University of Rhode Island

DigitalCommons@URI

Open Access Dissertations

2019

THREE-DIMENSIONAL PASSIVE-SOURCE REVERSE TIME MIGRATION OF CONVERT WAVES: A NEW METHOD TO IMAGE THE EARTH'S DISCONTINUITIES

Jiahang Li

University of Rhode Island, jiahang.li291@gmail.com

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

---

Recommended Citation

Li, Jiahang, "THREE-DIMENSIONAL PASSIVE-SOURCE REVERSE TIME MIGRATION OF CONVERT WAVES: A NEW METHOD TO IMAGE THE EARTH'S DISCONTINUITIES" (2019). *Open Access Dissertations*. Paper 841.

https://digitalcommons.uri.edu/oa_diss/841

This Dissertation is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

THREE-DIMENSIONAL PASSIVE-SOURCE REVERSE-TIME MIGRATION OF CONVERT
WAVES: A NEW METHOD TO IMAGE THE EARTH'S DISCONTINUITIES

BY

JIAHANG LI

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

OCEANOGRAPHY

UNIVERSITY OF RHODE ISLAND

2019

DOCTOR OF PHILOSOPHY DISSERTATION

OF

JIAHANG LI

APPROVED:

Dissertation Committee:

Major Professor   Yang Shen

Brian Savage

Gopu Potty

Nasser H. Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2019

ABSTRACT

At seismic discontinuities in the crust and mantle, part of the compressional-wave energy converts to shear wave, and vice versa. These converted waves have been widely used in receiver function (RF) studies to image discontinuity structures in the Earth. While generally successful, the conventional RF method has its limitations and is suited mostly to flat or gently dipping structures. Among the efforts to overcome the limitations of the conventional RF method is the development of the wave-theory-based, passive-source reverse-time migration (PS-RTM) for imaging complex seismic discontinuities and scatters. To date, PS-RTM has been implemented only in 2D in the Cartesian coordinate for local problems and thus has limited applicability. In Chapter 1, we introduce a 3D PS-RTM approach in the spherical coordinate, which is better suited for regional- and global-scale seismic imaging. New computational procedures are developed to reduce artifacts and enhance migrated images, including back-propagating the main arrival and the coda containing the converted waves separately, using a modified Helmholtz decomposition operator to separate the P and S modes in the back-propagated wavefields, and applying an imaging condition that maintains a consistent polarity for a given velocity contrast. The new approach allows us to use migration velocity models with realistic velocity discontinuities, improving accuracy of the migrated images. We present several synthetic experiments to demonstrate the method, using regional and teleseismic sources. The results show that both regional and teleseismic sources can illuminate complex structures and this method is well suited for imaging dipping interfaces and sharp lateral changes in discontinuity structures.

In Chapter 2, we test the 3D PS-RTM method with sparse and unevenly distributed seismic arrays using synthetic data and discuss the station spacing limitation in its application. A cubic spline interpolation method is used to interpolate the sparse and unevenly recorded seismic signal onto numerical grids to reconstruct the full wavefield. We found that as long as the station spacing is smaller than half apparent wavelength of the converted wave, and the noise level is comparable or even slightly larger than the converted wave, the cubic spline method is sufficient to interpolate the wavefield for the PS-RTM method. In Chapter 3, we apply the 3D PS-RTM method to the Yellowstone hotspot to image the mantle transition zone discontinuities. We develop a data regularization procedure that includes iterative deconvolution, principle component analysis and interpolation. The RTM volumes from different earthquakes are weighted and stacked based on the distribution of stations that record each earthquake. We discuss how the earthquake distribution, reference velocity model, dominant frequency and principle component analysis affect the image results. The resulting RTM model shows a depressed 410-km discontinuity, an uplifted 660-km discontinuity and a thinner-than-normal mantle transition zone thickness beneath Yellowstone, which may suggest a deep mantle plume rising from the lower mantle.

ACKNOWLEDGMENTS

I want to thank all the staff and faculty of the Graduate School of Oceanography, University of Rhode Island who helped me in both my research and life during the last 5 years. Mostly, I would like to thank my advisor Yang Shen for years of patient guidance. He introduced me to the reverse-time migration method and encouraged me to overcome different issues during years of study. Most importantly, he endured my poor writing skills and helped me to improve it time and time again. I could not finish this dissertation without his help. I also want to thank the members of the seismology lab in University of Rhode Island, for keeping me company and providing useful comments and advice for my research.

During my PhD study, I have grown stronger both mentally and physically because of my friend Kaikai Wu, who taught me to release the stress of work in exercising. The habit of exercise makes me an optimistic and persistent person, and will help me to pursue my dreams for the rest of my life.

Finally, I want to thanks my parent who had supported me with no doubt as always. Though I know they would not understand a word of what I write or present in my defense, the trust they have in me keeps me pursuing my interests. They have taught me to be a persevering person as they are, which is the most important lesson I have in my life.

