Structure of North American Mantle Constrained by Simultaneous Inversion of Multiple-Frequency \textit{SH}, \textit{SS}, and Love Waves

Yue Tian,\textsuperscript{1,2} Ying Zhou,\textsuperscript{3} Karin Sigloch,\textsuperscript{4} Guust Nolet,\textsuperscript{5} and Gabi Laske\textsuperscript{6}

\textsuperscript{1}Department of Geosciences, Princeton
Abstract. We simultaneously invert for the velocity and attenuation structure of the North American mantle from a mixed data set: $SH$-wave travel-time and amplitude anomalies, $SS$-wave differential travel-time anomalies, and Love-wave fundamental-mode phase delays. All data are measured for multiple frequency bands, and finite-frequency sensitivity kernels are used to explain the observations.

University, Princeton, New Jersey, USA.

2Now at Chevron Energy Technology Company, San Ramon, California, USA.

3Department of Geosciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA.

4Department of Earth and Environmental Sciences, Ludwig-Maximilians-Universität, Munich, Germany.

5Géosciences Azur, Université de Nice/Antipolis, Sophia Antipolis, France.

6Institute of Geophysics and Planetary Physics, U.C. San Diego, La Jolla, California, USA.
In the resulting $SH$ velocity model, a lower mantle plume is observed to originate at about 1500 km depth beneath the Yellowstone area, tilting about 40° from vertical. The plume rises up through a gap in the subducting Farallon slab. The $SH$ velocity model confirms high-level segmentation of the Farallon slab, which was observed in the recent $P$ velocity model [Sigloch et al., 2008]. Attenuation structure is resolvable in the upper mantle and transition zone; in estimating it, we correct for focusing. High correlation coefficients between $\delta \ln V_S$ and $\delta \ln Q_S$ under the Central and Eastern U.S. suggest one main physical source, most likely temperature. The smaller correlation coefficients and larger slopes of the $\delta \ln Q_S - \delta \ln V_S$ relationship under the Western U.S. suggest an influence of non-thermal factors such as the existence of water and partial melt. Finally, we analyze the influence of the different components of our data set. The addition of Love-wave phase delays helps to improve the resolution of both velocity and attenuation, and the effect is noticeable even in the lower mantle.
1. Introduction

Recent progress in high-resolution tomography has led to a clearer view of subduction history and slab geometry: subducted slabs may be torn apart and segmented [Nolet, 2009]. The Farallon slab shows evidence of segmentation under North America. Various hypotheses for the deformation of the Farallon slab have been proposed, including “slab buckling” [Humphreys, 1995] and “slab roll back” followed by detachment [Van der Lee and Nolet, 1997]. Sigloch et al. [2008] present the clearest evidence so far for high-level segmentation of the Farallon slab, an observation that was supported by P and S wave studies by Roth et al. [2008], Burdick et al. [2009], Tian et al. [2009], Xue and Allen [2010], and Obrebski et al. [2010]. The way in which the Farallon slab broke up into fragments has direct effects on subduction dynamics, by controlling the width of the slab and thus the trench migration rates [Schellart et al., 2007] which in turn influences the ability of the slab to enter the lower mantle [Goes et al., 2008]. The observed tears in the Farallon slab allow for mantle upwellings, and thus may contribute to explain the widespread and complex patterns of magmatism in the Western U.S. [Smith and Luedke, 1984]. Tomographic studies of the mantle beneath North America are necessary to examine how the Farallon slab deformed and fragmented. Our study serves this purpose.

Comparisons between attenuation and velocity heterogeneities provide further insight into the physical state of the mantle, because different physical sources for mantle heterogeneity, such as temperature [e.g., Karato, 1993; Jackson et al., 2002], water content [e.g., Karato, 2006], partial melt [e.g., Jackson et al., 2004; Faul et al., 2004], chemical composition [e.g., Lee, 2003], and grain size [e.g., Faul and Jackson, 2005], give differ-
ent attenuation-velocity relationships. There have been few tomographic studies on the attenuation structure of North American mantle. Lawrence et al. [2006] mapped the two-dimensional $P$ and $S t^*$ residuals under North America but with no depth constraints. Yang and Forsyth [2008] simultaneously inverted for one-dimensional $V_S$ and $Q_μ$ structure of the upper mantle beneath southern California. Hwang et al. [2009] estimated the two-dimensional $P$-wave $t^*$ structure under North America with no depth constraints. The first joint inversion for three-dimensional $V_P$ and $Q_P$ structure under North America was conducted by Sigloch et al. [2008]. The first joint inversion for three-dimensional $V_S$ and $Q_S$ structure under the Western U.S. was carried out by Tian et al. [2009], using multiple-frequency $SH$-wave travel times and amplitudes.

The current study extends the work by Tian et al. [2009] by including $SS$-wave delays and Love-wave phase data and imaging the mantle under North America. We simultaneously estimate the $SH$-wave velocity and attenuation structure under North America, taking advantage of three recent developments in seismology: (1) the large volume of data provided by the USArray which has densely sampled the Western U.S.; (2) finite-frequency sensitivity theory and the fast computation of the sensitivity kernels for body-wave travel times [Dahlen et al., 2000], focusing effects [Dahlen and Baig, 2002], attenuation [Nolet, 2008, Section 8.5], and for surface waves [Zhou et al., 2004]; (3) accurate techniques to measure frequency-dependent body-wave travel times and amplitudes [Sigloch and Nolet, 2006] and surface-wave phase delays [Laske and Masters, 1996]. The velocity model is used to study the mantle dynamics under North America. The velocity and attenuation models are compared to gain insight into the sources of mantle heterogeneities. Effects of adding new types of data, especially the Love-wave phase delays, are also investigated.
2. The Data Set

We use a mixed data set consisting of four different types of data, which are all measured at multiple frequencies from the transverse component: SH-wave travel-time anomalies, SH-wave amplitude anomalies, SS-wave differential travel-time anomalies, and Love-wave fundamental-mode phase delays. This data set is an extension of the SH data set built by Tian et al. [2009], to which we add: (1) Love-wave phase delays, (2) SS-wave differential delays, (3) more SH-wave measurements from most recent USArray stations. Love waves are particularly helpful. They provide the depth resolution which body waves lack, because in the shallow region the teleseismic body-wave paths are almost vertical. SS waves help to improve the resolution with reflection points covering regions not sampled by direct SH waves.

2.1. SH-Wave Travel-Time and Amplitude Anomalies

We construct a global data set of SH-wave travel-time anomalies

\[ \delta T_S = T_{S}^{\text{obs}} - T_{S}^{\text{pre}} \]  

and amplitude anomalies

\[ \delta \ln A_S = A_{S}^{\text{obs}} / A_{S}^{\text{pre}} - 1 \]  

(with ‘obs’ for observed and ‘pre’ for predicted data) in five frequency bands with center periods of 40 s, 20 s, 10 s, 5 s, 2.5 s, respectively (see Tian et al. [2009] for the passband filter responses). The measurements are obtained through cross-correlation, using the technique developed by Sigloch and Nolet [2006]. These measurements are deviations from predictions computed for the IASP91 \( V_S \) model [Kennett and Engdahl, 1991] extended with the PREM \( Q_S \) model [Dziewonski and Anderson, 1981]. Details of how to construct
the data set, and discussions of the frequency dependence and spatial distribution of the
data set are described by Tian et al. [2009]. In this paper, we update the SH-wave
data set with more recent USArray data until October 2008. The current SH data set
consists of 26296 wave paths, producing 98371 acceptable travel-time measurements in
the five frequency bands, and 74665 acceptable amplitude measurements in the three low-
frequency bands. For an idea of the SH data coverage in North America, see Figures 5c–e
and 6c–e.

The use of cross-correlation for the estimation of delays allows us to exploit the different
sensitivity of high and low frequencies. A small heterogeneity may not show up in a low-
frequency delay because of the effects of wavefront healing [Nolet and Dahlen, 2000], but
still be visible at high frequency. Delay dispersion thus provides information on the size of
velocity heterogeneity. The histograms of delays in Figure 1 show a frequency dependence
of the data: the median of travel-time anomalies and the median and standard deviation
of amplitude anomalies increase with frequency. To interpret the frequency dependence
of delays and amplitude anomalies, we use finite-frequency sensitivity kernels.

2.2. SS-Wave Differential Travel-Time Anomalies

We measure the SS-wave differential travel-time anomalies

\[ \delta(T_{SS} - T_S) = (T_{SS} - T_S)^{obs} - (T_{SS} - T_S)^{pre} \]  

in the same five frequency bands as for SH waves. The observed differential travel times
\((T_{SS} - T_S)^{obs}\) are measured by cross-correlating the observed SS waveform with the pre-
dicted SS waveform. The predicted SS waveform is computed by Hilbert transforming the
observed SH waveform, applying the attenuation operator to account for the difference
in attenuation along the SS and SH paths, and multiplication by $-1$ to account for the
development at the free surface. Both the observed and predicted SS waveforms are filtered
in each of the five frequency bands, and the cross-correlation is done in each band.

