottom. We extract two 1-D shear wave velocity (V_s) models, one from the Snake River Plain and one from the Idaho Batholith, from Schmandt *et al.* (2015) and use them to construct a 2-D model with a sharp velocity contrast (Fig. 1a). V_p and ρ are obtained from V_s using empirical relationships of Brocher (2005). As we focus on using 1-psi anisotropy to investigate lateral structure boundaries, the 2-D models are smoothed with $\sigma_z = 30$ km and $\sigma_x = 0.1$ m Gaussian kernels in the vertical and lateral directions, respectively. Two isotropic moment tensor sources with a dominant source frequency of 0.02 Hz are used to simulate 2-D wavefields propagating in opposite directions. The sources are located at 1 km depth, positioned on either side of the velocity boundary (i.e. x = 3000 km), and 2000 km away from it.

For each source, we perform a forward simulation and record waveforms for receivers on the surface. Example synthetic vertical Rayleigh-wave waveforms recorded by two receivers, each 35 km away from the structural boundary, on each side of the boundary, are shown in Fig. 1(b). For each source, we determine the differential phase traveltime between the two receivers based on the maximum amplitude of the waveform cross-correlation. Note that differences in the differential traveltimes are observed for waves propagating to the left versus to the right (Fig. 1c), hence the 1-psi apparent anisotropy. This observation is consistent with our previous results, in which we investigated the effect of interference between incident

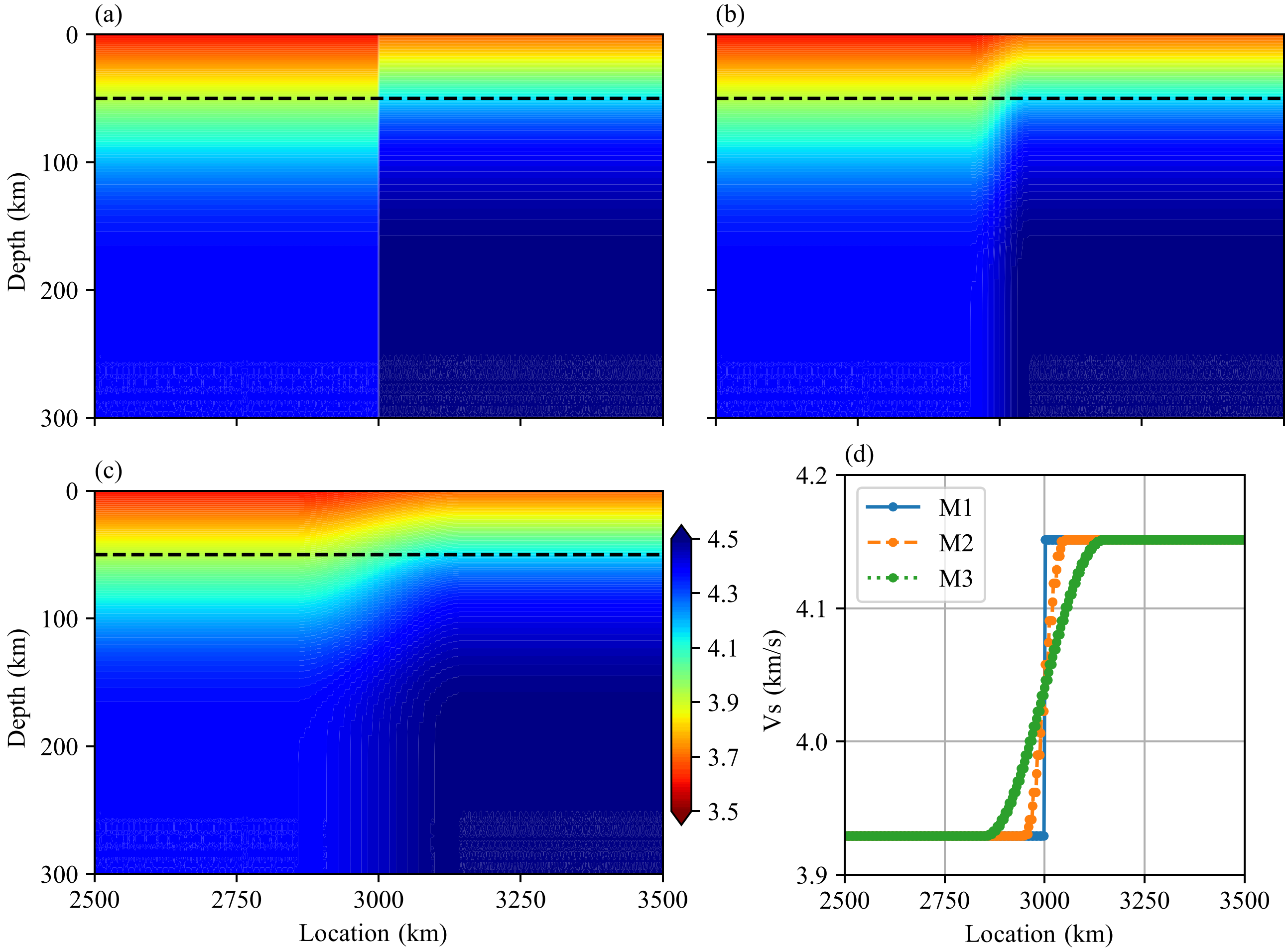

Figure 2. Velocity models with varying boundary widths. (a) Sharp boundary with 0 boundary width, same as Fig. 1(a). (b) Boundary width of 100 km. (c) Boundary width of 300 km. (d) Velocity depth slices at 50 km depth, M1-M3 corresponds to (a)–(c).

and scattered waves based on numerical simulations and normal mode summations (Zeng *et al.* 2024). For a laterally homogeneous model, the differential traveltimes for waves propagating in opposite directions are identical to one another as expected (Figs 1d–f).

2.2 Boundary sharpness and 1-psi anisotropy

Forward simulations based on laterally homogeneous and heterogeneous models, as shown in Fig. 1, demonstrate the causal relationship between structural boundaries and 1-psi anisotropy. To further understand how boundary sharpness affects the observation of 1 psi anisotropy, we perform simulations using velocity models with different boundary transition widths. Three representative velocity models are used, with 0, 100 and 300 km boundary transition widths d (Fig. 2d):

$$V_S(x,z) = V_{S1}(z) + (V_{S2}(z) - V_{S1}(z)) \times \frac{1-\cos\left[\left(\frac{x-x_0}{d}\right)\ \pi\right]}{2}$$
$$x_0 \le x \le x_0 + d, \tag{1}$$

where $V_S(x,z)$ is shear velocity within the boundary transition zone, $V_{S1}(z)$ and $V_{S2}(z)$ are shear velocities for the left and right 1-D models, x_0 denotes the horizontal location of the left end of the boundary. Shear velocities outside of the boundary transition zone remain the same as the left or right model.

In addition to testing transition width, we also test how 1-psi anisotropy varies with different periods (i.e. 30, 60 and 90 s) and different station spacings (i.e. 10, 30 and 70 km; Fig. 3). 1-psi anisotropy in 2-D can be described in terms of phase velocity using equation below,

$$A_1(x,\omega) = \frac{v_L(x,\omega) - v_R(x,\ \omega)}{(v_L(x,\omega) + v_R(x,\omega))/2}, \tag{2}$$

where $A_1(x,\omega)$ is the 1 psi amplitude at location x, ω is angular frequency, $v_L(x,\omega)$ and $v_R(x,\omega)$ are frequency-dependent apparent phase velocities at location x derived from the left and right sources, respectively. As phase velocities are measured using phase traveltimes, 1-psi anisotropy can also be described in terms of differential traveltime:

$$A_1(x,\omega) = \frac{\Delta t_R(x,\omega) - \Delta t_L(x,\omega)}{(\Delta t_L(x,\omega) + \Delta t_R(x,\omega))/2}, \tag{3}$$

where $\Delta t_L(x,\omega)$ and $\Delta t_R(x,\omega)$ are frequency-dependent apparent differential traveltimes between the two stations derived from the left and right sources, respectively.

