Geophysical Research Letter

Supporting Information for

Localized anisotropy in the mantle transition zone due to flow through slab gaps

Han Zhang[1*], Brandon Schmandt[1], Jin S. Zhang[1,2]

1. Department of Earth and Planetary Science, University of New Mexico, Albuquerque, NM, USA.
2. Institute of Meteoritics, University of New Mexico, Albuquerque, NM, USA.

Contents of this file

Additional Supporting Information (Files uploaded separately)

Text S1

The teleseismic shear wave splitting (SWS) measurements were first averaged at each station using circular averaging (Montagner et al., 2000). Then we created a $1^{\circ} \times 1^{\circ}$ grid dataset by averaging the station-based SWS measurements in a 200 km radius bin.

Text S2

The transverse energy minimization method uses a grid search framework to find the splitting parameters, fast-axis orientation and delay time, to minimize the energy of the SH component within a target time window (Long & Silver, 2009). The grid intervals for our fast-axis orientation and delay time are 1° and 0.1 s, respectively. The time windows were set as ± 4 s from the peak amplitude of the *Ps* phase on the stacked SV component. After finding the global minimum, under the assumption of Gaussian noise, it is conventional to report the uncertainty of the estimate using an F-test formulation (Walsh et al., 2013). Here we reported the 95% confidence level of our estimates.

However, given the fact that the noise may not follow a Gaussian distribution, the grid search results from low SNR phases (e.g. *P*410*s* and *P*660*s*) often suffer from an artificial global minimum unless further quality control is applied. We adopted a bootstrap based metric to qualify the reliability of the estimates based on how much they reduce the amplitude of the corrected SH stack compared to the raw SH stack (Fig. 1d). The metric was defined as

$$M = (RMS_{zero} - RMS_{min})/SD_{boot}$$

where the RMS_{zero} stands for the root-mean-square (rms) value of a zero-lag time SH stack, RMS_{min} is the rms value from the minimum energy parameters. SD_{boot} is the standard deviation of the rms values from 200 bootstrap resamples of the target SH window. A larger M means greater rms reduction and suggests a more reliable global minimum. Examples with various M values can be found in Fig. S5. Figure. 1d only gives the estimates from the *P*660*s* phase with M greater than 3.

Text S3

Since the *P*410*s* and *P*660*s* phases are much weaker than the *SKS* phase, stacking receiver functions from different azimuths is required to obtain stable measurements. Synthetic tests indicate that the polarity of the *Ps* phases on the SH component changes when the back azimuth crosses the fast- or slow-axis of anisotropy (Fig. S6). Therefore, to prevent deconstructive interference of traces from different azimuths during stacking, we flipped the SH traces in the 2nd and 4th quadrants using the fast-axis orientation estimated from the *P*660*s* phase. The flip was applied to both *Ps* phases because the null hypothesis of an isotropic mantle transition zone (MTZ) suggests that the two *Ps* phases share the same fast-axis orientation.

The amplitude ratio was then measured as the peak shear wave amplitude on the integrated stacked SH component over that on the stacked SV component (hereinafter referred to as P_{SH}/P_{SV}). Theoretically, under weak anisotropy (delay time less than a tenth of a period), the P_{SH}/P_{SV} amplitude ratio is proportional to the delay time (Montagner et al., 2000). Synthetic tests demonstrate that the positive correlation continues to moderate anisotropy (delay time up to one third of a period), where the amplitude ratio becomes increasingly sensitive as the magnitude of anisotropy increases (Fig. S1).

Bootstrap resampling, with 200 samples, was used to assess the uncertainties of the amplitude ratios assuming a normal distribution of the resampled stacks. The paired Cohen's distance was calculated using the following equation

$$Cohen's\ d = (\overline{AmpR_{660}} - \overline{AmpR_{410}})/SD_{diff}$$

where $\overline{AmpR}$ is the mean of measured amplitude ratios for the *P*410*s* and *P*660*s* phase, and SD_{diff} is the standard deviation of the two groups' difference. Using the corresponding paired 2-sample t-test, a two-tailed 68% confidence level would require Cohen's distance either greater than 1.0 or smaller than -1.0. A two-tailed 95% confidence level sets thresholds at ±2.0. The observed Cohen's distances were superimposed on the tomographic map in Figure. 2b.

Text S4

We explored the effects of two anisotropic layers with different fast-axis orientations using synthetic data (Fig. S3; Supporting Information S7). The anisotropy in the top layer (36 - 320 km) is 2.5%, and that of the bottom layer (360 - 556 km) is 1.5%. Accordingly, the delay times from the two layers are ~1.3 s and ~0.5 s respectively. The fast-axis orientation of the upper layer was fixed at 89˚, which is the mean fast-axis orientation estimated from the *P*660*s* phase (Fig. S6).