We measure any SS wave that has a corresponding acceptable measurement of $\delta T_S$ (see
Section 2.1) and that has an epicentral distance between $60^\circ$ and $88^\circ$. The quality of
acceptable $\delta T_S$ guarantees reliable estimates of the SH-wave window and source depth,
which are used to compute the predicted SS waveform. The lower limit of $60^\circ$ is to exclude
triplications of SS, and the upper limit of $88^\circ$ is to exclude D" and CMB diffracted SH
waves [Paulssen and Stutzmann, 1996]. Measurements are deemed acceptable by visual
inspection of the waveform fits after cross-correlation. We impose a lower limit of $0.9$ on
the cross-correlation coefficient. The two high-frequency bands (with center periods of 5
s and 2.5 s) are not used for tomography because the $(T_{SS} - T_S)_{\text{obs}}$ uncertainty estimated
from pairs of closely located events is much larger in these two bands ($\sim 1.8$ s) than
in the other frequency bands ($\sim 0.8$ s). The final SS-wave data set consists of 18919
measurements from 8270 wave paths in three frequency bands (with center periods of 40
s, 20 s, and 10 s). The SS data coverage in North America is shown in Figures 5b and 6b.

To examine the consistency of the measurements, Figure 2 plots $\delta(T_{SS} - T_S)$ at each
reflection point for unfiltered (“broadband”) data, which reflects anomalies in the shallow
subsurface beneath the reflection point. The map shows distinct large-scale patterns, and
they correlate with tectonic structure. For example, negative anomalies at the Canadian
shield and the Siberian shield, positive anomalies in western North America and western
Central America. The patterns are similar to those observed by Woodward and Masters
[1991] and Reid et al. [2001].
After applying crustal, ellipticity, and elevation corrections, the histogram of SS differential delays in each frequency band is plotted in Figure 3a–c. The histograms are skewed towards negative anomalies. One possible cause of this asymmetry is that more reflection points hit the region with fast shallow structure (see Figure 2). Figure 3d examines the dispersion of SS differential delays with respect to the lowest-frequency band. The dispersion is of the order of 1 s, which is of the same order as the uncertainty in SS differential delays. If the dispersion is completely due to random noise, one expects to observe a zero median and the same standard deviation in the 20-s and 10-s bands. The nonzero medians in both bands and the clear trend of larger dispersion with increasing frequency difference suggest that the observed dispersion is at least partially due to the finite-frequency effect, though at low frequency crustal effects may play a role [Ritsema et al., 2009].

2.3. Love-Wave Fundamental-Mode Phase Delays

For Love waves, we use the fundamental-mode minor-arc phase delays, measured at 11 frequencies (5, 6, 7, · · · , 15 mHz) with a multitaper technique [Laske and Masters, 1996]. This data set consists of 19485 phase delays from 1778 wave paths. It is part of the global surface-wave data set used by Zhou et al. [2006]. Figures 5a and 6a show the Love-wave data coverage in North America. The original phase delays used by Zhou et al. [2006] are for the reference model 1066A [Gilbert and Dziewonski, 1975]. To be consistent with the body waves, the phase delays are corrected to be for the reference model IASP91 [Kennett and Engdahl, 1991]. After crustal corrections, the histograms of the relative phase delays \( \delta \ln \phi = \delta \phi / \phi \) are plotted in Figure 4.
2.4. Data Corrections and Uncertainties

For body waves, when computing \( SH \) and \( SS \) travel times, the effects of ellipticity, crustal structure, and station elevations are taken into account [Tian et al., 2007]. The crustal corrections are computed using ray theory and the three-dimensional crust model CRUST2.0 [Bassin et al., 2000]. According to Ritsema et al. [2009] and Obayashi et al. [2004], the crustal correction difference \( \Delta(\delta t_C) \) between ray-theoretical and finite-frequency calculations are not negligible for long-periods waves. This is a potential source of uncertainties for \( SH \) and \( SS \) travel times. Specifically, for \( SH \) waves, Ritsema et al. [2009] show that the ratio between \( \Delta(\delta t_C) \) and the RMS of measured travel-time anomalies is 0.144, 0.146, and 0.184 for 40-s, 20-s, and 10-s period respectively. For \( SS \) waves, the ratio is 0.069, 0.110, and 0.055 for 40-s, 20-s, and 10-s period respectively. The uncertainties in \( \delta T_S \), \( \delta \ln A_S \), and \( \delta(T_{SS} - T_S) \) are estimated based on: (1) difference between two measurements from two closely located events at the same station; (2) variation of one measurement with cross-correlation window length; (3) quality of the waveform fit. As a result, the \( \delta T_S \) uncertainties range from 0.6 s to 1.2 s, the \( \delta \ln A_S \) uncertainties range from 0.12 to 0.22 (0.98 dB to 1.73 dB or -2.16 dB to -1.11 dB), and the \( \delta(T_{SS} - T_S) \) uncertainties range from 1.1 s to 1.7 s.

For Love waves, the crustal structure has a very large effect on phase delays [e.g., Zhou et al. 2006]. Crustal corrections are computed using ray theory and the model CRUST2.0 [Bassin et al., 2000], and are applied to the phase delay measurements before inversion. We adopt the phase delay uncertainties estimated by Zhou et al. [2006] based on pairs of closely located events, and slightly adjust them in order to control the relative importance of different data types in the joint inversion. The uncertainty for the relative phase delay
δφ/φ ranges from 0.74% to 0.93%. Regional crustal models in the Western U.S. have been developed based upon receiver function analysis and Pn tomography of USArray data, and discrepancies exist among models developed by different research groups [e.g., Wilson et al., 2010; Buehler and Shearer, 2010]. However, details in regional crustal structure do not have significant effects on Love wave phase delays at the periods we are interested in. For example, a 5-km difference in crustal thickness over a propagation distance of 40° would introduce a δφ/φ difference of less than 0.22% at 5-mHz frequency and a δφ/φ difference of less than 0.59% at 15-mHz frequency. Both are smaller than the estimated uncertainty of δφ/φ.

3. The Joint Finite-Frequency Tomographic System

3.1. The Continuous Linear System

In finite-frequency tomography, the linear inverse problem takes four forms corresponding to our four data types:

1. For SH-wave travel-time anomalies, the linear inverse problem is

\[ \delta T_{S}(s(\omega)) = \int \int K_{T_{S}}^{V}(s(\omega), r) \delta \ln V_{S}(r) \, d^{3} r + \delta T_{0} + \delta X_{0} \cdot \frac{\hat{p}}{V_{S}|_{X_{0}}}, \]  

with \( \delta \ln V_{S} \) representing velocity heterogeneities, and the sensitivity kernel \( K_{T_{S}}^{V} \) describing how the velocity heterogeneity affects SH-wave travel times [Dahlen et al., 2000]. Here \( s(\omega) \) is the frequency content of the observed waveform that is used to measure \( \delta T_{S} \). For our data set, we assume that this is equal to the frequency response of the passband filter. The last two terms are correction terms: \( \delta T_{0} \) represents the origin-time correction, vector \( \delta X_{0} \) represents the hypocenter correction, and \( \frac{\hat{p}}{V_{S}|_{X_{0}}} \) represents the slowness vector at hypocenter \( X_{0} \).
2. For **SH**-wave amplitude anomalies, the linear inverse problem is

\[ \delta \ln A_S(s(\omega)) = \iiint K_*^{AS}(s(\omega), r) \, \delta \ln V_S(r) \, d^3r + \iiint K_*^{AS}(s(\omega), r) \, \delta \ln Q_S^{-1}(r) \, d^3r \]

\[ + \delta \ln A_0 + \delta \ln A_r. \]  

(5)

Here \( \delta \ln Q_S^{-1} \) represents attenuation heterogeneities. The sensitivity kernel \( K_*^{AS} \) describes how the velocity heterogeneity affects **SH**-wave amplitudes [Dahlen and Baig, 2002], and \( K_*^{AS} \) describes how the attenuation heterogeneity affects **SH**-wave amplitudes [Nolet, 2008, Section 8.5]. \( \delta \ln A_0 \) is the source correction term accounting for the scalar moment errors and the uncertainty in source time function scaling in the measurement process [Sigloch and Nolet, 2006], and \( \delta \ln A_r \) is the station correction term accounting for sediment effects and adjustment of the instrument magnification.

3. For **SS**-wave differential travel-time anomalies, the linear inverse problem is

\[ \delta (T_{SS} - T_S)(s(\omega)) = \iiint (K_*^{TSS}(s(\omega), r) - K_*^{TS}(s(\omega), r)) \, \delta \ln V_S(r) \, d^3r, \]  

(6)

with \( K_*^{TSS} \) and \( K_*^{TS} \) representing the **SS** and **SH** travel-time-velocity sensitivities, respectively [Dahlen et al., 2000].

4. For Love-wave phase delays, the linear inverse problem is

\[ \delta \ln \phi(\omega) = \iiint K_*^{\phi}(\omega, r) \, \delta \ln V_S(r) \, d^3r + \frac{c \delta T_0}{\Delta}. \]  

(7)

The sensitivity kernel \( K_*^{\phi} \) describes how the velocity heterogeneity affects Love-wave phase delays [Zhou et al., 2004], and \( \omega \) is the angular frequency at which \( \delta \ln \phi \) is measured. In the last term, \( \delta T_0 \) is the origin-time correction, \( c \) is the phase velocity in km/s, and \( \Delta \) is the epicentral distance in km.