Zeng *et al.* (2024) showed that 1-psi anisotropy is caused by interference between the incident Rayleigh wave and the backscattered Rayleigh wave (short wavelength oscillation) and body waves (long wavelength oscillation). Their result is consistent with the current study with 0 km transition width (Figs 3a–c). At 90 s period (Fig. 3c),

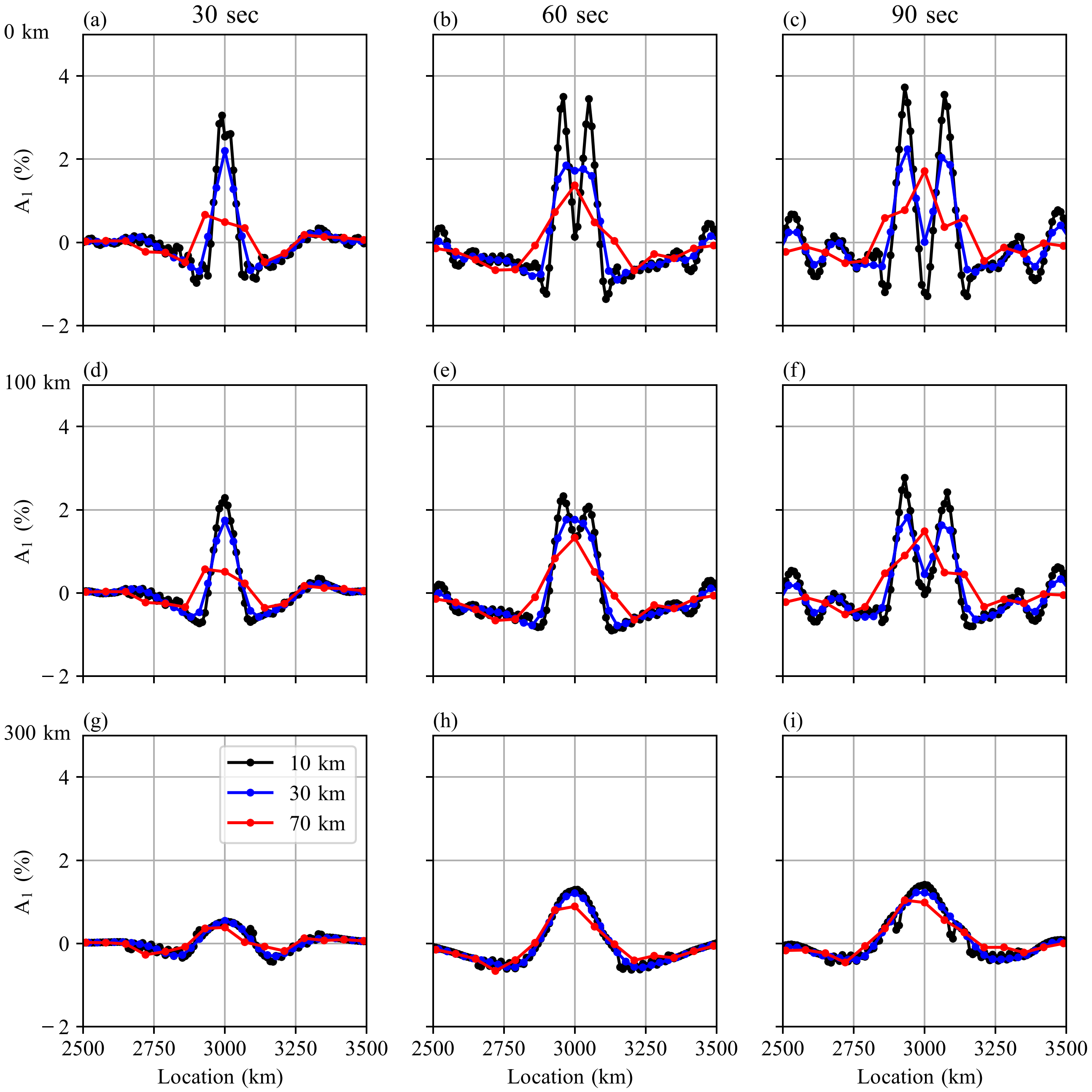

Figure 3. 1-psi anisotropy across the array with varying period and boundary width. The horizontal coordinate represents the midpoint of the station pair used for the measurement. (a) 1-psi anisotropy at 30 s period with model M1. (b) 60 s, M1. (c) 90 s, M2. (d), (e), (f) same as (a), (b), (c) but from M2. (g), (h), (i) same as above but from M3. Black lines are from 10 km station spacing. Blue lines are from 30 km station spacing. Red lines are from 70 km station spacing.

short-wavelength oscillations are observed for 10 and 30 km station spacing, which is the result of interference between incident and backscattered Rayleigh waves. Note that the 1-psi anisotropy results shown in Fig. 3 have been smoothed using a three-point moving average consistent with our prior approach to stabilize the measurements (Lin & Ritzwoller, 2011a; Zeng *et al.* 2024). Since the incident and backscattered waves share the same phase velocity, the wavelength of the 1-psi oscillation is about half of the Rayleigh-wave wavelength. Resolving this short wavelength oscillation therefore requires dense station spacing (i.e. station spacing smaller than 1/8 and 1/24 of the Rayleigh wave wavelength without and with the three-point smoothing, respectively). At 60 s period (Fig. 3b), short wavelength oscillations can only be observed with 10 km spacing. At 30 s period (Fig. 3a), no short wavelength oscillations can be observed even with 10 km station spacing. Whether or not short wavelength oscillations can be fully resolved, long-wavelength oscillations are apparent due to the interference between the incident Rayleigh wave and scattered body waves (Zeng *et al.* 2024).

When the boundary transition width is larger than half the Rayleigh-wave wavelength, the short wavelength oscillations become difficult to observe suggesting the backscattered Rayleigh wave is significantly diminished. The long wavelength 1-psi oscillation on the other hand is a more robust observation and less sensitive to the boundary sharpness and station spacing.

In general, the amplitude of 1-psi anisotropy decreases as the boundary transition width increases (Fig. 3). This is particularly