Since the top layer dominates the depth-integrated anisotropic effect, the fast-axis orientation estimations from the energy minimization method show no significant difference with the 95% confidence intervals. Such results suggest the fast-axis orientation measurements are not ideal to constrain differential orientations if the common layer dominates the total anisotropic effect. In contrast, the delay times show greater variations with respect to the differential orientations, but are still less sensitive when compared with the Cohen's distances.

Increasing the differential orientation between the two layers moves the resulting Cohen's distances from a constructive interference area (0˚ - 30˚) to a neutral area (40˚ - 60˚), and finally to a destructive interference area (70˚ - 90˚). Such results provide an alternative fit to the neutral observations and a potential explanation for the deconstructive observations in Fig. 2b.

Text S5

To illustrate a range of possible anisotropic structures beneath our study region, we constructed three types of forward seismic models in Fig. 3a. Several constraints were taken from the previous seismic model, mineral physics data, and our observation here from *P*660*s*. Firstly, the depths of mantle discontinuities come from previous migration results (Zhang & Schmandt, 2019). The depth-integrated delay time from the three models were set to match the mean estimated delay time from the *P*660*s* in the best resolved regions (light background in Fig. 1d), which is 1.4 s (Fig. S7). Such a setting makes the three models indistinguishable from *SKS* data alone.

Moreover, from bottom to top, we listed the detailed constraints in each layer below.

Lower mantle

- The agreement between the splitting parameters from the *P*660*s* and SWS suggests an isotropic lower mantle (Fig. 1d), which applies to all three models.

Mantle transition zone

1. As a control group, there is no transition zone anisotropy in model #1.
2. At the ambient mantle, mineral physics data suggested a nearly isotropic lower transition zone layer even lattice preferred orientation of ringwoodite was developed (Fig. S4). Therefore, model #2 has an isotropic lower MTZ layer. We further assumed a uniform upper MTZ layer with gradual transition at the 520 for simplicity.

3. When slab is present in the transition zone, atypical minerals such as phase E and akimotoite may contribute to the recorded anisotropic signal. Therefore, the deeper anisotropic layer in model #3 is set to match the depth extent (~380 km to ~620 km) of the slab suggested from tomographic results (Fig. 2c).

At the 410

1. As a control group, there is no anisotropy near the 410 in model #1.
2. At the ambient mantle, mineral physics data suggested a ~30% anisotropy drop across the 410 in a pyrolite model, which comes from the lower intrinsic elastic anisotropy of wadsleyite compared with olivine (Fig. S4). Under the assumption of simple shear deformation from mantle flow, the ~30% drop was honored in model #2.
3. When slab is present in the transition zone, the deformation may not follow the simple shear assumption. Therefore, we didn't implement any constraints at the 410 in model #3.

Upper mantle

- Surface wave studies in this region suggest ~1-2% azimuthal anisotropy within the first ~200 km (Wagner & Long, 2013), which applies to all three models. Due to the poor depth resolution in the upper mantle of our data, and for the sake of simplicity, we further assumed a uniform anisotropy strength in the upper mantle layer for all three models.

Crust

- There is no clear evidence of strong crustal anisotropy so we set it to be zero for simplicity.

While there are countless models that are consistent with the observational constraints, depth resolution is inadequate to constrain detailed structure. So, we consider three models to highlight potential anisotropy contributions from the upper mantle above 410 km, the upper transition zone where wadsleyite is stable, and the lower transition zone. The detailed parameters of our models are listed below.

1. Model #1 only contains an anisotropic layer extending from the Moho (at 36 km) to 400 km. The strength of anisotropy is 2.0% from the Moho down to 320 km and then linearly decreases to 0% at 400 km. The 80 km gradual

transition avoids strong artificial P-to-S conversions on the synthetics, which were not found in observational receiver functions.

2. Model #2 has a uniform anisotropic layer from the Moho down to the 410 discontinuity (at 428 km), which is underlain by another uniform layer of anisotropy in the upper MTZ (428-556 km). The strength of anisotropy in the upper mantle layer is 1.4% and that of the upper MTZ layer is 1.0%. The ~30% drop in the strength across the 410 discontinuity comes from the mineral physics constraints mentioned before. A 40 km thick 520 discontinuity (at 556 km) was set to avoid strong artifacts.
3. Model #3 has two separated anisotropic layers. The top layer (Moho-260 km) has a strength of 1.0% while the deeper one (380-620 km) has a maximum strength of 3.0%. To avoid artifacts from sharp contrast, the lower boundary of the top layer and the two boundaries of the second layer have gradual transitions over 80 km.

Using synthetic receiver functions (Supporting Information S7), figure 3a and S7 give the distribution of the amplitude ratios and the estimated splitting parameters from the three models.