3.2. The Discrete Linear System
We use a model parameterization in the form of a tetrahedral mesh with linear interpolation in between grid nodes. The general philosophy of adaptive tetrahedral meshing is described by Nolet [2008, Section 12.2]. Here we use the same tetrahedral mesh as in Sigloch et al. [2008] and Tian et al. [2009]. Beneath the U.S., the grid spacing of the mesh is about 70 km in the upper mantle, and increases to roughly 200 km at 660 km depth. Outside of the U.S., the grid spacing is about 200 km in the upper mantle and transition zone. Below 660 km, the grid spacing increases linearly to about 300 km at 2000 km depth. The mesh is produced with the Matlab software distmesh developed by Persson and Strang [2004].

With source and receiver correction terms and regularization, the complete discrete linear system of the tomographic problem can be written as

\[
\begin{pmatrix}
K_{V}^{T_{S}} & 0 & K_{C}^{T_{S}} & 0 & 0 \\
K_{V}^{A_{S}} & K_{Q}^{A_{S}} & K_{C}^{A_{S}} & 0 & 0 \\
K_{V}^{T_{SS} - T_{V}} & 0 & 0 & 0 & 0 \\
\epsilon_{1} W & 0 & 0 & 0 & 0 \\
0 & \epsilon_{1} W & 0 & 0 & 0 \\
0 & 0 & \epsilon_{1} W & 0 & 0 \\
0 & 0 & 0 & \epsilon_{1} W & 0 \\
\epsilon_{2} R W & 0 & 0 & 0 & 0 \\
0 & \epsilon_{2} R W & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
\delta T_{S} \\
\delta \ln A_{S} \\
\delta (T_{SS} - T_{S}) \\
\delta \ln V_{S} \\
\delta \ln Q_{S}^{-1} \\
C_{T_{S}}^{A_{S}} \\
C_{A_{S}}^{A_{S}} \\
C_{\phi}^{A_{S}} \\
\delta \ln \phi
\end{pmatrix} = \begin{pmatrix}
\delta V_{S} \\
\delta (T_{SS} - T_{S}) \\
\delta \ln A_{S} \\
\delta \ln \phi
\end{pmatrix}.
\] (8)

The first four rows correspond to equations (4)–(7). The vector on the right-hand side contains in sequence the four types of data, as described in Section 2. On the left-hand side, besides the velocity and attenuation heterogeneities, the model vector includes various correction terms: \(C_{T_{S}}^{A_{S}}\) represents hypocenter and origin-time corrections for \(SH\)-wave travel times (see equation (4)); \(C_{A_{S}}^{A_{S}}\) represents receiver and source corrections for \(SH\)-wave amplitudes (see equation (5)); \(C_{\phi}^{A_{S}}\) represents origin-time corrections for Love-wave phase delays (see equation (7)). The body-wave finite-frequency sensitivity kernels
\( K_{V}^{T_S}, K_{V}^{A_S}, K_{Q}^{A_S}, K_{V}^{T_S} \) are computed using the software by Tian et al. [2007a,b]. The Love-wave finite-frequency sensitivity kernels \( K_{V}^{\phi} \) are computed using the software by Zhou et al. [2006]. \( K_{C}^{T_S}, K_{C}^{A_S}, K_{C}^{\phi} \) are ad hoc matrices for corrections \( C^{T_S}, C^{A_S}, C^{\phi} \) (see equations (4), (5), (7)). \( W \) is a diagonal weighting matrix, with its diagonal element proportional to the tetrahedral volume associated with the corresponding grid point, and \( \epsilon_1 \) is the norm damping parameter. \( R \) is the Laplacian roughening operator [Nolet, 2008, Section 14.5] and \( \epsilon_2 \) is the smoothing parameter. Equation (8) shows that velocity and attenuation have coupled effects on \( SH \)-wave amplitudes, and \( SS \) and Love waves provide extra independent constraints on velocity. Our joint inversion interprets these data simultaneously.

As a measure of how strongly a particular tetrahedral volume in the Earth is sensed by the combined set of kernels, we compute the "kernel density", defined as

\[
D_j = \frac{\sum_i K_{ij}}{C V_j},
\]

(9)

where \( K \) stands for the tomographic matrices in equation (8), summation \( i \) is over all data, \( V_j \) is the average volume of the tetrahedron containing grid point \( j \), and \( C \) is a scaling constant such that the maximum \( D_j \) is 1. The factor \( 1/V_j \) removes the effect of nonuniform grid spacing. Figures 5–6 plot \( \log(D_j) \) for all the five sensitivity kernels in equation (8), at 100 and 600 km depth, respectively. Figures 5c–e and 6c–e show the kernel densities for \( SH \) waves. The dense coverage of the USAArray is reflected in the large values of \( SH \)-wave kernel densities in the Western U.S. For \( SH \)-wave amplitudes, with the same ray coverage, the sensitivity to attenuation is about 1.5 orders of magnitude weaker than that to velocity. As we go deeper, the body-wave sensitivity becomes wider although the strength slightly decreases near the geometrical rays. Figures 5a and 6a show
that Love-wave sensitivity is concentrated near the surface. Love and SS waves provide constraints on SH velocity outside of the U.S. region, where SH waves have poor coverage.

Within the U.S. region, the Love-wave and SS-wave sensitivity is at least one order of magnitude smaller than that of SH waves, because of the small number of measurements for Love and SS waves. Nonetheless, they provide extra constraints independent of SH waves: Love waves are more sensitive to vertical gradients in velocity than SH waves.

3.3. The Inversion Scheme

The joint linear system (8) has four different types of data with different orders of magnitude. Following a Bayesian philosophy [Nolet, 2008, Section 14.5], we scale the data with their estimated measurement uncertainties (see Section 2.4), and scale the model parameters with their “prior uncertainties”, which are estimated from results of existing tomographic studies. The model parameter “prior uncertainties” used in the final inversion are: 0.01 for \( \delta \ln V_S \), 0.2 for \( \delta \ln Q_S \), 30 km for SH-wave hypocenter corrections, 1 s for SH-wave origin-time corrections, 0.1 for SH-wave amplitude event corrections, 0.15 for SH-wave amplitude station corrections, and 1 s for phase-delay origin-time corrections.

These estimates of both data and model uncertainties carry a subjective uncertainty and were in effect slightly adjusted within their margin of error as we gained experience about data compatibilities during early inversion runs. Before inversion, the data and model parameters in equation (8) are divided by their uncertainties, and the kernel matrices (all the \( K \)'s in equation (8) ) are scaled correspondingly.

We expect a high resolution beneath the dense USArray, but have meshed the entire globe in order to absorb delays or focusing effects acquired along wave paths outside of North America. No unique solution exists and the linear system (8) is necessarily regular-
ized. We aim for a model which interprets the data within their margin of uncertainties (satisfying rows 1–4 of (8)), while staying close to the reference model (satisfying rows 5–9 of (8)), and remaining as smooth as possible (satisfying rows 10–11 of (8)). This trade-off between fitting the data and regularizing the model is controlled by $\epsilon_1$ and $\epsilon_2$.

The scaled and regularized linear system (8) is solved by a parallel version of LSQR [Paige and Saunders, 1982; Nolet, 1987] on a cluster. Outliers with residuals larger than three standard deviations after a first run with negligible norm damping and smoothing are removed.

In order to quantitatively assess the trade-off between fitting the data and regularization, we use $\chi^2/N$ as a measure for the misfit of the observed data to the values predicted by finite-frequency theory:

$$\frac{\chi^2}{N} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{d_i - \sum_j K_{ij}m_j}{\sigma_i} \right)^2,$$

where the summation over $i$ is over the first four matrix rows in (8), $K_{ij}$ is the submatrix formed by those rows, $\sigma_i$ is the estimated uncertainty of datum $d_i$, and $N$ is the total number of data. If $\sigma_i$ is correctly estimated, $\chi^2/N$ should be close to 1 for a model that fits the data, but which is not forced to fit the data closer than one standard deviation on average. Meanwhile, we use the L2-norm of the model vector $\|m\|$ as a measure for the deviation from the reference model, and use the L2-norm ratio $\|Rm\|/\|m\|$ as a measure for the roughness of the model, where $m$ is the model vector, $\delta \ln V_S$ or $\delta \ln Q_S^{-1}$ in equation (8), and $R$ is the Laplacian roughening operator [Nolet, 2008, Section 14.5].

Figure 7 shows the trade-off between fitting the data and regularizing the model, for velocity and attenuation respectively. The preferred model has $\chi^2/N = 0.928$, and is near the corner of the trade-off curves. The norms of the scaled $\delta \ln V_S$ and $\delta \ln Q_S^{-1}$ are
of the same order, suggesting that the estimates for prior uncertainties in velocity and attenuation are reasonable.