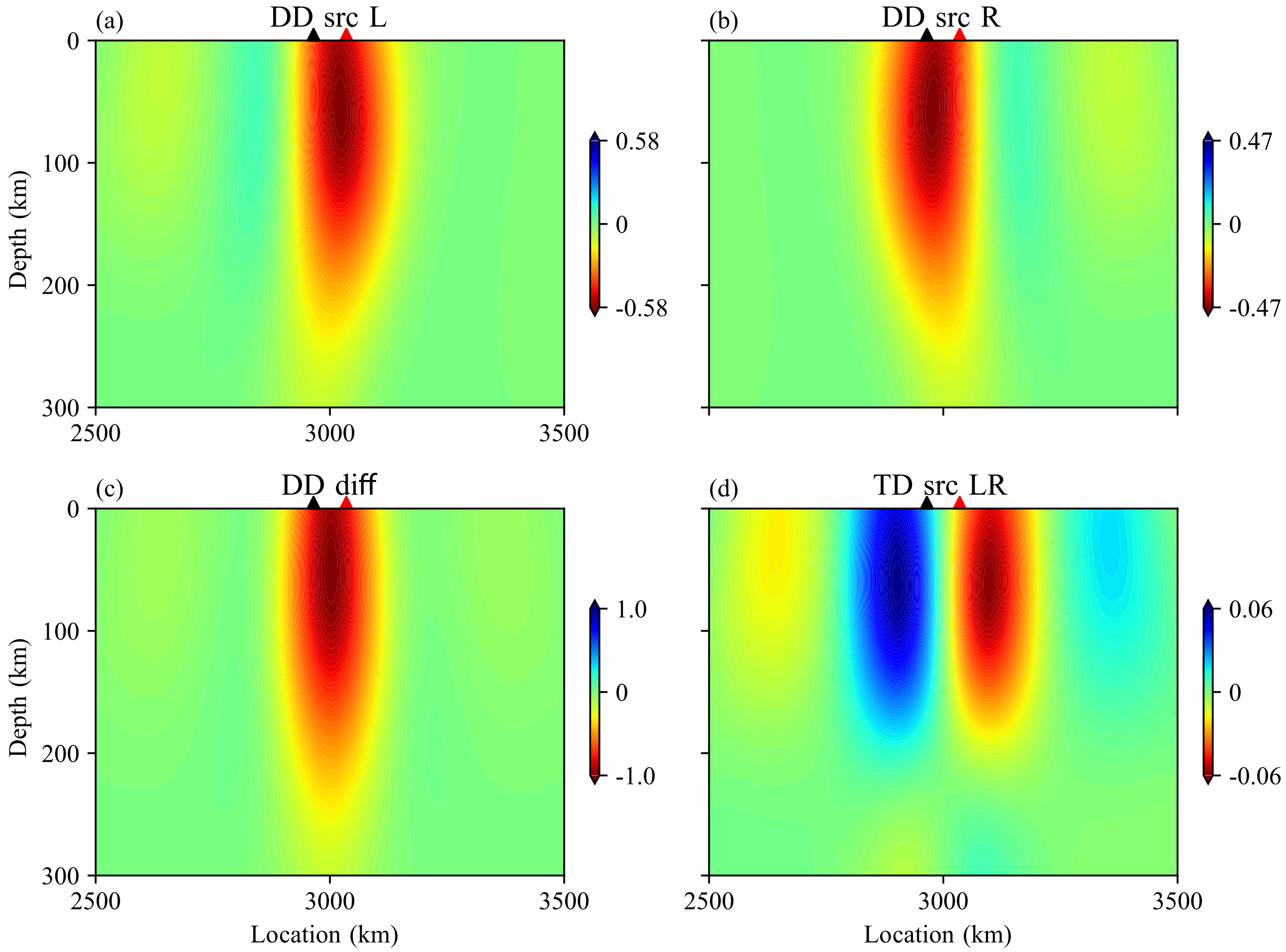

Figure 4. Sensitivity kernels of 2 stations with 70 km spacing filtered around 60 s (±20 per cent). Target model is Fig. 1(a), starting model is Fig. 1(d). (a) Double-difference sensitivity kernel from left source. (b) Double-difference sensitivity kernel from right source. (c) Sum of (a) and (b). (d) Triple-difference sensitivity kernel.

apparent for the short wavelength oscillation when the station spacing is small. This is consistent with the expectation that sharper structure boundaries produce stronger backscattering, where the degree of sharpness should be evaluated based on the ratio of the wavelength to the transition width (Stolyarov 1978). Most of the short-wavelength oscillations diminish when the transition width exceeds half the Rayleigh-wave wavelength. Note that for the same period, body waves have longer wavelengths and higher apparent velocities than Rayleigh waves, which might explain why the long wavelength 1-psi oscillation seems to be less affected by reducing boundary sharpness.

The result of our forward calculations suggests that the sharpness of a structure boundary can potentially be resolved when station spacing is sufficient to resolve the full 1-psi surface-wave backscattering interference pattern. In this case, a sharper boundary is expected to produce a larger 1-psi amplitude owing to stronger reflections. The exact reflection coefficient will depend on both the transition width to wavelength ratio and the impedance contrast of the two media (Benz & Vidale 1993; Aki & Richards 2002, pp. 128–149). We note that while wavelength is often considered a limiting factor for resolution in traditional ray-based traveltime tomography (Bolton and Masters, 2001; Montelli *et al.* 2004; Simmons *et al.* 2010), this is not necessarily the case for 1-psi measurements, which is finite frequency in nature (Dahlen *et al.* 2000; Hung et al. 2000; De Hoop and Van der Hilst 2005; Van der Hilst and De Hoop 2005).

3 SENSITIVITY KERNELS OF TRIPLE-DIFFERENCE TRAVELTIMES

In this section, we construct the equivalent 1-psi sensitivity kernel through the adjoint method (Tromp *et al.* 2005). As described in eq. (3), 1-psi anisotropy is equivalent to the ratio between the difference and the average of differential traveltimes between a station pair for waves propagating in opposite directions. As small changes in average velocity do not affect the value of 1-psi in a major way, we construct a new triple-difference misfit function χ^{td} for the adjoint method using the numerator, the difference in differential traveltimes, alone. For a simple case of two receivers and two sources from left and right:

$$\chi^{\mathrm{td}} = \frac{1}{2}\sum_{k}\left[\Delta\Delta\Delta t_{LR12}(\omega_k)\right]^2, \tag{4}$$

where

$$\begin{aligned}\Delta\Delta\Delta t_{LR12} = &\left(\left(t_{L2}^{\mathrm{syn}} - t_{L1}^{\mathrm{syn}}\right) - \left(t_{R1}^{\mathrm{syn}} - t_{R2}^{\mathrm{syn}}\right)\right)\\ &- \left(\left(t_{L2}^{\mathrm{obs}} - t_{L1}^{\mathrm{obs}}\right) - \left(t_{R1}^{\mathrm{obs}} - t_{R2}^{\mathrm{obs}}\right)\right),\end{aligned} \tag{5}$$

is the triple-difference traveltime at each angular frequency ω_k. The dependence of traveltime on frequency is assumed but omitted in eq. (5) and after for concision. In eq. (5), L and R represent left and right sources, 1 and 2 represent left and right receivers, and

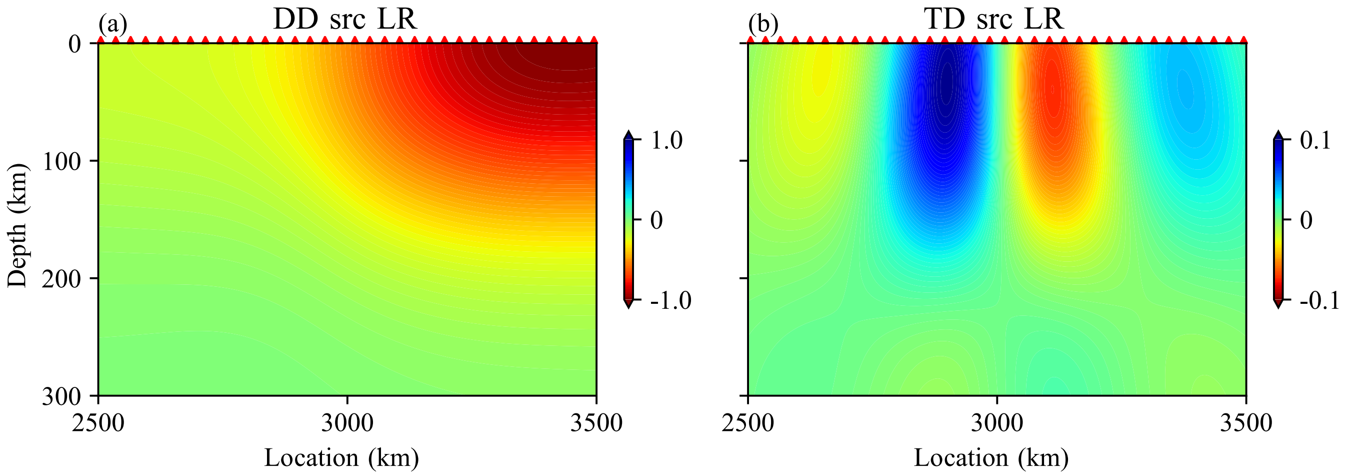

Figure 5. Sensitivity kernel from all the stations, with 30 km station spacing. Station pairs with cross-correlation coefficients larger than 0.95 are used. Target model is Fig. 1(a), starting model is Fig. 1(d). (a) Double-difference kernel. (b) Triple-difference kernel.

syn and obs represent synthetics based on the reference model and observations based on the true model, respectively (e.g. t_{L2}^{syn} is the synthetic traveltime from the left source to the right receiver derived from the reference model). There are triple differences involved in the misfit function: the difference between two adjacent stations (1 and 2), the difference between sources from opposite directions (L and R) and the difference between synthetics and observations (syn and obs). In contrast to the triple difference misfit function established here, the previously established double different misfit function χ^{dd} for two sources can be expressed as:

$$\chi^{\mathrm{dd}} = \frac{1}{2} \sum_{k} \left[\Delta\Delta t_{L12}(\omega_k)^2 + \Delta\Delta t_{R21}(\omega_k)^2 \right], \tag{6}$$

where

$$\Delta\Delta t_{L12} = \left(t_{L2}^{\mathrm{syn}} - t_{L1}^{\mathrm{syn}}\right) - \left(t_{L2}^{\mathrm{obs}} - t_{L1}^{\mathrm{obs}}\right), \tag{7}$$

$$\Delta\Delta t_{R21} = \left(t_{R1}^{\mathrm{syn}} - t_{R2}^{\mathrm{syn}}\right) - \left(t_{R1}^{\mathrm{obs}} - t_{R2}^{\mathrm{obs}}\right). \tag{8}$$

It is worth noting that although the differential traveltimes between two stations are generally much larger than the difference between the left and right differential traveltimes for the same station pair (e.g. Fig. 1c), the triple-difference misfit can be comparable to the double-difference misfit when a good smooth reference model is used.