PREFACE

The following dissertation entails the key steps of the development of a new imaging method called 3D passive source reverse time migration, using converted P wave. The dissertation is written in manuscript format, comprised of the three following manuscripts:

Manuscript one, “Three-dimensional passive-source reverse-time migration of converted waves: The method” describes the theory of 3D passive-source reverse-time migration and its advantages over traditional receiver function methods. This manuscript was published in Journal of Geophysical Research: Solid Earth, December 2018.

Manuscript two, “Numerical tests of 3D passive-source reverse-time migration of interpolated wavefields recorded by sparse seismic arrays ” presents some numerical tests results using sparse seismic array. It provides guidance for future seismic network design and the limitation of 3D passive-source reverse-time migration. Part of this manuscript is published as supplementary material in Journal of Geophysical Research: Solid Earth, December 2018.

Manuscript three, “3D passive-source reverse-time migration imaging of the mantle transition zone beneath the Yellowstone hotspot ” demonstrates the advantage of our method in imaging the mantle transition zone discontinuities in Yellowstone hotspot region. This manuscript is prepared, with submission to Earth and Planetary Science Letters expected in August 2019.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS.....	iv
PREFACE.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES	vii
CHAPTER 1, Three-dimensional passive-source reverse-time migration of converted waves: The method.....	1
CHAPTER 2, Numerical tests of 3D passive-source reverse-time migration of interpolated wavefields recorded by sparse seismic arrays	37
CHAPTER 3, 3D passive-source reverse-time migration imaging of the mantle transition zone beneath the Yellowstone hotspot.....	60

LIST OF FIGURES

Figure 1.1 Synthetic velocity model and reference velocity model

Figure 1.2 Synthetic seismograms

Figure 1.3 Comparison of forward and backward propagated wavefields

Figure 1.4 Comparison of snapshots before and after separating P and P coda

Figure 1.5 Snapshots of separated P and S mode wavefields

Figure 1.6 Image results using the conventional imaging condition

Figure 1.7 Image results using different sources

Figure 1.8 Image results using explosive sources and teleseismic sources

Figure 2.1 Slab model, reference model and image result

Figure 2.2 Seismic station distribution with different station spacing

Figure 2.3 Image result using 5km spacing seismic array without interpolation

Figure 2.4 Comparison of synthetic and interpolated seismograms

Figure 2.5 Image results using different station spacing and interpolation methods

Figure 2.6 Seismograms with different signal to noise ratios

Figure 2.7 Image results using seismograms with different signal to noise ratios

Figure 3.1 Topography map of the Yellowstone region

Figure 3.2 Map of the distribution of earthquakes used in this study

Figure 3.3 Example of three component receiver functions using iterative deconvolution

Figure 3.4 Example of PCA effects on the seismogram

Figure 3.5 Map of seismic station distribution of a particular earthquake

Figure 3.6 Vertical slices of the image result along latitudinal and longitudinal directions

Figure 3.7 Vertical slices of Wand and Pavlis (2006) image result along latitudinal and longitudinal directions

Figure 3.8 Map of 410- and 660-km discontinuity topography and thickness of the MTZ

Figure 3.9 Vertical slice of image result along latitude 40N.

Figure 3.10 Image results using different groups of earthquakes

Figure 3.11 Map of 410- and 660-km discontinuity topography and thickness of the MTZ using different reference models

Figure 3.12 Image results using higher frequency wave (4s)

Figure 3.13 Image results using only the first PCA component of vertical and radial receiver functions

CHAPTER 1

Three-dimensional passive-source reverse-time migration of converted waves: The method

By

Jiahang Li¹, Yang Shen¹ and Wei Zhang²

Published in 2018, *Journal of Geophysical Research: Solid Earth*

123. <https://doi.org/10.1002/2017jb014817>

-
1. *Graduate school of Oceanography, University of Rhode Island, 215 S Ferry Rd, Narragansett, RI 02882, USA*
  2. *Department of Earth and Space Sciences, Southern University of Science and Technology, 1088 Xueyang Ave, Nanshan Qu, Shenzhen Shi, Guangdong Sheng, 518055, China*

1.1 Introduction

The method of stacking receiver functions at the common conversion points (CCP) (e.g., Dueker and Sheehan, 1997; Shen et al., 1998; Ryberg and Weber, 2000; Abe et al., 2011) has been widely used to image velocity discontinuity interfaces of the Earth's interior, such as the Moho discontinuity and mantle transition zone boundaries (e.g., Rondenay et al., 2000; Ai et al., 2007; Kind et al., 2012; Tauzin et al., 2013; Gao and Liu, 2014). Stacking of multiple receiver functions over finite lateral and depth dimensions is necessary to enhance the signals of the converted waves as individual receiver functions often have low signal-to-noise ratios. Because of this spatial averaging, the CCP method assumes locally horizontal or sub-horizontal velocity discontinuity interfaces within the common conversion patch (point) used in stacking (Dueker and Sheehan, 1997) and does not account for the effects of wave diffraction and scattering caused by lateral variations in the interfaces (Chen et al., 2005a). Consequently, the CCP method is inadequate for imaging complex structures such as scatters, sharp changes in discontinuity structures, and high-angle dipping or faulted interfaces (e.g., vertical Moho offsets or slabs), as demonstrated by Sheehan et al. (2000) and Cheng et al. (2016).