4. The $SH$ Velocity Model of the Mantle beneath North America

In this section, we discuss several interesting features of our $SH$ velocity model of the mantle, which is the deviation from the IASP91 $S$ velocity model [Kennett and Engdahl, 1991]. A complete catalog of the velocity maps is given in Figures 19–21 (right columns).

4.1. Subduction

The fast velocity anomalies in our model confirms the two-stage subduction under North America and the existence of tears in the Farallon slab observed by Sigloch et al. [2008]. Figure 8 shows the whole subduction system under North America. Figures 9–10 show the subduction under the Western U.S. and the tearing of the slab. To facilitate comparison with the earlier $P$-wave inversions, we use the same notation ($S1$, $N1$, $S2$, $N2$, $F1$, $F2$, $SG$) as Sigloch et al. [2008] to identify anomalies.

Fast anomalies are observed at 100–200 km depth under the Cascades (Figure 19, right column; $S0$, $N0$ in Figures 9–10), indicating the most recently subducted Juan de Fuca Plate. It is subducting at a steep angle ($S0$ in Figure 8AA’). The Juan de Fuca slab continues eastward, penetrating the transition zone ($S1$, $N1$ in Figures 8–10), and reaching the lower mantle ($S2$, $N2$ in Figures 8–9). Under the Western U.S. down to $\sim$ 800 km depth, no large fast anomaly is observed south of $\sim$ 37°N ($SW$ in Figures 8–10), which is the present Mendocino edge of the Juan de Fuca slab. This region with absence of subduction is known as the slab window, produced by the breakup and movement of the Farallon slab in the last $\sim$ 28 My [Atwater and Stock, 1998]. Although Schmandt and
Humphreys [2010a] found small-scale high velocity anomalies possibly representing the
slab coming to rest in the transition zone, our tomographic image indicates that the last
large (∼600 km) piece of the Farallon Plate south of the Mendocino Triple junction has
sunk to at least 900 km depth. If the slab window formed at ∼28 Ma [Atwater and Stock,
1998], this implies a vertical sinking rate of 3.2 cm/yr, which is comparable to the speed
of the Farallon plate at ∼28 Ma [Engebretson et al., 1985].

Under eastern North America, large volumes of fast anomalies (Figures 20-21, right
columns; F1, F2 in Figure 8) are observed parallel to the young Juan de Fuca slab de-
scribed above. They extend as far south as at least 25°N, occupying the mantle beneath
eastern North America down to at least 1400 km. This probably represents the ancient
Farallon slab, and it is well separated from the Juan de Fuca slab, as can be seen from
Figure 8: the western tip of F1 is riding above the eastern tips of S2 and N2 with a
gap in between. The coexistence of the western and eastern subduction systems was first
observed by Sigloch et al. [2008], and interpreted in terms of a big break at 50–40 Ma
between F1 and S1. Consistent with their $V_p$ model, we observe a flat F1 in the transition
zone (Figure 8AA’), and a relatively steeper subduction angle of S1, which makes it easier
for S1 to penetrate the transition zone.

Taking advantage of the dense USArray coverage, we are able to obtain a high resolution
under the Western U.S. and thus study the detailed structure of the Juan de Fuca slab.
The most striking feature is the slab gap (SG in Figures 9–10), which is a continuous
1600-km-long trail void of fast anomalies, and which divides the Juan de Fuca slab into
a northern part (N0, N1, N2) and a southern part (S0, S1, S2). It starts at ∼120 km
depth near (45°N, 238°E), extends northeastward as it sinks deeper, and ends at ∼800
km depth near (47°N, 254°E) (Figure 20, right column, 800 km). Because there is no surface evidence showing the absence of subduction near (45°N, 238°E), the slab gap is more likely to have formed at ~ 120 km depth rather than at surface. We exclude the kinematic segmentation of the monolithic Farallon plate starting at ~ 30 Ma as the cause of the slab gap. Most of this breakup of the Farallon plate occurred south of the Mendocino fracture zone, and the plate to the north (the Juan de Fuca plate) remained relatively intact [Atwater and Stock, 1998]. The slab gap revealed by our velocity model is inside the Juan de Fuca plate, so it is less likely to relate to the breakup of the monolithic Farallon plate. In addition, the slab gap reaches only 800 km depth, which suggests that the slab gap is younger than 30 My using the Farallon plate velocity by Engebretson et al. [1985]. One possible cause of the slab gap could be the different subduction angles of the northern and southern parts of the Juan de Fuca plate, as shown in Figure 9, S1 enters the transition zone at a further east location than N1. Our velocity model shows that the subduction angle is ~ 60° (from surface) for S1, and ~ 70° for N1, and the thickness of S1 and N1 is ~ 70km. A back-of-the-envelope calculation with this geometry suggests that S1 and N1 start to diverge (no overlap) at ~ 350 km, which is much deeper than the observed 120 km depth. This indicates that there are other causes contributing to the formation of the slab gap in the meantime, such as the interaction between the Yellowstone plume and the slab [Obrebski et al., 2010]. This slab gap was first discovered by Sigloch et al. [2008]. In their P velocity model, the slab gap extends longer (2500 km long) and deeper (1200 km deep) than in our model. The slab gap is also recognizable in other recent tomographic models [e.g., Obrebski et al., 2010; Schmandt and Humphreys, 2010b], although it is not discussed in these studies. The slab hole at shallow depth near
45°N under the Cascades that was observed in several recent studies [Roth et al., 2008; Tian et al., 2009; Burdick et al., 2009; Schmandt and Humphreys, 2010b] overlaps the upper tip of the slab gap, and thus is actually the top part of the slab gap.

Evidence for more breaks or detachment of the slab is observed. At 700–1200 km depth, F2 is divided into two blob-like segments, along an east-west oriented break near ∼ 40°N (Figure 8AA’ and Figures 20–21, right columns). The \( V_P \) model [Sigloch et al., 2008] shows a break at similar location and depth, but extending broader in the north-south direction and shorter in the east-west direction than in our \( V_S \) model. S2 has a blob-like tail down to ∼ 1600 km depth (Figure 8). Blob-like slabs in the lower mantle are also observed by seismic tomography under northern Kuril and Mariana [Fukao et al., 2001]. Such droplet-like slabs are modeled by numerical simulation with a large viscosity contrast across the 660-km discontinuity [Tagawa, 2007]. The droplet-like shape suggests strong deformation and perhaps detachment of the slab as it penetrates the 660-km discontinuity. The discontinuity between S0(N0) and S1(N1) (Figures 8AA’ and 9) is the evidence of a further small-scale tear in the Juan de Fuca slab.

Different hypotheses explaining the Farallon slab segmentation have been proposed to explain the post-Laramide magmatism in the Western U.S. The “slab buckling” model [Humphreys, 1995] predicts east-west oriented Farallon slab remnants. Anomalies S1 at ∼ 40°N and N1 at ∼ 47°N can not be the north and south sides of the buckle which is now near 36°N. The “slab roll back and detachment” model [Van der Lee and Nolet, 1997] and the “two-stage subduction” model [Sigloch et al., 2008] predict north-south oriented Farallon slab remnants. Our \( SH \) velocity model largely satisfies predictions from the subduction history of the Farallon slab proposed by Sigloch et al. [2008], such as the
separation between anomalies S1 and F1, possibly resulting from a two stage subduction. Meanwhile we observe blob-like slab segments in the lower mantle and other smaller-scale slab tears, all indicating high-level segmentation of the Farallon slab.

The subduction-related features discussed above are reasonably well resolved, including the subduction under the Cascades (Figure 22, 100–200 km), the slab window (Figure 22, 100–800 km), the coexisting two subduction systems (Figure 23AA’), the slab gap under the Western U.S. (Figure 22, 200–800 km, Figure 23BB’), the break in F2 (Figure 22, 800–1200 km), the droplet-like geometry of S2 and F2 (Figure 23AA’), and the discontinuity between S0 and S1 (Figure 23AA’).

4.2. The Yellowstone Plume

A strong slow anomaly (Y0) with large velocity gradient is observed under the Yellowstone Caldera (Figures 11a and 12). It has a diameter of \( \sim 200 \) km and reaches 200 km depth. Y0 connects to a 200-km wide slow anomaly (Y1) in the transition zone, which is centered at \( \sim 1^\circ \) north of Y0 (Figure 12AA’). To the southwest, Y0 abuts a belt of strong slow anomalies (SR0) under the eastern Snake River Plain (Figures 11a and 12BB’). Down in the lower mantle, a plume-shaped 300-km wide conduit (Y2) is observed with its top directly under Y1. The plume comes from south, tilting \( \sim 40^\circ \) from vertical, reaches as deep as 1500 km, but spreads out at 700–1100 km depth. Figure 23 CC’–DD’ shows that these velocity features are well resolved. A tilting plume under the Yellowstone that extends to the lower mantle is also observed by Obrebski et al. [2010] and Schmandt and Humphreys [2010b], but the plume comes from southwest in their models. The \( V_S \) model of Obrebski et al. [2010] also shows that the plume spreads out at 600–900 km depth, a feature quite similar to that in our model. The image of a tilting whole-mantle plume
is quite different from previous tomographic images, especially in the lower mantle. Regional studies have shown an upper-mantle plume from northwest with a smaller tilting angle [Yuan and Dueker, 2005; Waite et al., 2006], and other studies have revealed no plume-like features in the lower mantle [Montelli et al., 2006; Sigloch et al., 2008; Burdick et al., 2009].