Text S6

We used a grid search method to invert preferred models for explaining the observed amplitude ratios. At each stacking point, we used noise-free synthetics with a ray parameter and back-azimuth distribution identical to the observation. Such a process ensures the magnitudes of anisotropy within the two layers are the only factors affecting the amplitude ratios.

In the model #2 parameterization, the top layer has a uniform anisotropy between the Moho and 320 km depth. The second layer starts from 400 km and extends down to 556 km. The 30% anisotropy drop across the 410 (at 428 km) is kept in the second layer. From 320 km to 400 km, a gradual transition between the two layers was applied to avoid strong artifacts. The two magnitudes used in the grid search are the uniform anisotropy in the top layer and the maximum anisotropy in the upper transition zone layer (Fig. 3a, #2).

In the model #3 parameterization, the top layer extends from the Moho to 260 km. The second layer starts at 380 km and extends to 620 km. Gradual transitions over 80 km at the boundaries are applied. The two magnitudes in grid search are the maximum anisotropy within the two layers (Fig. 3a, #3).

The P_{SH}/P_{SV} of the $P410s$ and $P660s$ phases from each of the models were calculated using noise-free synthetics. The two standardized squared deviations from the observed means were summarized as a misfit term.

$$Q = (\Delta AmpR_{410}/SD_{410})^2 + (\Delta AmpR_{660}/SD_{660})^2$$

where $\Delta AmpR$ is the difference between observed and predicted amplitude ratios for the $P410s$ and $P660s$. SD is the standard deviation of observed amplitude ratios. Given the assumption of normally distributed amplitude ratios, the misfit term Q approximately follows a chi-square distribution with a 2 degrees of freedom (Fig. S8). Accordingly, the 68% confidence intervals of our best fit model were reported using the chi-square distribution in Fig. 3b and c.

When interpreting the inverted magnitudes of anisotropy in the upper mantle and transition zone layers, please keep in mind that only the depth-integrated anisotropy in each layer is constrained rather than the magnitude of anisotropy at a specific depth. Consequently, there is a tradeoff between the strength and thickness of anisotropy within either the upper mantle or transition zone. The upper mantle layer is thicker, so the potential tradeoff is larger. For instance, if anisotropy were concentrated in only a 100 km depth interval of the upper mantle (e.g., mantle lithosphere or asthenosphere), then the actual magnitude of anisotropy would be locally greater than our estimates. The tradeoff may span a smaller range of values in the transition zone layer because it is thinner. For a pyrolite composition, almost all anisotropy is expected in the wadsleyite stability field from about 410 to 520 km (Fig. S4). We consider it less likely that only a subset of the wadsleyite depth interval contains anisotropy because it is smaller and does not include major rheological contrasts. Thus, the optimized values for the transition zone layer are a more localized constraint on the actual magnitude of anisotropy (Fig. 3b and c).

Text S7

The synthetic receiver functions were generated using a reflectivity method (Levin & Park, 1997). Identical to the processes applied on the observed data, the synthetic receiver functions were rotated to the P-SV-SH coordinate and were filtered between 0.07 Hz and 0.25 Hz. We then contaminated the synthetics with pre-event noise collected from observed USArray data. We assumed a P wave SNR of 5 when adding the noise to the synthetics. Normal moveout correction was applied to the synthetics using the input velocity model. Using a ray parameter and back-azimuth distribution identical to the whole PNW dataset, the noisy synthetics produce the forward modeling results in Fig. 3a, Fig. S3 and Fig. S7.

