Supplementary Data for “**Seismic Architecture of the Lithosphere-Asthenosphere System in the Western United States from Joint Inversion of Body and Surface Wave Observations: Distribution of Partial Melt in the Asthenosphere”**

**Section 1.** Breadth of the NVG

The widths of the NVG used for the preferred inversion are shown in Figure S1. The reader is referred to the main text and to *Hopper and Fischer*, 2018 for details.

Diagram

Description automatically generated

**Figure S1.** Width of the NVG used for the preferred inversion.

**Section 2. Factors controlling the recovery of the gradient below the Moho**

In the main text (Figure 6b), a model with a negative gradient in *Vs* below the Moho is recovered with a positive gradient unless head-waves are included in the inversion. The error is significant and could be interpreted as real structure should this occur when inverting the data. Here, we perturb the synthetic model from Figure 6b in 4 ways to test what factors control the recovery of this gradient. The first two perturbations are to the crust, namely thinning the crust (Figure S2a) and making the gradient in *Vs* in the low crust less steep (Figure S2b). Though the second perturbation is a subtle change the model, both perturbations correct the recovery of the gradient below the Moho. In contrast, changing the mantle structure either by making the NVG deeper (Figure S2c) or making the minimum *Vs* below the NVG faster (Figure S2d) do not alter the recovery of the gradient below the Moho. While all possible perturbations are not explored here, we conclude that this effect is primarily caused by steep gradients in *Vs* in thick crust. The difference between inversions with and without a head-wave constraint (Figure 12b) suggest that this or some similar condition is often reached in the Western US. Since the biggest difference is reached beneath the Rocky Mountains, where the crust is thick, the condition of steep gradients in the lower half of thick crust seems a plausible explanation.

Chart, line chart

Description automatically generated

**Figure S2. Recovery of perturbed versions of the model in Figure 6b of the main text.** Colors are the same as in Figure 6 of the main text, and the perturbations are discussed in the Supplementary Section S2. Note that in Figure 6b of the main text, the sub-Moho gradient is erroneous (pink line in all panels) unless head-wave constraints are enforced. Changes at crustal depths (panels a,b) can prevent the error, while perturbing the NVG (panels c, d) has no effect on the positive sub-moho gradients in the pink models.

**Section 3. Construction of random synthetic models**

To construct smooth profiles of *Vs* within each layer via a sum of Chebyshev polynomials, random values for different features of the models are drawn from the distributions given in Table S1. A 1-km thick Moho is assumed in all cases, while assumed NVG thickness ranges from 10 to 50 km. A model is redrawn if any of the following conditions are violated: the Moho has a velocity contrast <1% or >25%, the NVG is shallower than the Moho, or the NVG has a velocity contrast >-1% or <-20%. At the base of the model (a depth of 410 km), we fix *Vs* and the first derivative to 4.769 and 7e-4 (km/s)/km to match the values in PREM (Dziewonski & Anderson, 1981). Two parameters are also randomized that are not modified during the inversion – the width of the NVG and the *Vp/Vs*. In both cases, a random value is drawn twice – one value used to generate the data and one value assumed during the inversion. The size of the mismatches is given in Table S1, and the effect of these mismatched parameters are thus included in the final uncertainties reported in the main text. This propagates errors in the assumed values of the width of the NVG and the Vp/Vs ratio.

**Table S1. Bounds for parameters defining the synthetic models.**

|  |  |
| --- | --- |
| **Parameter, units** | **Bounds** |
| Depth to Moho, km | 15/60 |
| Depth to NVG, km | 40/120 |
| NVG width, km | 5/50 |
| Vs, at the surface, km/s | 2/3.5 |
| Vs, above the Moho, km/s | 3.5/4.25 |
| δVs/δz, at the surface, (km/s)/km | 0/1e-2 |
| δVs/δz, above the Moho, (km/s)/km | 0/1e-2 |
| Vs, below the Moho, km/s | 4.25/5.0 |
| Vs, above the NVG, km/s | 4.0/4.5 |
| δVs/δz, below the Moho, (km/s)/km | -5e-2/1e-2 |
| δVs/δz, above the NVG, (km/s)/km | -2e-2/2e-2 |
| Vs, below the NVG, km/s | 3.7/4.4 |
| δVs/δz, below the NVG, (km/s)/km | -5e-3/1e-2 |
| **Mismatched parameters** | **Standard Deviation** |
| Vp to Vs ratio | 0.03 |
| NVG width, km | 15, minimum of 10 |

Once the parameters are drawn, we define four non-dimensional quantities that range from -1 to 1 as