The dimensions of Y0, Y1, and Y2 agree well with the expected plume radius from other studies [e.g., Montelli et al., 2004; Steinberger and Antretter, 2006]. The large southwards tilting angle of Y2 is surprising because it would require a very strong mantle wind. Both Steinberger and O’Connell [1998] and Steinberger [2000] predict indeed a southward flow in the mid-mantle beneath the Yellowstone from numerical modeling of mantle flow. The spreading of SR2 can be explained either if the 660-km discontinuity acts as a barrier, as is sometimes observed for plumes [Nolet et al., 2006], or if the slab fragment S1 acts as a barrier for the upring Y2 (Figure 11b). Y2 might have distorted in two ways in response to the barrier. Part of Y2 might have navigated its path around S1 and found its way up through the slab gap, and the other part of Y2 might have smeared southwestward along the base of S1 and formed SR2. Though much of this explanation is speculative, it is clear that we witness a complex interaction between upwellings and downwellings in this part of the mantle.

Much of Y0 and the eastern edge of SR0 have \( \delta \ln V_S < -5\% \) above 120 km (Figure 12BB’), and the peak anomaly is \(-7\%\) at \(\sim 50\) km depth directly under the Yellowstone Caldera. According to Cammarano et al. [2003] and Goes and Govers [2000], \( \delta \ln V_S < -5\% \) requires a thermal anomaly of \( \delta T > 250\) K, and the \(-7\%\) velocity anomaly under the caldera requires an \( \delta T \) of at least 350 K if it is a purely thermal effect. Tomographic
studies with data from local transportable array experiments in the Yellowstone region [Schutt et al., 2008; Schmandt and Humphreys, 2010b] show even slower velocities under the Yellowstone Caldera and the eastern Snake River Plain, which indicate even higher temperatures. These high temperatures reach the solidus of peridotite at 50–100 km depth [Goes and van der Lee, 2002], and thus may cause partial melt in these regions, especially under the caldera. The partial melt is probably responsible for the magmatism in the Yellowstone and eastern Snake River Plain.

4.3. The New England Slow Anomaly

Slow velocity anomalies up to −3% are observed under the New England Province from the surface to the bottom of the transition zone (Figure 13AA’), and they are well resolved (Figure 23 EE’). Low velocity in this region down to the transition zone has been reported by Van der Lee and Nolet [1997], and has been confirmed by recent body-wave studies [Sigloch, 2008; Burdick et al., 2009] and surface-wave studies [Bedle and van der Lee, 2009]. According to Cammarano et al. [2003], under pure thermal effect, a −3% velocity anomaly at 100 km depth and a −1% velocity anomaly at 500 km depth require a thermal anomaly of at least 150 K and 220 K, respectively. This reaches the lower limit of the central temperature anomaly of an active mantle plume [e.g., Nolet et al., 2006; Steinberger and Antretter, 2006]. On the other hand, the New England Province has been tectonically inactive for the last ∼ 100 My. The Montegerian hotspot passed through this region and created the Montegerian Hills at about 125 Ma [Sleep, 1990]. Hence the thermal anomaly in this region should be weak at present, if not zero. Therefore, the large amplitude of the velocity anomaly is hard to be explained by a pure thermal effect.
This suggests non-thermal origin of the observed slow velocity, possibly including both chemical heterogeneity and the presence of water/partial melt [Van der Lee et al., 2008].

4.4. The Colorado Plateau

The $SH$ velocity model shows fast anomalies under the Colorado Plateau at 0–200 km depth (Figure 13BB'). This high-velocity structure is thicker in the southwest than in the northeast, which is also observed by $P$-wave tomography except with an overall thinner root [Sigloch, 2008]. This uneven thickness is resolved as shown in Figure 23FF'. These fast anomalies probably represent a thick lithospheric root. Since the 2-km elevation of the Colorado Plateau largely exceeds the elevation that is expected to be supported by depleted cratonic lithosphere, some other form of support, static and/or dynamic, is needed. Various suggestions have been made, including warming of heterogeneous lithosphere [Roy et al., 2009], edge-driven convection [Karlstrom et al., 2008; Van Wijk et al., 2010], dynamic modelling of mantle convection [Moucha et al., 2009; Liu and Gurnis, 2010], and small-scale mantle convection represented by drip-like high-velocity bodies in tomographic images [Schmandt and Humphreys, 2010b]. In our $SH$ velocity model, we observe slow anomalies under southern California and northern Baja California in the lower mantle down to 1000 km depth. These slow anomalies overlap the warm mantle upwelling proposed by Moucha et al. [2009] which is responsible for the uplift in the central Basin and Range province and Colorado Plateau. The thicker root in the southwest coincides with the higher accumulative dynamic topography in the southwest plateau for the last 25 My from the mantle convection model [Moucha et al., 2009].

In the upper mantle, at 100 km depth (Figure 13BB', map view), anomalies as low as $-6\%$ are observed under the boundary between the Colorado Plateau and the Basin and
Range. The large magnitude of the slow anomalies suggest partial melt as discussed in Section 4.2 or the existence of water, which might be a result of hydration of the mantle beneath the Basin and Range due to the subduction of the Farallon slab under this region over 30 My ago [Atwater and Stock, 1998]. The slow anomaly rapidly changes to a fast anomaly as we go southeast from the Basin and Range to the Colorado Plateau, with a velocity contrast of 9.0% over 150 km. This large velocity gradient is resolved in our model as shown in Figure 22 (100 km). Sine et al. [2008] propose that this large velocity gradient defines a boundary between altered Paleozoic lithosphere (under the Basin and Range) and unaltered Proterozoic lithosphere (under Colorado Plateau).

5. The Attenuation Model of the Mantle beneath North America

A catalog of the attenuation maps is given in Figures 19–21 (left columns). These maps show that a large portion of the attenuation signal is located in the Western U.S., because the attenuation resolution in the Central and Eastern U.S. is low (Figure 27) due to poor data coverage (Figures 5e and 6e). Figure 24 shows that the attenuation resolution (especially amplitude recovery) starts to decrease at 600 km depth and the anomalies become unresolvable at 800 km depth and below. In the following, we discuss the relationship between attenuation and velocity anomalies in the upper mantle and transition zone beneath the U.S., which may provide extra constraints on the physical sources of mantle heterogeneities.

If we assume an approximate linear relationship $\delta \ln Q_S = k \delta \ln V_S + b$, then the correlation coefficient between $\delta \ln Q_S$ and $\delta \ln V_S$ describes the strength of this linear relationship. Because different physical sources (e.g., temperature, water content, partial melt, chemical composition, grain size) of mantle heterogeneities give different slopes of the linear
relationship, a low correlation coefficient suggests coexistence of multiple physical sources.

Even the same physical source has a depth-dependent property and thus produces a depth-dependent slope, so we study the correlation coefficient and slope at each depth rather than the average value. Figure 14a plots the variation of the signed correlation coefficient with depth for the Western U.S. (WUS, west of 255°E) and the Central and Eastern U.S. (EUS, east of 255°E), respectively. All correlation coefficients are positive, meaning a dominant coexistence of slow (fast) velocity and high (low) attenuation. Such positive correlation is also observed in global models derived from surface waves [e.g., Romanowicz, 1990; Artemieva et al., 2004; Dalton and Ekström, 2006] and in regional studies [e.g., Roth et al., 2000]. A positive correlation is predicted for the effects of most physical sources, such as temperature [e.g., Karato, 1993; Jackson et al., 2002], water content [e.g., Karato, 2006], partial melt with grain boundary sliding [e.g., Jackson et al., 2004; Faul et al., 2004], and grain size [e.g., Faul and Jackson, 2005]. The large correlation coefficient for EUS suggests one major physical source of attenuation and velocity heterogeneities under EUS, which is most likely temperature [e.g., Shito et al., 2006]. WUS has a lower correlation coefficient than EUS, suggesting that non-thermal sources play a more important role in forming heterogeneities under WUS, especially at the larger depth.

Further insight into the physical state of the mantle comes from examining the variation of the slope $k = \partial(\delta \ln Q_S)/\partial(\delta \ln V_S)$ with depth (Figure 14b). Because $\delta \ln Q_S$ and $\delta \ln V_S$ do not have a strict linear relationship, we treat them as random variables and estimate the slope as the direction of the major axis of the error ellipse. A formal estimate for the uncertainty in the slope is obtained using a 10% resampling technique (jackknifing, Tian et al. [2009]). Figure 14 shows that small slope uncertainties correspond to large
correlation coefficients (e.g., under EUS), and this is expected from the situation of one major physical source. Large slope uncertainties correspond to small correlation coefficients (e.g., around 350 km depth under WUS), and are expected from the coexistence of multiple physical sources. The correlation coefficient at 600 km depth under WUS is so low that a simple linear relationship is no longer sufficient. Therefore we do not discuss the slope at this depth. The slope for EUS mainly reflects the behavior of thermal effects, as discussed above. The slope for WUS is significantly larger than that for EUS, suggesting non-thermal physical sources under WUS that produce larger slopes than temperature, especially around 350 km depth. Such sources are possibly increasing water content [Karato, 2006] and partial melt with grain boundary sliding [Jackson et al., 2004; Faul et al., 2004].