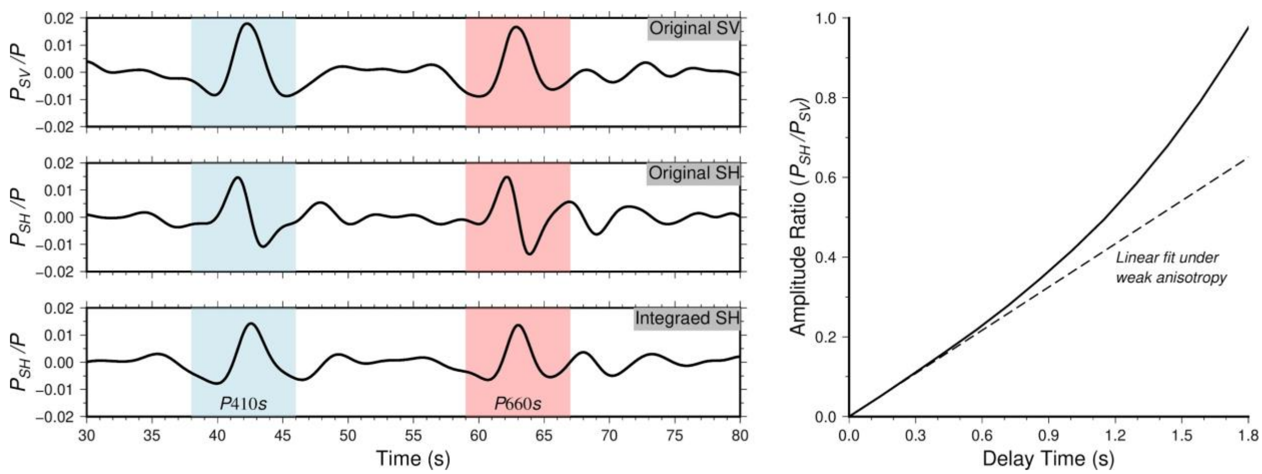

Fig. S1. Synthetic receiver function from an anisotropic upper mantle model and the relationship between the P_{SH}/P_{SV} amplitude ratios and delay times. The amplitude ratio P_{SH}/P_{SV} is measured as the peak shear wave amplitude on the integrated SH component over that on the original SV component. The amplitude ratios are approximately proportional to the delay times under weak anisotropy. At moderate anisotropy, the amplitude ratios become increasingly sensitive to the magnitude of anisotropy.

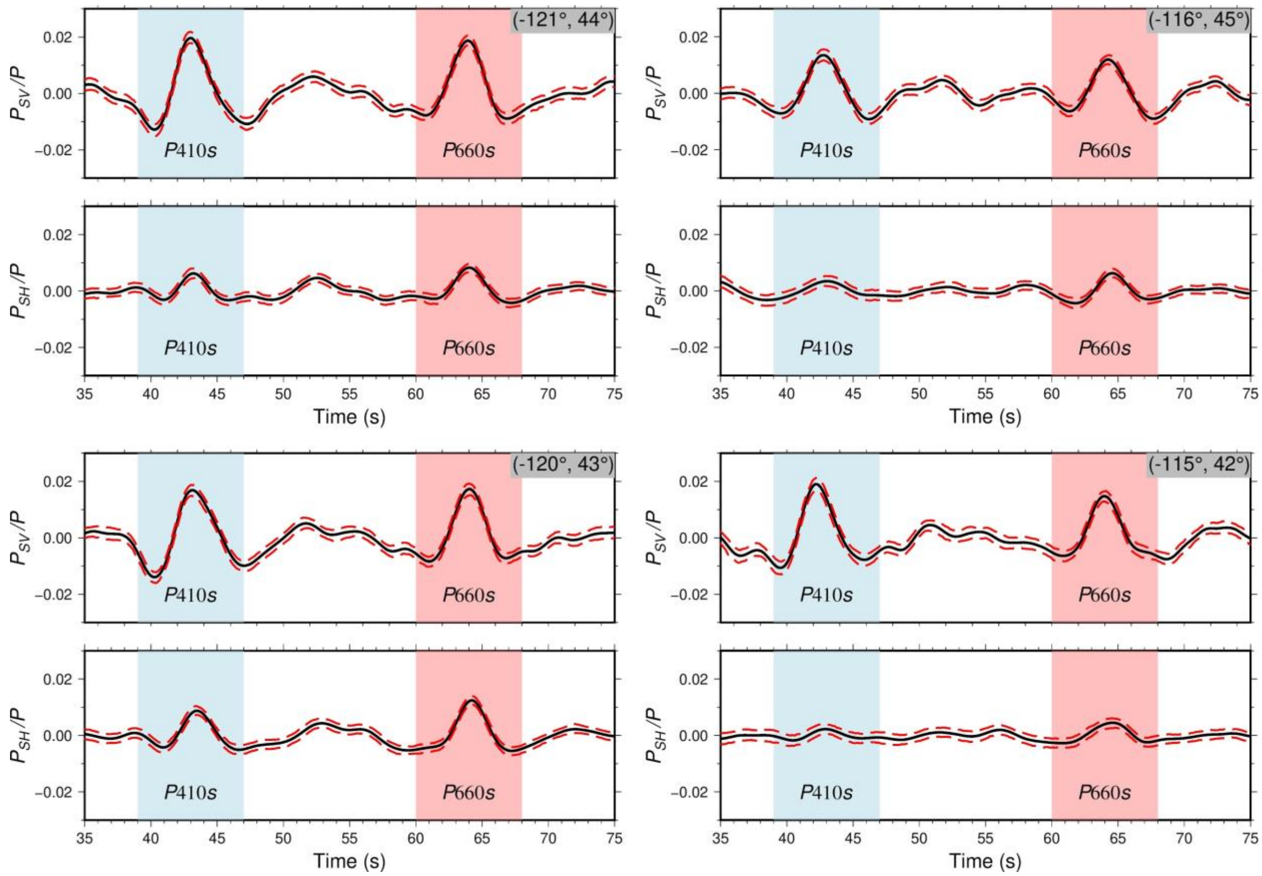

Fig. S2. Additional receiver function examples that require mantle transition zone anisotropy. The locations are labeled at the top right corner. The red dashed line represents the 95% confidence level of the stacked traces from bootstrap resampling.