(S1)

(S2)

(S3)

(S4)

where z is the depth in km, v is shear-wave velocity in km/s, δ is the first derivative of velocity in km/s, Δ indicates the difference of the variable between the top and the bottom of the layer, and subscript 0 and 1 indicates the value at the top and bottom of the layer, respectively. The hat superscript indicates a non-dimensionalized value. We now express the non-dimensional velocity at a non-dimensional depth in terms of the summation of the first four Chebyshev polynomials of the first kind as

(S5)

where , … coefficients that need to be found. By solving (S5) at -1 and 1, then differentiating (S5) against and again solving at -1 and 1, we find four equations per layer that relate through to the boundary conditions defined above. From this system we find the solutions for the coefficients as

(S6)

(S7)

(S8)

(S9)

from which we can solve (S6) and redimensionalize with (S1)-(S4) to obtain a smooth velocity profile for each layer. The 500 models that result from this process are shown in Figure S3.

Chart

Description automatically generated

**Figure S3. 500 randomly generated synthetic models**. Three randomly selected models are plotted in black for demonstration, and the remaining 497 models are plotted in gray.

Finally, the correlation between the velocities and slopes below the NVG are strongly correlated in the inversion of the data, but not in the inversion of the synthetics (see main text for a discussion). This shows that correlation observed in the data is not an artifact of our regularization procedure (if the second derivative were overdamped, then the correlation between Vs at the base of the NVG and the slope below the NVG would approach 1 since Vs is pinned to a constant value at base of the model). The individual data points are shown in Figure S4.

Chart, scatter chart

Description automatically generated

**Figure S4**. Synthetic velocities against slopes below the NVG before (top) and after (below) inversion.

**Section 4. Comparison between the receiver functions predicted by this study and by Shen and Ritzwoller, 2016**

Our inversion constrains the structure of the Moho in part with information extracted from the model of Shen and Ritzwoller, 2016. To ensure that our model is consistent with the source of these constraints, we compare the receiver functions predicted for our model with the receiver functions predicted for the model of Shen and Ritzwoller, 2016 (Figure S5). Receiver functions are predicted with reflectivity (Park, 1996; Levin & Park, 1997) at a slowness of 0.06 s/km and low-pass filtered to 1 Hz to mimic the frequency content used in Shen and Ritzwoller, 2016. Our models do not include the sediment layer of Shen and Ritzwoller, 2016, and so the predicted receiver functions from their model show large amplitude phases at short time lags caused by both direct and reverberating phases in the sediment layer. Figure S5 shows the comparison at 3 sites chosen to have relatively minimal sedimentary interference, but the sediment layer at all locations leads to some reverberating phases. In all three chosen locations, however, there is good agreement between the timing and amplitude of the Moho phase predicted by both models.

A picture containing chart

Description automatically generated**Figure S5. Comparison between the predicted receiver functions for two studies**. In each panel, the left-hand column shows the model at the given location in the preferred model of this study (black) and in Shen and Ritzwoller, 2016 (blue). The right-hand column shows the predicted P receiver functions for both models. Dashed lines mark the depth of the converted phase from the Moho in the receiver function associated with our model.

**Section 5. Consistency between the observed and predicted S receiver functions.**

We note that the gradient in *Vs* in some regions is greater between the Moho and NVG than within the NVG itself (Figure 11) and show in the supplemental materials that these more peculiar models can adequately reproduce the arrival times of *Sp* phases. We generated synthetic *Sp* receiver functions from reflectivity synthetics (Park, 1996; Levin & Park, 1997) for each velocity model and measured the arrival time of the negative *Sp* phase using the weighted-average approach used in *Hopper and Fischer* (2018). This tests the consistency of our models with the original *Sp* receiver functions, as we did not include the shape of the waveforms in our inversion. The mean squared misfit normalized by the standard error, , to the *Sp* times is 0.72 (Supplementary Figure S6), which we judge acceptable (cf. Figure 7). The misfit to the *Sp* data is slightly greater in parts of the Snake River Plain and Maryvale fields than elsewhere, though the misfits are broadly acceptable everywhere except for directly beneath the Yellowstone hotspot (where exceeds 10 and the model should not be considered reliable). The for this test at the locations shown in Figure 11a are 1.3, 3.1, and 0.98, from the north-west to south-east, thus these values are each within twice the standard error and wide NVGs below sharply negative sub-moho gradients are plausible.

Diagram

Description automatically generated**Figure S6.** Fit of the travel time to the *S*-to-*p* converted phase predicted by our models to the times observed by *Hopper and Fischer*, 2018.