Although only differences for a single receiver pair and source pair is demonstrated here, the misfit function above and adjoint source equations afterwards can be easily generalized to the case of multiple receiver pairs and source pairs, by summing over all possible receiver pair and source pair combinations. This triple difference concept proposed here is built on studies of double difference tomography (Yuan *et al.* 2016; Tong *et al.* 2024) that have shown advantage over absolute difference tomography. In the double difference approach, the difference between stations reduces the influence of source uncertainties and systematic errors. For triple differences, the additional difference between sources in the opposite directions emphasizes the interference effect from scattering at boundaries.

The gradient of the new triple-difference misfit function relative to reference model perturbations can be expressed as:

$$\delta\chi^{\mathrm{td}} = \sum_{k} \Delta\Delta\Delta t_{LR12}\, \delta\left(\left(t_{L2}^{\mathrm{syn}} - t_{L1}^{\mathrm{syn}}\right) - \left(t_{R1}^{\mathrm{syn}} - t_{R2}^{\mathrm{syn}}\right)\right), \tag{9}$$

where δt^{syn} denotes the traveltime perturbation due to model perturbations. Here we assume that the equation can be linearized (Dahlen *et al.* 2000; Hung et al. 2000; Tromp *et al.* 2005), that is, the perturbation of differential traveltime is equivalent to the difference of individual traveltime perturbations:

$$\delta\left(\left(t_{L2}^{\mathrm{syn}} - t_{L1}^{\mathrm{syn}}\right) - \left(t_{R1}^{\mathrm{syn}} - t_{R2}^{\mathrm{syn}}\right)\right) = \left(\delta t_{L2}^{\mathrm{syn}} - \delta t_{L1}^{\mathrm{syn}}\right) - \left(\delta t_{R1}^{\mathrm{syn}} - \delta t_{R2}^{\mathrm{syn}}\right). \tag{10}$$

Based on the Born approximation (Hudson 1977; Wu and Aki 1985) and reciprocity of Green's tensors (Aki and Richards 2002; Dahlen and Tromp 2020), the adjoint source for the triple-difference misfit can then be calculated. In practice, we substitute the double-difference traveltime term in the double-difference adjoint source (Yuan *et al.* 2016) by the triple-difference traveltime to form the triple-difference adjoint source. This requires access to the waveforms from two sources on opposite sides at the same time. So we forward calculate the waveforms from the two sources and store the waveforms first, and then calculate the adjoint sources from both source-pair and station-pair combinations.

The time domain integral of forward and adjoint wavefields yields the gradient of the misfit function in the spatial domain, which is the finite frequency sensitivity kernel (Tromp *et al.* 2005; Bozdağ *et al.* 2011). The sensitivity kernel describes how model perturbations affect the misfit, thus providing information on how the reference model should be updated to minimize the target misfit function. In the example sensitivity kernel construction below, we use the laterally inhomogeneous model shown in Fig. 1(a) as the true model to generate observed waveforms and the laterally homogeneous model shown in Fig. 1(d) as the starting model to generate synthetic waveforms. Despite the double and triple difference misfit functions are calculated based on the same observations and synthetics, because of the difference in the misfit evaluation based on eqs (4)–(8), the double- and triple-difference kernels are significantly different from each other. The double-difference sensitivity kernel is sensitive to the average structure between the two stations, whether from single sources (Figs 4a and b) or from multiple sources (Fig. 4c; Tape *et al.* 2007). In contrast, and as expected, the triple-difference sensitivity kernel is sensitive to the velocity contrast across the boundary, as shown by the opposite sign of the sensitivity kernel across the boundary (Fig. 4d).

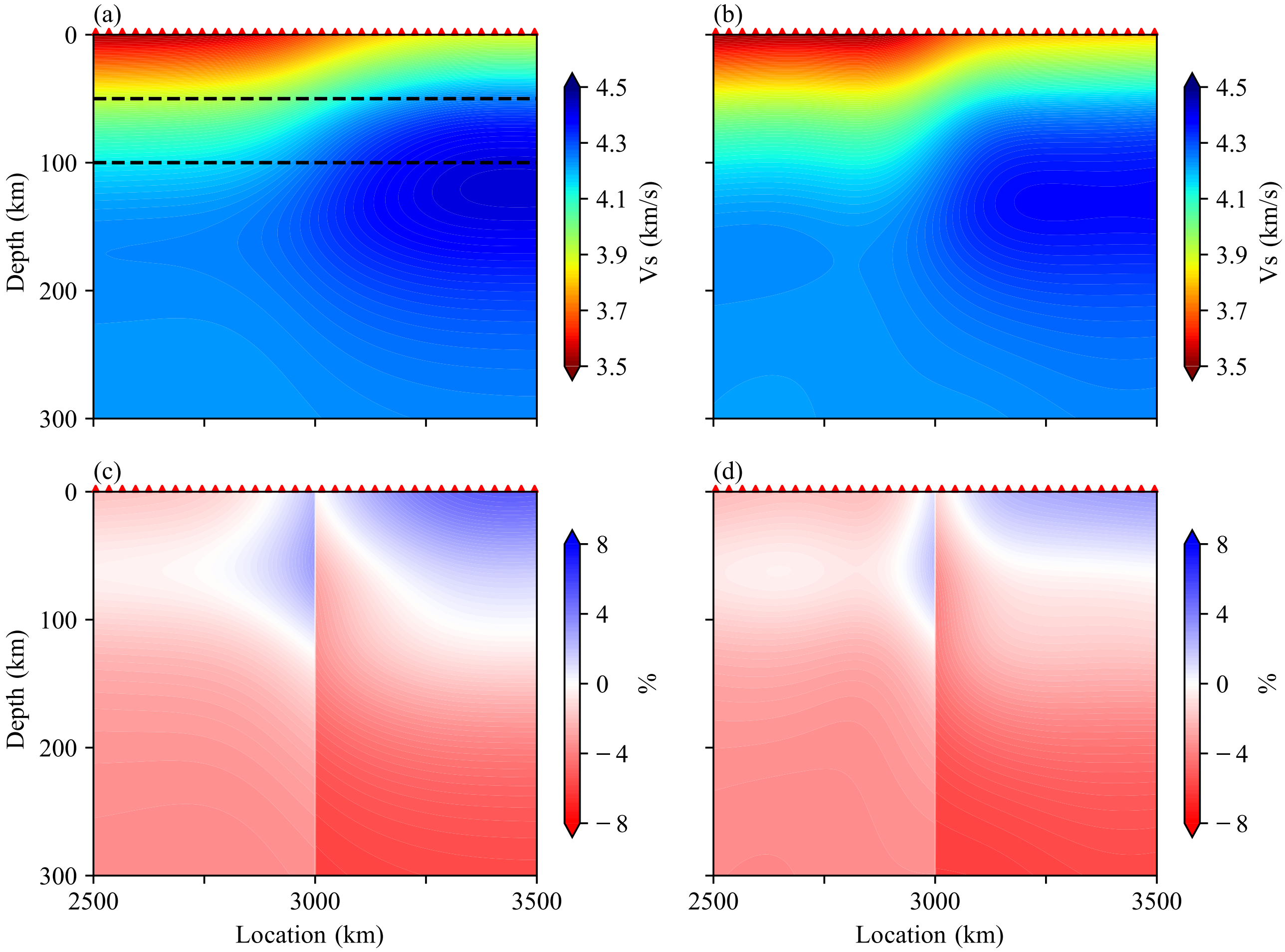

Figure 6. First iteration of inversions for (a) absolute V_s model using the DD method (b) same as (a) but using the DD + TD method. (c) the relative difference between (a) and true V_s model. (d) the relative difference between (b) and true V_s model.