Extra constraints on the physical state of the mantle may be obtained by examining the lateral variation of the $\delta \ln V_S - \delta \ln Q_S$ relationship. Figure 15 shows maps of the normalized $C = \delta \ln V_S \times \delta \ln Q_S$ in the upper mantle and transition zone under North America. Positive values of $C$ are dominant, consistent with the observed positive correlation coefficients (Figure 14a). Large-scale positive $C$ exists in various tectonic regions, e.g., in the Wyoming craton at 100–200 km depth, under the magmatically active Basin and Range at 100–200 km depth, in the subducted slabs under Iowa, Illinois, Missouri at 400–600 km depth. Positive correlation under the old continents and magmatic regions on a global scale is reported by Dalton et al. [2009]. Note that large positive $C$ does not necessarily indicate a large correlation coefficient, because there may be multiple physical sources with different (but all positive) slopes $k$, which in combination gives a relatively low (but positive) correlation coefficient. Relatively large-scale anti-correlation between $\delta \ln V_S$ and
δlnQ_S (negative C) is also observed, and one can only speculate about its cause. Low velocity and low attenuation coexist under the northwest coast at 200–400 km depth. This is likely due to the smearing of the slab-related, low attenuation feature in the attenuation model. The coexistence of low velocity and low attenuation under the Central Valley at 100–400 km depth is more puzzling, though it may perhaps be explained by the effect of partial melt with a melt-squirt mechanism [Hammond and Humphreys, 2000a,b]. A third example of anti-correlation is the coexistence of high velocity and high attenuation under the Northern Rocky Mountains at 100–200 km depth. Calculations show that for natural peridotites, an increase in Mg# significantly increases \( V_S \) [Lee, 2003]. On the other hand, attenuation is expected to be much less sensitive to major element variations [Karato, 2006].

6. Conclusions

We perform a joint inversion on multiple-frequency \( SH \)-wave delays and amplitude anomalies, \( SS \)-wave differential delays, and Love-wave fundamental-mode phase delays, to simultaneously obtain the velocity and attenuation structure under North America. Besides the intrinsic frequency dependence of Love-wave phases, frequency dependence is also observed for all the body-wave data in our study, indicating the presence of anomalies smaller than the width of the Fresnel zones.

Our \( SH \) velocity model very much confirms the same fast anomalies as in the \( P \) velocity model by Sigloch et al. [2008], including the two separate subduction systems under North America and the slab gap under the Western U.S. It also reveals further evidence of high-level segmentation and deformation of the slab, including the droplet-like fragments of \( S2 \) and \( F2 \) in the lower mantle, a tear in \( F2 \) under 700 km, and discontinuities around
410 km depth in the slab under the Western U.S. With the more recent USArray data, we are able to extend the high velocity resolution further east compared with Sigloch et al. [2008] and Tian et al. [2009], and start to see a glimpse of the Yellowstone plume. A lower mantle plume Y2 is observed to originate from about 1500 km depth, and rises up through the slab gap. We propose that a southward mantle flow produces the 40° tilting angle (from vertical) of Y2.

Attenuation is only resolved in the upper mantle and transition zone under the U.S. Large positive correlation coefficients between $\delta \ln V_S$ and $\delta \ln Q_S$ are observed under the Central and Eastern U.S., suggesting one major physical source of mantle heterogeneities, most likely temperature. Smaller correlation coefficients and larger $\delta \ln Q_S - \delta \ln V_S$ slopes are observed under the Western U.S., suggesting that non-thermal physical sources, probably the existence of water and partial melt (with grain boundary sliding), are playing a more important role, especially around 350 km depth. Anti-correlation is observed in the upper mantle under the Central Valley and the Northern Rocky Mountains.

The inclusion of Love-wave phase delays helps improve velocity resolution in the lower mantle by better constraining the upper mantle. It also influences the attenuation image indirectly by changing the velocity image and thus the focusing effect, resulting in weaker attenuation anomalies above 200 km and stronger attenuation anomalies below 500 km. The misfits of Love-wave and SS-wave data are larger than those of SH-wave data.
Acknowledgments. We thank Frederik Simons, Nadine McQuarrie, Tom Duffy, and Eugene Humphreys for helpful discussions. Yue Tian was supported by the US National Science Foundation under grant numbers EAR-0309298 and EAR-0105387. Ying Zhou was supported by the US National Science Foundation under grant numbers EAR-0809464. Guust Nolet thanks the European Research Council for financial support through the Globalseis Advanced Grant. Part of the maps are produced with GMT [Wessel and Smith, 1998].
References


Goes, S., and R. Govers (2000), Shallow mantle temperatures under Europe from $P$ and

Goes, S., and S. van der Lee (2002), Thermal structure of the North Ameri-
can uppermost mantle inferred from seismic tomography, *J. Geophys. Res.*, 107,

Goes, S., F. A. Capitanio, and G. Morra (2008), Evidence of lower-mantle slab penetration

Hammond, W. C., and E. D. Humphreys (2000a), Upper mantle seismic wave velocity:

Hammond, W. C., and E. D. Humphreys (2000b), Upper mantle seismic wave attenuation:


Hwang, Y. K., J. Ritsema, and S. Goes (2009), Spatial variations of $P$ wave attenuation in

Jackson, I., J. D. F. Gerald, U. H. Faul, and B. H. Tan (2002), Grain-size-
sensitive seismic wave attenuation in polycrystalline olivine, *J. Geophys. Res.*, 107,

Jackson, I., U. H. Faul, J. D. Fitz Gerald, and B. H. Tan (2004), Shear wave attenu-
ation and dispersion in melt-bearing olivine polycrystal: 1. Specimen fabrication and


Schmandt, B., and E. Humphreys (2010b), Complex subduction and small-scale convection revealed by body-wave tomography of the western United States upper mantle,