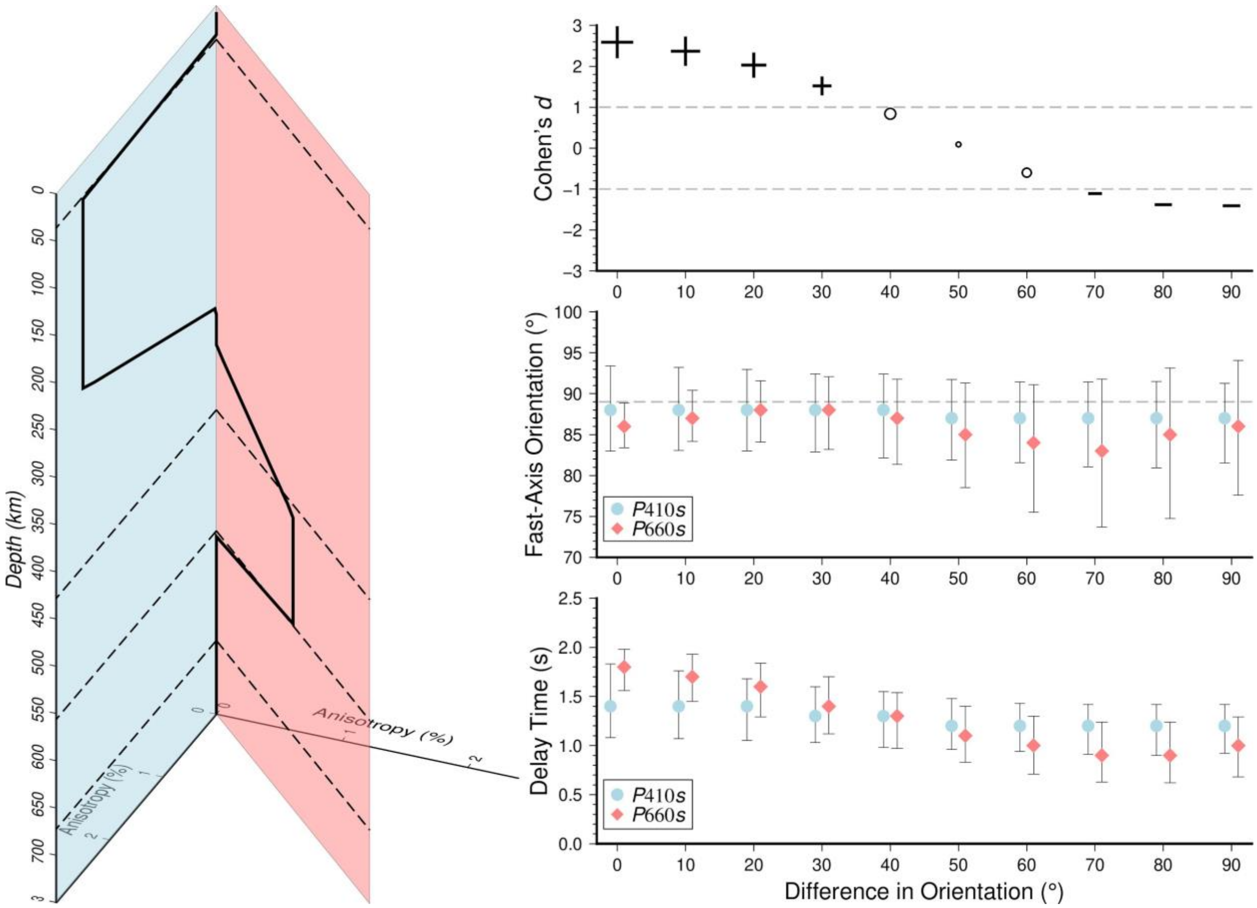

Fig. S3. Synthetic results of two anisotropic layers with different orientations. Cohen's distance shows the greatest sensitivity to the variation of differential orientations among the three measurements.