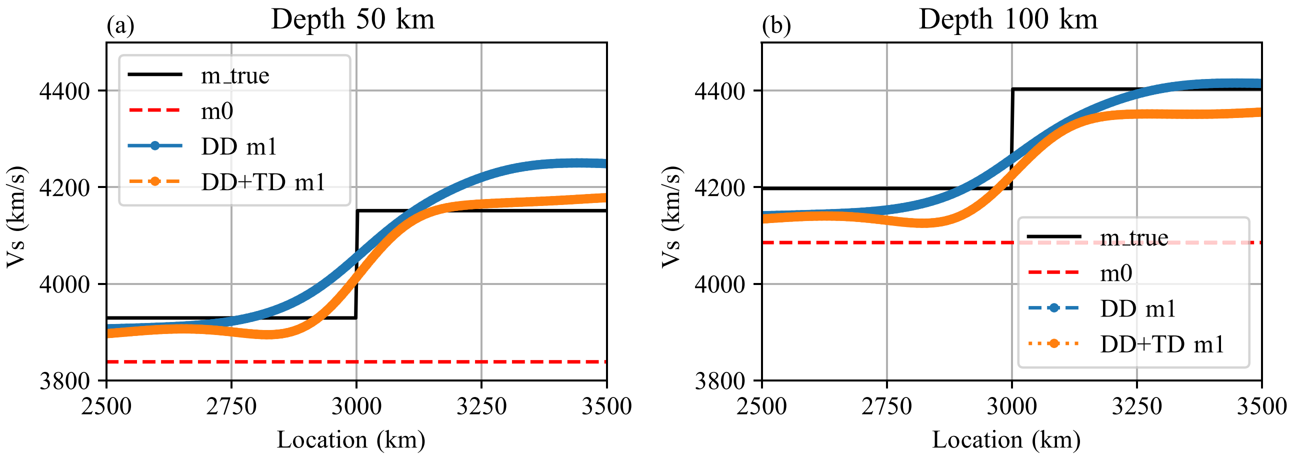

Figure 7. Depth slices of models. (a) V_s at 50 km depth. The black solid line is from the true model. The dashed red line is from the starting model. The blue and orange lines are from the DD and DD + TD inversions at the 1st iteration, respectively. (b) Same as (a) but at 100 km depth.

For multiple station pairs, the total misfit function and sensitivity kernel is the sum of all individual single station pair misfit functions and sensitivity kernels (sum over receiver indices i & j).

$$\chi^{td} = \frac{1}{2}\sum_{i}\sum_{j>i}\sum_{k}\left[\Delta\Delta\Delta t_{LRij}(\omega_k)\right]^2. \tag{11}$$

Fig. 5 shows the double- and triple-difference sensitivity kernel when we consider station pairs across a 30 km spaced array with correlation coefficient larger than 0.95. As the starting model (Fig. 1d) is overall slower than the true model (Fig. 1a), the double-difference sensitivity kernel (Fig. 5a) emphasizes how increasing the velocity, particularly on the right hand side, would

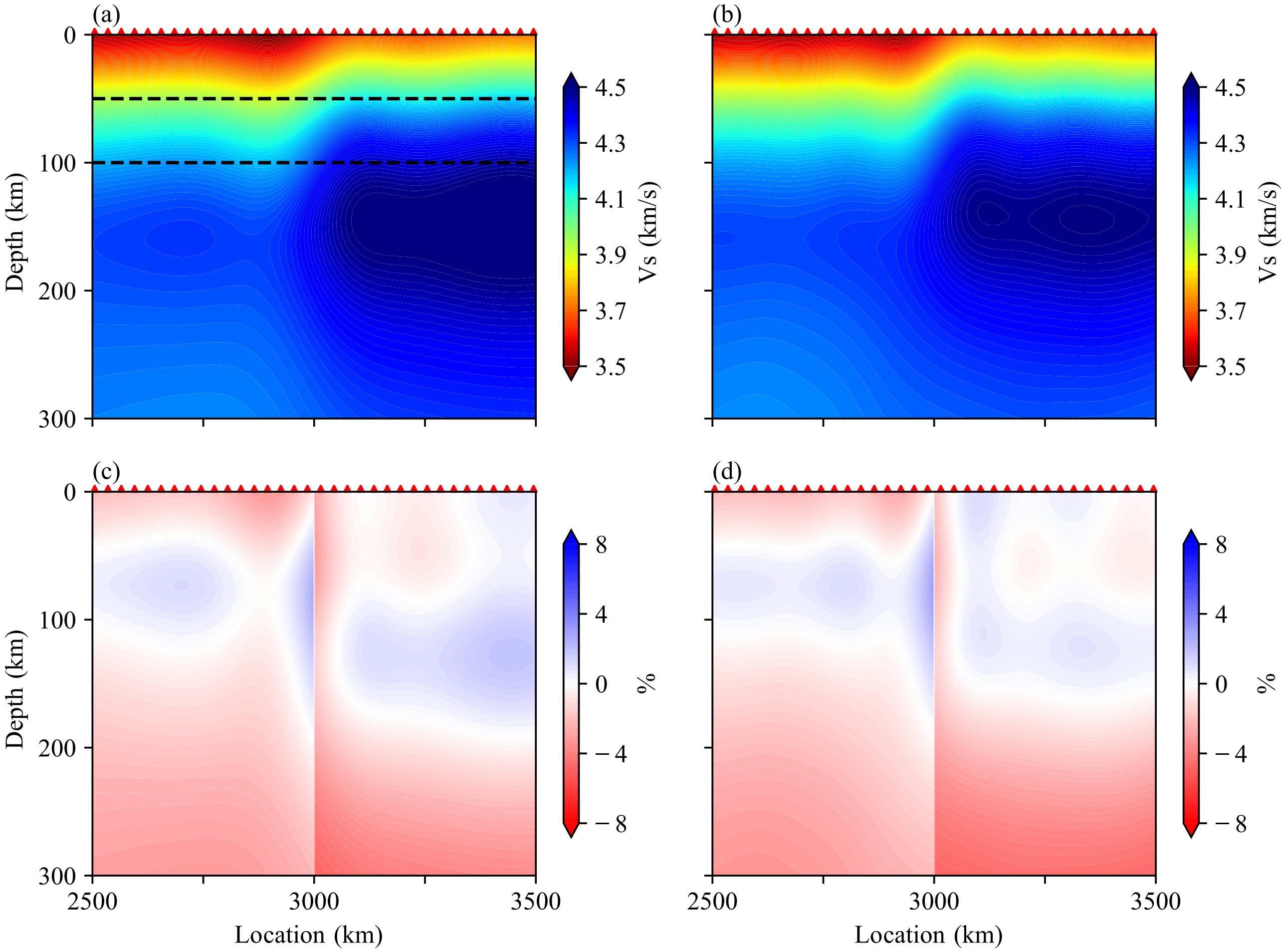

Figure 8. (a) The absolute V_s model at the fifth iteration of inversions using the DD method. (b) Same as (a) but using the DD + TD method. (c) The relative difference between (a) and the true V_s model. (d) The relative difference between (b) and the true V_s model.

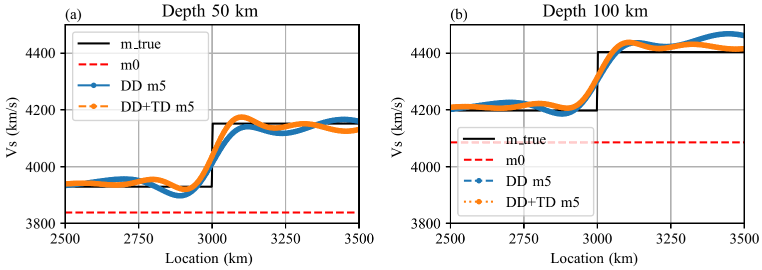

Figure 9. Depth slices of models (a) V_s at 50 km depth. The black solid line is from the true model. The dashed red line is from the starting model. The blue and orange lines are from the DD and DD + TD inversions at the 5th iteration. (b) V_s at 100 km depth.

be the most effective way to reduce the double difference misfit. Although the double-difference kernel gives a sense of the velocity contrast across the boundary, the sensitivity is extremely smoothed. The triple-difference sensitivity kernel (Fig. 5b), on the other hand, emphasizes the velocity contrast across the boundary instead of the absolute velocity changes. We therefore conclude that a tomographic inversion incorporating both double- and triple-difference traveltimes is likely the most effective approach for capturing both the overall absolute velocity structure and the velocity contrasts simultaneously.