Figure 1. Histograms of $SH$-wave measurements used in this study (before applying ellipticity, elevation, and crustal corrections). a-e: Travel-time anomalies defined as the difference between the observed arrival time $T_{S}^{\text{obs}}$ and the predicted arrival time $T_{S}^{\text{pre}}$. f-h: Amplitude anomalies defined as the ratio between the observed amplitude $A_{S}^{\text{obs}}$ and the predicted amplitude $A_{S}^{\text{pre}}$, in decibel. The title indicates the center period of the frequency band ($T_{c}$) and the median of the measurements in that band ($med$).
Figure 2. Broadband $SS$-wave differential travel-time anomalies $\delta(T_{SS} - T_S)$ plotted at each reflection point. Positive values reflect slow velocity anomalies in the vicinity of the reflection point, and negative values indicate fast velocity anomalies.
Figure 3.  a-c: Histograms of SS-wave differential travel-time anomalies $\delta(T_{SS} - T_S)$ used in this study (after applying ellipticity, elevation, and crustal corrections). $T_c$ indicates the center period of the frequency band. d: Dispersion of $\delta(T_{SS} - T_S)$. On the horizontal axis, each frequency band is represented by its center period. The vertical axis is the difference between $\delta(T_{SS} - T_S)$ in one band and $\delta(T_{SS} - T_S)$ in band 40 s. Only wave paths that have acceptable measurements in all three bands are used to produce the plot. The black dot, red bar, blue bar, and green bar represent the median, 68% interval, 90% interval, and 95% interval of the dispersion, respectively.
Figure 4. Histogram of Love-wave relative phase delays $\delta \phi / \phi$ at 11 frequencies (5, 6, 7, · · · , 15 mHz) used in this study, after applying crustal corrections.
Figure 5. The kernel density (see equation (9)) for the five sensitivity kernels in equation (8) at 100 km depth. The title indicates the kernel type. The central column is for $SH$-wave velocity kernels. The right column is for $SH$-wave attenuation kernels. The upper left panel is for Love-wave kernels, and the lower left panel is for $SS$-wave kernels.
Figure 6. The kernel density (see equation (9)) for the five sensitivity kernels in equation (8) at 600 km depth. The title indicates the kernel type. The central column is for $SH$-wave velocity kernels. The right column is for $SH$-wave attenuation kernels. The upper left panel is for Love-wave kernels, and the lower left panel is for $SS$-wave kernels.
Figure 7. Trade-off curve for velocity (circles) and attenuation (asterisks) models. Left: Trade-off between fitting the data and norm damping. Right: Trade-off between fitting the data and smoothing the model. $\chi^2/N$ is defined by equation (10), $m$ is the model vector $\delta \ln V_S$ or $\delta \ln Q_S^{-1}$ in equation (8), $R$ is the Laplacian roughening operator [Nolet, 2008, section 14.5], $\| \cdot \|$ represents the L2-norm. Different points on the same curve correspond to different values of $\epsilon_1$ or $\epsilon_2$ in equation (8). The dotted line indicates the preferred model with $\chi^2/N = 0.928$. 
Figure 8. Great-circle cross sections of the subduction system under North America. The black dot and $0^\circ$ represent the mid-point of the great circle arc. In the map views, green lines delineate tectonic boundaries, dotted black lines delineate political boundaries, and the depth is indicated at the lower left corner. In the cross-section views, the 410-km and 660-km discontinuities are indicated, and each grid along the arc represents $1^\circ$. SW denotes the slab window south of the Mendocino Fracture Zone, where no subduction exists. Cross section AA’ shows the two separate subduction systems: the younger one in the west ($S_0$, $S_1$, $S_2$) and the older one in the east ($F_1$, $F_2$); C denotes the craton. Cross-section BB’ shows that $F_1$ lies above $S_2$ and $N_2$. Labels $S_1$, $N_1$, $S_2$, $N_2$, $F_1$, $F_2$ identify the same structure as in Sigloch et al. [2008].
Figure 9. Three-dimensional view of the subduction system under the Western U.S., looking from the east. Plotted is the isosurface of $\delta \ln V_s = +0.6\%$. The extent and geometry of the structure only changes modestly as we shift the contour level between 0.5% and 0.8%. Red represents the subduction in the upper mantle (S0, N0), green in the transition zone (S1, N1), and purple in the lower mantle (S2, N2). Fast anomalies that are not deemed subducted material (e.g., the craton, the Colorado Plateau root) are not displayed. SG represents the slab gap, which is a continuous trail void of fast anomalies and divides the slab into the northern part (N0, N1, N2) and the southern part (S0, S1, S2). SW denotes the slab window south of the Mendocino Fracture Zone, where no subduction exists. Labels S1, N1, S2, N2, SG identify the same structure as in Sigloch et al. [2008].
Figure 10. Great-circle cross sections of the subduction system under the Western U.S. The black dot and 0° represent the mid-point of the great circle arc. In the map views, green lines delineate tectonic boundaries, dotted black lines delineate political boundaries, and the depth is indicated at the lower left corner. In the cross-section views, the 410-km and 660-km discontinuities are indicated, and each grid along the arc represents 1°. SW denotes the slab window south of the Mendocino Fracture Zone, where no subduction exists. Cross section AA’ goes through the subducted slab (S0, N0) and the slab gap (SG) in the upper mantle. Cross section BB’ goes through the subducted slab (S1, N1) and the slab gap (SG) in the transition zone. Labels S1, N1, SG identify the same structure as in Sigloch et al. [2008].
Figure 11. Three-dimensional view of the plume-slab interaction system under the Yellowstone and the eastern Snake River Plain, looking from the southwest. Red represents the isosurface of slow anomalies $\delta \ln V_S = -0.5\%$, and blue represents the isosurface of fast anomalies $\delta \ln V_S = +0.8\%$. The extent and geometry of the structure only changes modestly as we shift the contour level between 0.5\% and 0.8\%. The green triangle represents the location of the Yellowstone Caldera. For slow anomalies above 660 km, only those under the Yellowstone and the eastern Snake River Plain (40°-46°N, 244°-251°E) are displayed. Shallow fast anomalies that are not deemed subduction (e.g., the craton and the Colorado Plateau root) are not displayed. a: Slow anomalies only. Under the eastern Snake River Plain, SR0 is in the upper mantle and SR2 is in the lower mantle. Under the Yellowstone Caldera, Y0 is in the upper mantle, Y1 in the transition zone, and Y2 in the lower mantle. b: Superimposing the fast anomalies on top of the slow anomalies. See Figure 9 for a description of the fast anomalies. Labels S1, N1, S2, N2, SG identify the same structure as in Sigloch et al. [2008].
Figure 12. Great-circle cross sections of the velocity structure under the Yellowstone Caldera (YC) and the eastern Snake River Plain (SRP). The black dot and 0° represent the location of the Yellowstone Caldera and the mid-point of the great circle arc. In the map views, green lines delineate tectonic boundaries, dotted black lines delineate political boundaries, and the depth is indicated at the lower left corner. In the cross-section views, the 410-km and 660-km discontinuities are indicated, and each grid along the arc represents 1°. Cross section AA’ goes through the Yellowstone plume trail (Y2, Y1, Y0). Cross section BB’ goes through the eastern Snake River Plain (SRP) and the slow anomaly beneath it (SR2), and the northern edge of the Yellowstone plume trail (Y2, Y1, Y0). S1, S2, N2, SG are parts of the subduction system under the Western U.S. (see Figure 9), and they identify the same structure as in Sigloch et al. [2008].
Figure 13. Great-circle cross sections of selected small-scale velocity features. The black dot and 0° represent the mid-point of the great circle arc. In the map views, green lines delineate tectonic boundaries, dotted black lines delineate political boundaries, and the depth is indicated at the lower left corner. In the cross-section views, the 410-km and 660-km discontinuities are indicated, and each grid along the arc represents 1°. Cross section AA’ goes through slow anomalies under the New England Province (NEP). Cross section BB’ goes through the Colorado Plateau (CP) and the Basin and Range (BR).
Figure 14.  

(a): Variation of signed correlation coefficient between $\delta \ln V_S$ and $\delta \ln Q_S$ with depth. The vertical bar represents the 95% confidence interval.  

(b): Variation of slope $\partial (\delta \ln Q_S)/\partial (\delta \ln V_S)$ with depth. At each depth, the slope is estimated as the direction of the major axis of the error ellipse. The vertical bar gives the slope uncertainty, which is estimated with a 10% resampling technique (jackknifing, Tian et al. [2009]). The slope at 600 km depth under WUS is out of the axis range. In both (a) and (b), blue is for the Western U.S. (WUS, west of 255°E) and red is for the Central and Eastern U.S. (EUS, east of 255°E). Points with $|\delta \ln V_S| < 0.5\%$ or $|\delta \ln Q_S| < 5\%$ are not used in producing the plots.
Figure 15. Maps showing $\delta \ln V_s \times \delta \ln Q_s$ at 100–600 km depth. The value is normalized to have a maximum of 1. Points with $|\delta \ln V_s| < 0.5\%$ or $|\delta \ln Q_s| < 5\%$ are not used in producing the figure, in order to avoid that damping effects dominate the correlation.
Supporting nonprint material

Appendix A: The Joint Inversion System

Our tomographic system (8) is a joint system in that four different types of data ($\delta T_S$, $\delta \ln A_S$, $\delta \ln \phi$, $\delta (T_{SS} - T_S)$) are used to simultaneously invert for velocity and attenuation heterogeneities. Velocity is constrained by all four types of data and attenuation is constrained by $SH$-wave amplitudes. The relative importance of velocity and attenuation heterogeneities in explaining $SH$-wave travel-time and amplitude anomalies and the contribution of $SH$-wave amplitudes to constrain the velocity model are discussed by Tian et al. [2009]. This study extends the work of Tian et al. [2009] by including Love-wave and $SS$-wave data. Here we examine how well each type of data is explained and study the contribution of $\delta \ln \phi$ and $\delta (T_{SS} - T_S)$ to constrain velocity and attenuation models.

A1. Misfit for Each Data Type

Figure 16 describes how well each type of data is explained, with $\chi^2/N$ defined by equation (10). It is not surprising that $SH$-wave delays have the smallest misfit, because direct body-wave travel times are affected by fewer factors and are easier to explain than other measurements (e.g., amplitudes, travel times of later phases) extracted from the seismograms. Love-wave phase delays and $SS$-wave differential delays have the largest misfits, even higher than the misfit of $SH$-wave amplitudes. This is likely due to the fact that the number of $SS$-wave or Love-wave measurements is only $\sim 19\%$ of the number of $SH$-wave travel-time data and $\sim 25\%$ of the number of $SH$-wave amplitude data (see also Figures 5–6). Thus the velocity model is “biased” to satisfy the major portion of the data, which is the $SH$-wave measurement.
A2. Contribution of Love and SS Waves to Constraining Velocity and Attenuation Models

Since the velocity and attenuation models are dominantly guided by the \textit{SH}-wave data, one may ask the question what the contribution of Love and \textit{SS} waves is to constraining velocity and attenuation models. To answer this question, we remove Love-wave or \textit{SS}-wave data, and do a joint tomography on the remaining sub-dataset. A comparison between the models obtained from the complete data set and from the sub-dataset is shown in Figures 17–18. The three inversions have about the same $\chi^2/N$, i.e., they explain the data equally well, to make a fair comparison.

At 100–400 km depth, the addition of Love or \textit{SS} waves gives large velocity amplitudes in the periphery of the U.S. (Figure 17, 100 km). This is because at shallow depth, the peripheral region is covered by Love and \textit{SS} waves but not by \textit{SH} waves (Figures 5–6). At 100 km depth, somewhat surprisingly, Love waves help improve the lateral resolution (comparing the left and right columns in Figure 17), e.g., imaging the Cascades subduction zone and the Colorado Plateau root with larger anomaly amplitudes. At 400–1000 km depth, \textit{SS} and Love waves mainly help better resolve the fast velocity structure (slabs) under the Western and Eastern U.S., respectively (Figure 17, 500 and 700 km). Interestingly, below 1000 km, the addition of Love waves show a stronger influence than adding \textit{SS} waves (Figure 17, 1200 km), although the Love-wave kernel become negligible below 400 km (Figure 6a). This indicates that it is important to have Love waves to constrain the shallow velocity model, which in turn reduces ambiguity in the deep velocity model.