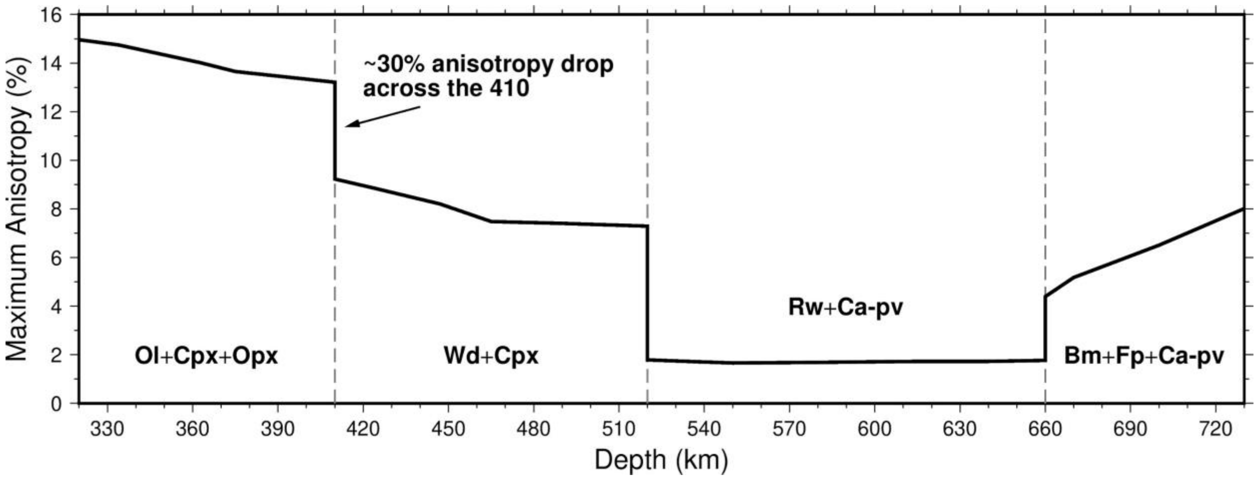

Fig. S4. Maximum shear wave anisotropy indicated from a pyrolite model. There is an ~30% anisotropy drop across the 410 discontinuity primarily due to the significantly lower intrinsic anisotropy of wadsleyite compared with olivine. Only the minerals contributing to the calculated anisotropy are labeled in the figure: olivine (Ol), clinopyroxene (Cpx), orthopyroxene (Opx), wadsleyite (Wd), ringwoodite (Rw), Ca-perovskite (Ca-pv), bridgmanite (Bm), ferropericlase (Fp).

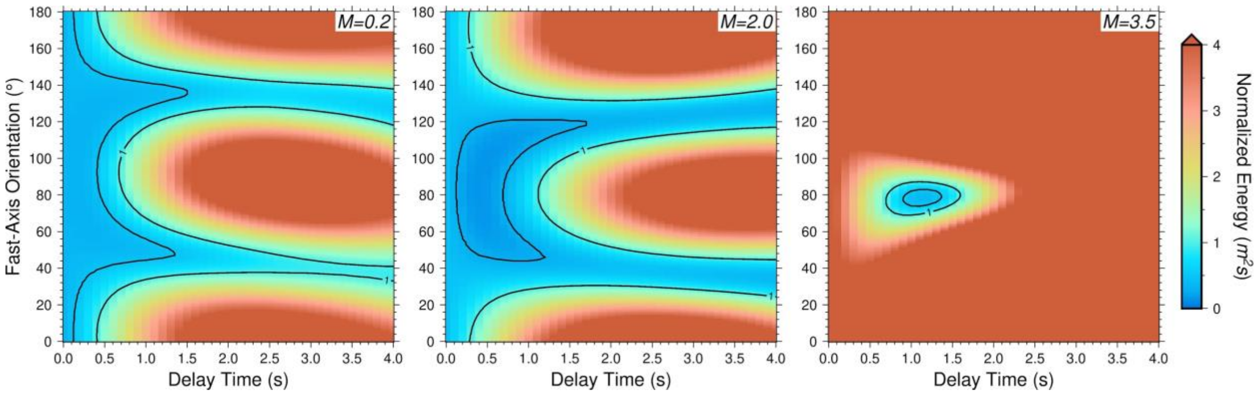

Fig. S5. Examples of the energy minimization results with various *M* values. The *M* values are labeled at the top right corner. The contoured line with normalized energy equal to 1 represents the 95% confidence interval from the F-test. The inner contour represents the 68% confidence interval.