Like other finite frequency sensitivity kernels (Yoshizawa & Kennett 2005; Lin & Ritzwoller 2010), the triple-difference traveltime sensitivity kernel has observable sidelobes (e.g. Figs 4d and 5b) that are away from its primary sensitivity near the structure boundary. The presence of sidelobes arises from the single-scattering approximation (Dahlen *et al.* 2000; Zhou *et al.* 2004) inherent in the adjoint-state approach we adapted (Tromp *et al.* 2005). Because of the finite frequency effect, scattered energy away from the correct boundary location could also interfere with the direct wave and result in 1-psi anisotropy near the boundary. While the single

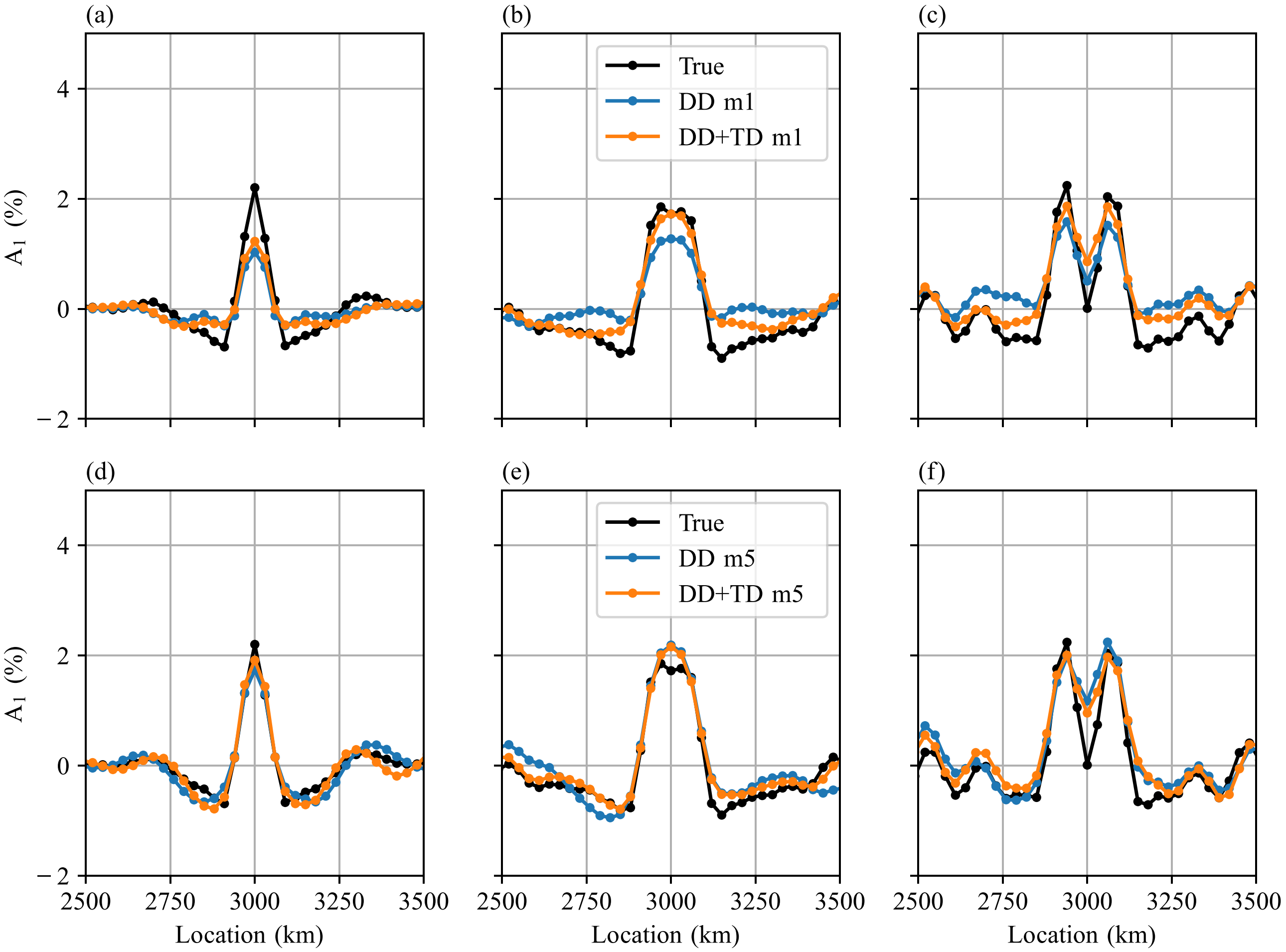

Figure 10. 1-psi anisotropy across the array with varying period with 30 km station spacing. The horizontal coordinate represents the midpoint of the station pair used for the measurement. (a) 1-psi anisotropy at 30 s period. Black line is from measurement from true model simulation. Blue line is from the DD inversion at the 1st iteration. Orange line is from the DD + TD inversion at the 1st iteration. (b) same as (a) but at 60 s period. (c) same as (a) but at 90 s. (d), (e), (f) are same as (a), (b), (c) but with models from inversions at the fifth iteration.

scattering approximation could break down when the true model is substantially different from the reference model (Panning *et al.* 2009; Zhou *et al.* 2011), the multiscattering wavefield phenomenon (Friederich *et al.* 1993) in most cases could still be resolved under the single scattering framework through an iterative process.

The triple-difference traveltime sensitivity kernels constructed in this study are smoothed by convolving Gaussian functions (Tape *et al.* 2007; Tape *et al.* 2010), in order to remove the spurious amplitudes in the vicinity of the sources and receivers. The scale length of the smoothing used is 70 km in both x and z direction, which is around the shortest wavelength of the Rayleigh waves. In Section 4, we compare the double- plus triple-difference inversions versus double-difference inversions to show the improvement by including triple differences. We note that the focus of this study is to demonstrate the applicability of incorporating 1-psi anisotropy measurements and the triple difference misfit into a tomographic inversion. The choice of the 70 km universal smoothing parameter is rather conservative, which limits the overall achievable resolution of the final image. We intend to more thoroughly investigate regularization and process optimization (Fichtner & Trampert 2011; Bozdağ *et al.* 2016) when we apply the method to real data in future studies.

4 TRIPLE DIFFERENCE ADJOINT INVERSIONS

Here we demonstrate the result of adjoint tomography (Tromp *et al.* 2005) by incorporating both double- and triple-difference measurements into the misfit function. To balance the contribution between the two types of measurements, we use the weighted sum of the double-difference (DD) and triple-difference (TD) misfits as the total misfit, where the weighting is determined based on the inverse of the measurement variance. We evaluate the performance of the DD + TD inversion against the inversion based on DD misfit alone.