In Figure 18, no significant difference is observed between the left and central columns, indicating that adding \textit{SS} waves does not affect the attenuation image much. On the other hand, compared with the right column, the left column has a smaller amplitude at 100 km depth and a larger amplitude at 600 km depth. This indicates that the addition of Love waves redistributes
the strength of the attenuation anomalies in the inversion results, resulting in weaker anomalies in the shallow region (100–200 km depth) and stronger anomalies in the deeper region (500–700 km depth). The influence of Love waves on velocity and attenuation images is consistent with the conclusion that focusing effect dominates over attenuation effect in explaining amplitude anomalies [Tian et al., 2009]. The change of the velocity image explained by the addition of Love waves changes the amount of amplitude anomalies due to focusing effect, and thus changes the amount of amplitude anomalies to be explained by attenuation.

Appendix B: A Catalog of Model Maps

A catalog of the preferred velocity and attenuation models under North America down to 1600 km depth is shown in Figures 19–21.

Appendix C: Resolution Tests

Resolution tests are performed to provide some sense of the reliability of the model. The complete input model of the resolution test includes the designed velocity and attenuation anomalies and the randomly produced correction terms (see equation 8). The synthetic data are produced by multiplying the kernel with the input model. Added to the synthetic data is randomly produced noise which has a normal distribution. The joint inversion is implemented on the perturbed synthetic data to produce the output model.

C1. Recovery of the Tomographic Models

We perform a resolution test with the tomographic velocity and attenuation models as input. The signal-to-noise ratio (ratio between the RMS of the synthetic data before adding noise and the RMS of the noise) is 2.35. The recovered models are shown in Figures 22–24. This gives us a good sense of what features are reliable, or at least what features are not reliable. For velocity
(Figure 22), at 100 km depth, the anomalies in the Central U.S. are not well recovered due to the low kernel density there (Figure 5c). At 100–300 km depth, the anomalies in the periphery of the U.S. are poorly recovered, indicating that these features are not reliable. At 400-1200 km depth, most velocity anomalies are well recovered, with a slightly smaller amplitude. Below 1200 km, the amplitude of the anomalies is less well recovered. Figure 23 shows that the velocity features discussed in Section 4 are reasonably well recovered. For attenuation (Figure 24), the anomalies are fairly well recovered above 500 km, with a slightly smaller amplitude or slightly changed geometry. At 500–800 km depth, the attenuation resolution obviously decreases, especially in the Eastern U.S. Below 800 km, the attenuation structure is not resolved.

C2. Lateral Resolution for Velocity

Figure 25 shows the lateral resolution for velocity. The input velocity anomalies are regularly spaced Gaussian balls, with the ball dimensions adapted to different resolving power in different regions. The signal-to-noise ratio is 3.67. The multiple-frequency data from the dense USArray network provide very high lateral resolution in the Western U.S. The Central and Eastern U.S. has good resolution, although lower than the Western U.S. The periphery of the U.S. has poor resolution, because it is constrained by SS and Love waves, which provide only a small portion of the data set (see Figures 5–6).

C3. Depth Resolution for Velocity

Figure 26 shows the depth resolution for velocity. The input velocity anomalies have three stripes: fast velocity at 0–200 km depth, slow velocity at 400–800 km depth, and fast velocity at 1000–1800 km depth. The signal-to-noise ratio is 4.73. Vertical leakage occurs between the stripes. In general, the model is better resolved under the Western U.S. than under the Central
and Eastern U.S., and the shallow region has higher resolution than the deep region. As we go from south (AA’) to north (CC’), the resolution slightly decreases.

C4. Lateral Resolution for Attenuation

Figure 27 shows the lateral resolution for attenuation. The input attenuation anomalies are regularly spaced Gaussian balls, with the ball dimensions adapted to different resolving power in different regions. The signal-to-noise ratio is 3.74. Similar to the velocity resolution, because of the dense USArray coverage, the resolution in the Western U.S. is higher than that in the Central and Eastern U.S., especially in terms of amplitude recovery. There is almost no resolution for attenuation outside of the U.S. (Figures 19-21, left columns) because the attenuation is only constrained by SH waves, which have little coverage outside of the U.S. (Figures 5e and 6e). Compared with velocity, attenuation has much larger resolvable length and worse amplitude recovery. This confirms that it is very hard to do Q tomography. According to Figure 27, we can only trust the first-order features in our tomographic attenuation model.
Figure 16. The value of $\chi^2/N$ for each data type.
Figure 17. Velocity models from joint tomography with different data sets. Left: with all four types of data, $SH$-wave delays and amplitude anomalies, $SS$-wave differential delays, Love-wave phase delays. Center: removing $SS$-wave data, with three types of data, $SH$-wave delays and amplitude anomalies, Love-wave phase delays. Right: removing Love-wave data, with three types of data, $SH$-wave delays and amplitude anomalies, $SS$-wave differential delays. Geologic and political boundaries are plotted. Depth and the reference model value are shown at the lower left corner of each map. For the models from left to right, the $\chi^2/N$ is 0.93, 0.93, 0.98, respectively.
Figure 18. Attenuation models from joint tomography with different data sets. Left: with all four types of data, $SH$-wave delays and amplitude anomalies, $SS$-wave differential delays, Love-wave phase delays. Center: removing $SS$-wave data, with three types of data, $SH$-wave delays and amplitude anomalies, Love-wave phase delays. Right: removing Love-wave data, with three types of data, $SH$-wave delays and amplitude anomalies, $SS$-wave differential delays. Geologic and political boundaries are plotted. Depth and the reference model value are shown at the lower left corner of each map. For the models from left to right, the $\chi^2/N$ is 0.93, 0.93, 0.98, respectively.
Figure 19. Map views of the preferred tomographic models of $\delta \ln Q_S$ (left) and $\delta \ln V_S$ (right). at 100, 200, 300, and 400 km depth. Geologic and political boundaries are plotted. Depth and the reference model value are shown at the lower left corner of each map.
Figure 20. Map views of the preferred tomographic models of $\delta \ln Q_S$ (left) and $\delta \ln V_S$ (right) at 500, 600, 700, and 800 km depth. Geologic and political boundaries are plotted. Depth and the reference model value are shown at the lower left corner of each map.
Figure 21. Map views of the preferred tomographic models of $\delta\ln Q_s$ (left) and $\delta\ln V_s$ (right) at 1000, 1200, 1400, and 1600 km depth. Geologic and political boundaries are plotted. Depth and the reference model value are shown at the lower left corner of each map.
Figure 22. Resolution test with the tomographic velocity and attenuation models as input. Maps of the recovered velocity model are shown. For input models, see Figures 19–21 (right columns). Geologic and political boundaries are plotted. Depth is indicated at the lower left corner of each map.
Figure 23. Resolution test with the tomographic velocity and attenuation models as input. Cross sections of the recovered velocity model are shown, with AA' being the recovered cross section of Figure 8AA', BB' for Figure 10BB', CC' and DD' for Figure 12AA' and BB', EE' and FF' for Figure 13AA' and BB'.
Figure 24. Resolution test with the tomographic velocity and attenuation models as input. Maps of the recovered attenuation model are shown. For input models, see Figures 19–21 (left columns). Geologic and political boundaries are plotted. Geologic and political boundaries are plotted. Depth is indicated at the lower left corner of each map.
Figure 25. Lateral resolution for velocity. The depth of each map is indicated at the lower left corner. The input anomalies are regularly spaced Gaussian balls. At 100 km depth, the $1/e$ diameter of the Gaussian ball is 150 km under the Western U.S., 276 km under the Central and Eastern U.S., and 900 km outside of the U.S. At 400 km depth, the $1/e$ diameter of the Gaussian ball is 190 km under the Western U.S., 450 km under the Central and Eastern U.S., and 900 km outside of the U.S. At 800 and 1200 km depth, the $1/e$ diameter of the Gaussian ball is 480 km under the U.S., and 850 km outside of the U.S. The spacing between adjacent Gaussian balls is large enough to avoid overlapping of the balls.
Figure 26. Depth resolution for velocity. Three great-circle cross sections through North America are shown. The black dot and 0° represent the mid-point of the great circle arc. In the map view, solid black lines delineate tectonic boundaries, and dotted black lines delineate political boundaries. In the cross-section views, each grid along the arc represents 1°. The input model has three stripes: fast velocity at 0–200 km depth, slow velocity at 400–800 km depth, and fast velocity at 1000–1800 km depth.
Figure 27. Lateral resolution for attenuation. The depth of each map is indicated at the lower left corner. The input anomalies are regularly spaced Gaussian balls. At 100 km depth, the $1/e$ diameter of the Gaussian ball is 386 km under the Western U.S., and 600 km under the Central and Eastern U.S. At 400 km depth, the $1/e$ diameter of the Gaussian ball is 600 km. At 800 km depth, the $1/e$ diameter of the Gaussian ball is 1000 km. The spacing between adjacent Gaussian balls is large enough to avoid overlapping of the balls.