**Section 6. Predictions for slopes in *Vs* from changes in grain size**

In the main text, polygonal regions are given to mark the predictions for slopes in *Vs* from changes in grain size. These polygons bound the individual results given in Table S2. Since the effect of grain size saturates above ~1 cm, larger gradients in grain size do not always translate to larger gradients in *Vs* over the whole interval, and so the results are not linearly dependent and do not easily admit clear display in Figure 13c. Calculations that would predict grains below 1 mm at any depth are not included, and so there are fewer entries at smaller grain sizes.

**Table S2. Predictions for the slope in *Vs* from gradients in grain size**

|  |  |  |  |
| --- | --- | --- | --- |
| **With JF10** |  |  |  |
| **Mean grain size, mm** | **Grain size slope, mm/km** | ***Vs* below the NVG, km/s** | **Slope in *Vs*, (km/s)/km x 10-3** |
| 1 | 0 | 4.213 | 2.659 |
| 2 | 0 | 4.284 | 2.491 |
| 3 | 0 | 4.261 | 2.430 |
| 4 | 0 | 4.273 | 2.376 |
| 4 | 111 | 4.224 | 3.598 |
| 5 | 0 | 4.279 | 2.344 |
| 5 | 111 | 4.252 | 3.067 |
| 6 | 0 | 4.285 | 2.312 |
| 6 | 111 | 4.265 | 2.867 |
| 7 | 0 | 4.289 | 2.300 |
| 7 | 111 | 4.276 | 2.690 |
| 7 | 222 | 4.234 | 3.645 |
| 8 | 0 | 4.294 | 2.273 |
| 8 | 111 | 4.282 | 2.597 |
| 8 | 222 | 4.257 | 3.191 |
| 9 | 0 | 4.297 | 2.260 |
| 9 | 111 | 4.285 | 2.574 |
| 9 | 222 | 4.268 | 2.992 |
| 10 | 0 | 4.300 | 2.248 |
| 10 | 111 | 4.292 | 2.492 |
| 10 | 222 | 4.276 | 2.885 |
| 10 | 333 | 4.239 | 3.658 |
| **With YT16** |  |  |  |
| 1 | 0 | 4.128 | 2.747 |
| 2 | 0 | 4.242 | 2.613 |
| 3 | 0 | 4.283 | 2.546 |
| 4 | 0 | 4.321 | 2.475 |
| 4 | 111 | 4.162 | 6.553 |
| 5 | 0 | 4.343 | 2.427 |
| 5 | 111 | 4.257 | 4.828 |
| 6 | 0 | 4.363 | 2.380 |
| 6 | 111 | 4.298 | 4.169 |
| 7 | 0 | 4.373 | 2.356 |
| 7 | 111 | 4.332 | 3.587 |
| 7 | 222 | 4.196 | 6.665 |
| 8 | 0 | 4.389 | 2.310 |
| 8 | 111 | 4.354 | 3.289 |
| 8 | 222 | 4.270 | 5.173 |
| 9 | 0 | 4.397 | 2.288 |
| 9 | 111 | 4.363 | 3.210 |
| 9 | 222 | 4.309 | 4.459 |
| 10 | 0 | 4.404 | 2.266 |
| 10 | 111 | 4.381 | 2.957 |
| 10 | 222 | 4.332 | 4.127 |
| 10 | 333 | 4.211 | 6.624 |

**Section 7. Sensitivity to the shape of melting region**

In detail, the smooth surface-wave constraints and model parameterization likely limit our ability to precisely define the width of the melt zone. If the triangular distribution of melt described in the main text (that is, peak melt fraction occurs at the NVG with a linear decrease towards zero with increasing depth) is correct, our inversion will accurately reconstruct the resulting *Vs* profile. We show this in Figure S7 by creating a synthetic velocity profile and reducing the velocity beneath the NVG at 70 km depth with 0.25% of melt that tapers to 0% at 120 km. We invert the predicted data for this synthetic model in the same procedure for used for the test shown in Figure 6, and the inverted model agrees well with the input model (red lines in Figure S7). However, if melt were to accumulate more tightly beneath the NVG into a melt-rich layer, our inversion would produce a profile that issmoother than the true model (yellow and blue lines in Figure S7). As the thickness or total melt fraction in a narrow layer increase, the minimum *Vs* decreases and the gradient increases in the recovered models in ways that qualitatively captures the increasing volume of melt beneath the NVG, but not in a way that accurately describes the melt-migration dynamics. This limitation reflects the intrinsic sensitivity of the data.

A picture containing chart

Description automatically generated 