In the 1st iteration, the DD + TD inversion result clearly outperforms the DD only inversion in resolving the boundary sharpness (Fig. 6). The DD + TD model shows a narrower transition zone when compared with DD model between 0 and 100 km depth and better reflects the true model. Closer inspection reveals that the advantage of the DD + TD inversion is particularly evident at a depth of 50 km (Fig. 7a), which the 50-s Rayleigh waves dominantly excited by the virtual sources are most sensitive to. The absolute velocity on both side of the boundary as well as the sharpness of the boundary is better resolved by the DD + TD inversion. At 100 km (Fig. 7b), although the inverted boundary remains sharper using DD + TD, the velocity structure away from the boundary is better

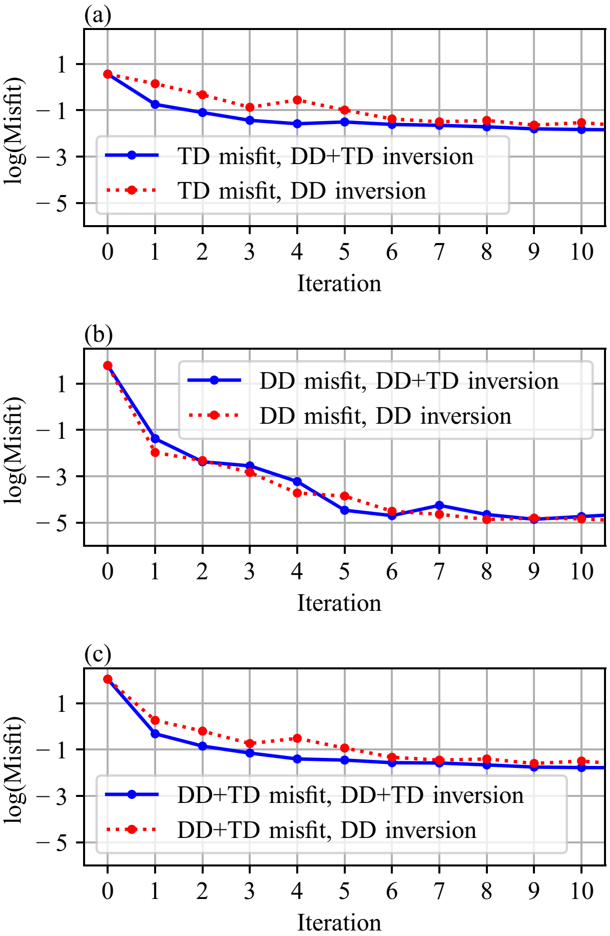

Figure 11. Misfit functions. (a) Triple-difference misfits of the DD + TD inversion models denoted by solid blue lines. Triple-difference misfits of the DD inversion models are denoted by red dotted lines. X-axis is the iteration number, y-axis is the natural log of misfit. (b) same as (a) but for double-difference misfits. (c) same as (a) but of the overall misfit combing double- and triple-difference measurements.

resolved by DD. This is perhaps not surprising considering the DD misfit mostly is affected by absolute velocities instead of boundary sharpness.

By the 5th iteration of inversion, the difference between the DD + TD and DD models (Fig. 8) becomes smaller compared to the 1st iteration, as both begin to converge toward the true model. Nevertheless, the DD + TD inverted V_s profiles at 50 and 100 km depth (Fig. 9) continue to resolve the velocity structure and the boundary sharpness slightly better than the DD inversion. Structures at different depths are resolved progressively through iterations. And the inclusion of TD generally helps to converge to the true model faster, especially near the boundary.

Fig. 10 compares the predicted and observed 1-psi anisotropy from DD + TD and DD inversions at the 1st and 5th iterations. At the 1st iteration, DD + TD inversion fit the 1-psi anisotropy amplitude much better at 60 s (Fig. 10b) than the DD only inversion. At 30 s, DD + TD slightly fits better (Fig. 10a). At 90 s, DD + TD fits worse than DD right at the boundary but fits better for locations away from the boundary (Fig. 10c). At the 5th iteration, the DD + TD and DD inversions generally fit the 1-psi anisotropy similarly. However, the DD + TD inversion fits the 1-psi anisotropy amplitude slightly better for locations away from the boundary at 60 s (Fig. 10e) and for locations right at the boundary at 30 s and 90 s (Figs 10d and f).

Fig. 11 shows the TD, DD and total misfit functions through inversion iterations. The TD misfit from the DD + TD inversion is consistently smaller than that from the DD inversion throughout the iterations, but particularly in the early stages (Fig. 11a). Compared to TD misfit, differences in the DD misfit between the two inversions are less obvious. While the DD only inversion does have a lower DD misfit in the 1st iteration, the TD + DD inversion overall achieves similar DD misfit throughout the iterations (Fig. 11b). This suggests that the TD measurements are not intrinsically contradictory to the DD measurements; therefore, incorporating TD into the inversion does not necessarily compromise the DD misfit. The overall DD + TD misfit from the DD + TD inversion is consistently smaller than that from the DD inversion through all iterations, mainly due to the improvement of the TD misfit (Fig. 11c). As both the TD + DD and DD inversions eventually converge toward a similar model, no obvious difference in misfits are observed after the fifth iteration. As any given TD measurement can be derived from a pair of DD measurements, a model that can explain all DD measurements will automatically also explain all TD measurements. Nevertheless, including TD measurements accelerates convergence toward the final model by enhancing the resolution of boundary sharpness in the early iterations.

We note that the magnitude of the DD misfit heavily depends on how close the reference model is from the true model. When the initial reference velocity is not accurate, the DD misfit could be a few orders of magnitude larger than the triple difference misfit (Fig. 11). However, as the model improves and becomes more accurate through the inversion, the DD misfit becomes comparable to the TD misfit in later iterations. When applying the method to real data, realistic uncertainty estimates (Yan *et al.* 2024) should be incorporated when determining the DD and TD misfits. Previous studies have shown that reliable 1-psi anisotropy can be observed near major structural boundaries in various regions, with measurement uncertainties suppressed by including many independent observations (Lin & Ritzwoller 2011a; Mauerberger *et al.* 2021; Zeng *et al.* 2024). Similar to DD, TD method should have the same advantage of reducing influence from the source signature (Yuan *et al.* 2016). Incorporating the 1-psi anisotropy constrained by real data into the tomography scheme proposed here remains a subject for future study.

5 CONCLUSIONS

In this study, we perform forward wavefield simulations for velocity models with boundary widths of 0, 100 and 300 km and measure 1-psi anisotropy over 30-, 60- and 90-s periods, with 10, 30 and 70 km station spacings. These results show that 1-psi anisotropy amplitudes and patterns are sensitive to the boundary transition width. The sharpness of the structural boundary can potentially be resolved using 1-psi apparent anisotropy when the station spacing is sufficiently dense. We incorporate 1-psi anisotropy into seismic tomography by defining a new triple-difference traveltime misfit function. The corresponding sensitivity kernel and inversion are realized through adjoint methods. The triple-difference sensitivity kernel shows better illumination of the velocity boundary compared with the double difference method. We combine triple-difference and double-difference misfit functions to perform seismic inversion. Synthetic inversion results show that the DD + TD inversion resolves the boundary better in early iterations, with a faster convergence rate compared with DD inversion. For subsequent studies using real data sets, more complexities will appear, including the noise, uncertainties and imperfect array geometries. While the applicability, advantages and limitations of triple difference traveltime adjoint tomography are yet to be tested in such settings, but we expect the improved convergence near boundaries to be similar to that in our synthetic tests and therefore provide some benefits to using triple-difference measurements.

ACKNOWLEDGMENTS

This project is funded by National Science Foundation (EAR 1753362 & 2438772). QZ would like to thank Wenyong Pan for helps on running seisDD, Michael Ritzwoller for suggestions on forward simulations and Brandon Schmandt for financial support on attending AGU Annual Meeting 2025.

CONFLICT OF INTEREST

The authors acknowledge that there are no conflicts of interest recorded.

DATA AVAILABILITY

All the examples are numerical synthetics. We use seisDD package (https://github.com/yanhuay/seisDD) from Yuan *et al.* (2016) for numerical simulations.

REFERENCES

Aki, K. & Richards, P.G. 2002. *Quantitative Seismology,* 2nd edn, Univ. Science Books.

Artemieva, I.M. 2009. The continental lithosphere: reconciling thermal, seismic, and petrologic data, *Lithos,* **109**(1-2), 23–46.

Benz, H.M. & Vidale, J.E. 1993. Sharpness of upper-mantle discontinuities determined from high-frequency reflections, *Nature,* **365**(6442), 147–150.

Bolton, H. & Masters, G., 2001. Travel times of P and S from the global digital seismic networks: Implications for the relative variation of P and S velocity in the mantle, *Journal of Geophysical Research: Solid Earth,* **106(B7),** 13527–13540.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D. 2016. Global adjoint tomography: first-generation model, *Geophys. Suppl. Mon. Not. R. Astr. Soc.,* **207**(3), 1739–1766.

Bozdağ, E., Trampert, J. & Tromp, J. 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.,* **185**(2), 845–870.

Brocher, T.M. 2005. Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. seism. Soc. Am.,* **95**(6), 2081–2092.