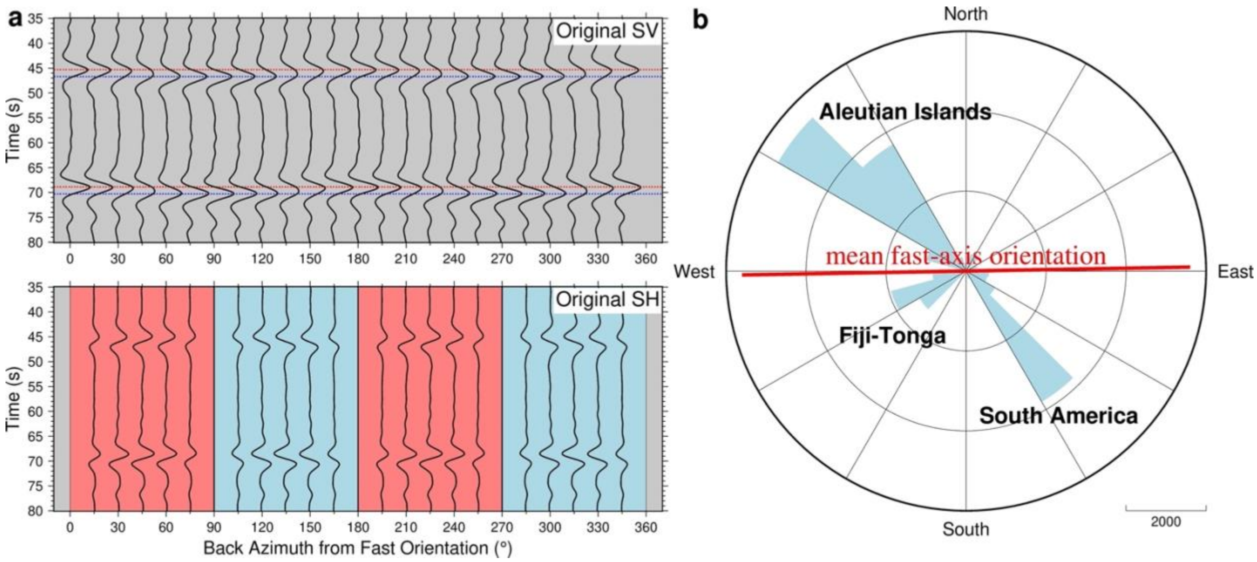

Fig. S6. Synthetic receiver functions and back azimuth distribution of the PNW dataset. (**a**) The top panel shows synthetic receiver functions on the SV component while the bottom panel shows the SH components. The SH components flip polarity after the back-azimuth crosses the fast- or slow-axis. (**b**) The mean fast-axis orientation (89˚) estimated from the *P*660*s* phase is shown by the red solid line. Three regions (Aleutian Islands, South America, and Fiji-Tonga) contributed the majority of the receiver functions in this study.