Cui, C., Lei, W., Liu, Q., Peter, D., Bozdağ, E., Tromp, J. & Pugmire, D. 2024. GLAD-M35: a joint *P* and *S* global tomographic model with uncertainty quantification, *Geophys. J. Int.,* **239**(1), 478–502.

Dahlen, F. & Tromp, J., 2020. Theoretical global seismology. *In*..

Dahlen, F.A., Hung, S.H. & Nolet, G. 2000. Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.,* **141**(1), 157–174.

De Hoop, M. V. & Van der Hilst, R. D., 2005. On sensitivity kernels for 'wave-equation'transmission tomography, *Geophysical Journal International,* **160(2),** 621–633.

Fichtner, A. & Trampert, J. 2011. Hessian kernels of seismic data functionals based upon adjoint techniques, *Geophys. J. Int.,* **185**(2), 775–798.

Fichtner, A., Kennett, B.L. & Trampert, J. 2013. Separating intrinsic and apparent anisotropy, *Phys. Earth planet. Inter.,* **219,** 11–20.

Fichtner, A., Kennett, B.L., Tsai, V.C., Thurber, C.H., Rodgers, A.J., Tape, C. & Lin, F.C. 2024. Seismic tomography 2024, *Bull. seism. Soc. Am.,* **114**(3), 1185–1213.

Friederich, W., Wielandt, E. & Stange, S. 1993. Multiple forward scattering of surface waves: comparison with an exact solution and born single-scattering methods, *Geophys. J. Int.,* **112**(2), 264–275.

Fukao, Y., Widiyantoro, S. & Obayashi, M. 2001. Stagnant slabs in the upper and lower mantle transition region, *Rev. Geophys.,* **39**(3), 291–323.

Helffrich, G.R., Stein, S. & Wood, B.J. 1989. Subduction zone thermal structure and mineralogy and their relationship to seismic wave reflections and conversions at the slab/mantle interface, *J. geophys. Res.: Solid Earth,* **94**(B1), 753–763.

Hudson, J. A., 1977. Scattered waves in the coda of P, *Journal of Geophysics,* **43(1),** 359–374.

Hung, S. H., Dahlen, F. A. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimeséII. Examples, *Geophysical Journal International,* **141(1),** 175–203.

Hutchings, S.J., Koper, K.D., Burlacu, R., Zeng, Q., Lin, F.C. & Zandt, G. 2025. Upper mantle earthquakes along the edge of the Wyoming Craton, *Geophys. Res. Lett.,* **52**(9), e2024GL114073.

Lin, F.C. & Ritzwoller, M.H. 2010. Empirically determined finite frequency sensitivity kernels for surface waves, *Geophys. J. Int.,* **182**(2), 923–932.

Lin, F.C. & Ritzwoller, M.H. 2011a. Apparent anisotropy in inhomogeneous isotropic media, *Geophys. J. Int.,* **186**(3), 1205–1219.

Lin, F.C. & Ritzwoller, M.H. 2011b. Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure, *Geophys. J. Int.,* **186**(3), 1104–1120.

Lin, F.C., Moschetti, M.P. & Ritzwoller, M.H. 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.,* **173**(1), 281–298.

Liu, Q. & Gu, Y.J. 2012. Seismic imaging: from classical to adjoint tomography, *Tectonophysics,* **566-567,** 31–66.

Liu, X. & Ritzwoller, M.H. 2025. The effect of Rayleigh–Love coupling in an anisotropic medium, *Geophys. J. Int.,* **241**(2), 1204–1225.

Liu, X. 2020. Finite-frequency sensitivity kernels for seismic noise interferometry based on differential time measurements, *J. geophys. Res.: Solid Earth,* **125**(4), e2019JB018932.

Mauerberger, A., Maupin, V., Gudmundsson, Ó. & Tilmann, F. 2021. Anomalous azimuthal variations with 360° periodicity of Rayleigh phase velocities observed in Scandinavia, *Geophys. J. Int.,* **224**(3), 1684–1704.

Montagner, J.P. & Nataf, H.C. 1986. A simple method for inverting the azimuthal anisotropy of surface waves, *J. geophys. Res.: Solid Earth,* **91**(B1), 511–520.

Montelli, R., Nolet, G., Masters, G., Dahlen, F. A. & Hung, S. H., 2004. Global P and PP traveltime tomography: rays versus waves, *Geophysical Journal International,* **158(2),** 637–654.

Pan, W., Innanen, K.A. & Wang, Y. 2020. SeisElastic2D: an open-source package for multiparameter full-waveform inversion in isotropic-, anisotropic-and visco-elastic media, *Comput. Geosci.,* **145,** 104586.

Panning, M. & Romanowicz, B. 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.,* **167**(1), 361–379.

Panning, M.P., Capdeville, Y. & Romanowicz, B.A. 2009. Seismic waveform modelling in a 3-D Earth using the Born approximation: potential shortcomings and a remedy, *Geophys. J. Int.,* **177**(1), 161–178.

Schmandt, B., Lin, F.C. & Karlstrom, K.E. 2015. Distinct crustal isostasy trends east and west of the Rocky Mountain Front, *Geophys. Res. Lett.,* **42**(23), 10–290.

Simmons, N. A., Forte, A. M., Boschi, L. & Grand, S. P. 2010. GyPSuM: A joint tomographic model of mantle density and seismic wave speeds. *Journal of Geophysical Research: Solid Earth,* **115(B12)**. doi: 10.1029/2010JB007631.

Smith, M.L. & Dahlen, F.A. 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, *J. geophys. Res.,* **78**(17), 3321–3333.

Stich, D. & Morelli, A. 2007. Reflection of seismic surface waves at the northern Apennines, *Earth planet. Sci. Lett.,* **259**(1-2), 149–158.

Stolyarov, S.N. 1978. Influence of sizes of transitional layers on wave reflection from moving interfaces, *Radiophys. Quantum Electron.,* **21**(2), 118–123.

Tape, C., Liu, Q. & Tromp, J. 2007. Finite-frequency tomography using adjoint methods—Methodology and examples using membrane surface waves, *Geophys. J. Int.,* **168**(3), 1105–1129.

Tape, C., Liu, Q., Maggi, A. & Tromp, J. 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180**(1), 433–462.

Tong, P., Li, T., Chen, J. & Nagaso, M. 2024. Adjoint-state differential arrival time tomography, *Geophys. J. Int.,* **236**(1), 139–160.

Tromp, J., Komatitsch, D. & Liu, Q. 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.,* **3**(1), 1–32.

Tromp, J., Tape, C. & Liu, Q. 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160**(1), 195–216.

Van der Hilst, R. D. & De Hoop, M. V., 2005. Banana-doughnut kernels and mantle tomography, *Geophysical Journal International,* **163(3),** 956–961.

Woodhouse, J.H. & Dziewonski, A.M. 1984. Mapping the upper mantle: three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. geophys. Res.: Solid Earth,* **89**(B7), 5953–5986.

Wu, R. S. & Aki, K., 1985. Scattering characteristics of elastic waves by an elastic heterogeneity, *Geophysics,* **50(4),** 582–595.

Yan, M., Yao, H., Lei, T., Luo, S. & Feng, J. 2024. Linear array double difference adjoint ambient noise tomography of the central Tanlu fault zone, eastern China, *J. geophys. Res.: Solid Earth,* **129**(7), e2024JB028791.

Yoshizawa, K. & Kennett, B.L.N. 2005. Sensitivity kernels for finite-frequency surface waves, *Geophys. J. Int.,* **162**(3), 910–926.

Yuan, Y.O., Simons, F.J. & Tromp, J. 2016. Double-difference adjoint seismic tomography, *Geophys. J. Int.,* **206**(3), 1599–1618.

Zeng, Q., Lin, F.C. & Tsai, V.C. 2024. Spurious Rayleigh-wave apparent anisotropy near major structural boundaries: a numerical and theoretical investigation, *Geophys. J. Int.,* **239**(2), 901–913.

Zhou, Y., Dahlen, F.A. & Nolet, G. 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.,* **158**(1), 142–168.

Zhou, Y., Liu, Q. & Tromp, J. 2011. Surface wave sensitivity: mode summation versus adjoint SEM, *Geophys. J. Int.,* **187**(3), 1560–1576.