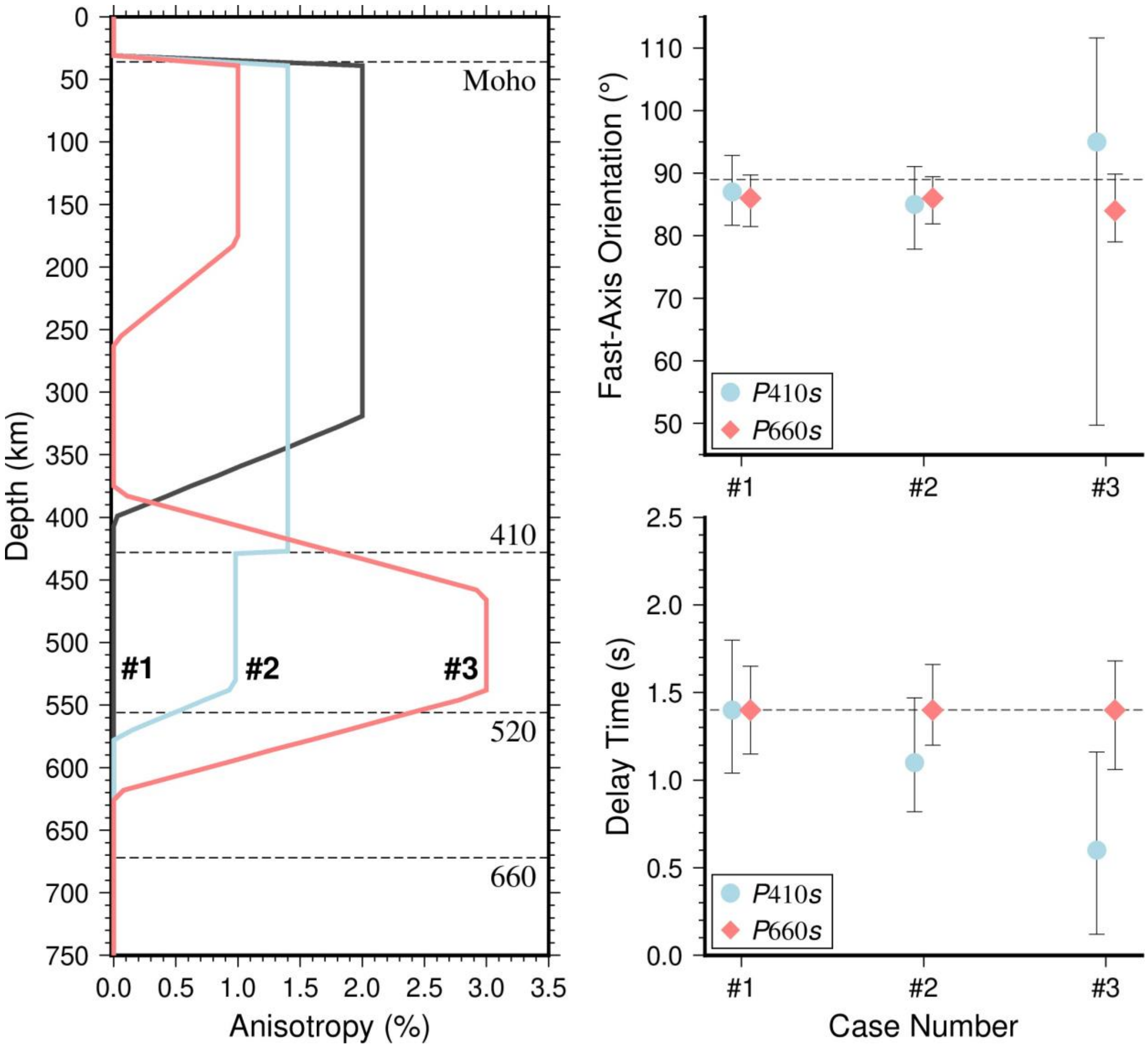

Fig. S7. Forward models and splitting estimations from energy minimization method. All three input models have a fast-axis orientation of 89˚ and a delay time from *P*660*s* of ~1.4 s. The energy minimization method successfully recovers the input with the 95% confidence intervals.

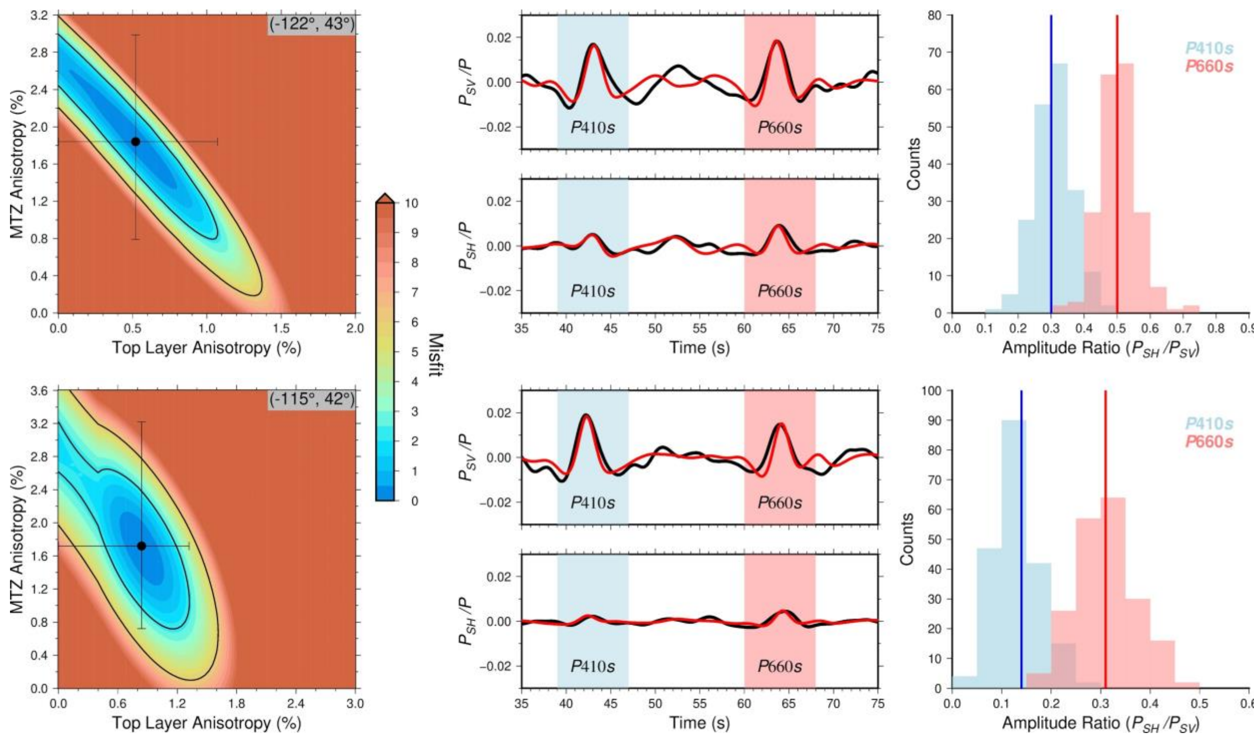

Fig. S8. Examples of misfit distribution from grid search inversion and best fit model. The left panel shows the distribution of the misfit Q. The locations are labeled at the top right corner. The two black contour lines denote the 68% and 95% confidence interval of the estimated parameters respectively. We represented the intervals in Fig. 3b and c using the error bar boxing the 68% area. The middle panel gives the observed receiver functions (black) and the synthetic waveforms (red) from the best fit model. The right panel shows the observed amplitude ratio distributions and the predictions from the best fit model (vertical lines).

Movie. S1. Isosurface of 1% high P wave velocity anomalies beneath the PNW. The locations showing significant anisotropy are denoted by solid spheres at 500 km depth. The white ones indicate constructive interference while the black ones represent deconstructive interference.