Several other techniques have been developed to use converted waves recorded by seismic arrays to image crustal and mantle discontinuity structures: Bostock et al. (2001) developed a 2D imaging method using an inverse scattering theory based on

generalized radon transform; Liu and Levander (2013) used generalized transform based on Kirchhoff approximation; Poppeliers and Pavlis (2003a, 2003b) used plane wave decomposition to transform the recorded data set into ray parameter and back-azimuth domain plane waves and migrated them separately; Ryberg and Weber (2000) introduced post-stack migration of receiver functions; Revenaugh (1995), Sheehan et al. (2000), Levander et al. (2013), and Cheng et al. (2016, 2017) developed pre-stack Kirchhoff migration to image scatters and velocity discontinuities; Chen et al. (2005a, 2005b) presented a wave-equation migration method, which back-propagates the CCP stacked receiver functions with one-way phase screen propagator.

Using synthetic waveforms generated by finite-difference simulations, Shang et al. (2012) developed a 2D passive-source reverse-time migration (PS-RTM) method to image complex structures with converted waves. Different from the conventional active-source, reverse-time migration (RTM) method in industry (Baysal et al. 1983; McMechan 1983; Chang and McMechan 1987 and 1994), which uses both the source-side (forward) and receiver-side (time-reverse, back-propagated) wavefields, the PS-RTM method uses only the receiver-side wavefield. Therefore, as in receiver function studies, the PS-RTM images the structures beneath the receivers using converted waves without a detailed knowledge of the sources. By comparing the CCP and PS-RTM results for a synthetic model with an offset in the Moho structure, Shang et al. (2012) demonstrated the advantage of this wave-equation-based migration method in imaging complex structures. Compared to Kirchhoff migration,

the wave-equation-based migration method is computationally more expensive. But for complex structures, it theoretically has advantages over Kirchhoff migration as it accounts for finite-frequency wave effects and overcomes, for example, multipathing in the propagating wavefield that may affect Kirchhoff migration. However, the 2D PS-RTM method of Shang et al. (2012) has limited applicability for several reasons. First, it can only use linear seismic arrays with earthquakes from a narrow azimuth range in the direction of the linear array. Second, since it does not account for diffracted and scattered waves outside of the 2D plane, its application is mostly suited to 2D structures perpendicular to the linear seismic array. Finally, a smoothed reference model was used by Shang et al. (2012) to back-propagate seismic records in order to minimize spurious phase conversions at the interfaces in the back-propagated wavefields (see more discussion in the Method section). This artificial smoothing of the velocity structure inevitably alters the wavefield near sharp velocity discontinuities and thus may introduce errors in the migrated images.

With the advance of computational power and deployment of dense three-component, areal seismic arrays, it is increasingly feasible and necessary to use wave-equation-based migration methods to image the Earth structure without the locally horizontal subsurface assumption and accurately account for the effects of wave propagation in realistic 3D heterogeneous structures. In this study, we extend the method of Shang et al. (2012) from 2D to 3D, using a 3D finite-difference elastic wave solver (Zhang et al., 2012) in the spherical coordinate, which is suited for simulating wave propagation at regional and global scales. Unlike Shang et al.

(2012), we separate the main P arrival and P-to-S converted waves in both the data (time) and image (spatial) domains. This allows us to use reference models with realistic, sharp velocity discontinuity structures. We adapt a new imaging condition that maintains a consistent polarity of the converted phases for a given velocity contrast (from low-to-high or high-to-low velocities) for sources with different azimuths and incidence angles, making it more straightforward to stack images obtained from different earthquake sources. We focus on the P-to-S converted waves in this study, though S-to-P conversions can be applied in a similar way. In the following sections, we describe the new 3D PS-RTM method and show several numerical examples to demonstrate the potentials of this method in imaging complex 3D velocity discontinuity interfaces.

1.2 Method

1.2.1 Summary of the computational procedure

Our 3D PS-RTM method contains the following steps:

Step 1: Separate the main P arrival and its coda containing converted S-wave energy by windowing them in the data (time) domain;

Step 2: Back-propagate the time-reversed, three-component P-wave seismogram and the coda containing the converted S wave separately to reconstruct two back-propagated elastic wavefields;

Step 3: Apply a modified Helmholtz decomposition operator to further isolate the P mode in the back-propagated P wavefield and S mode in the back-propagated coda wavefield, respectively, in the 3D spatial domain;

Step 4: Apply an imaging condition using the P and S modes to generate an image. A newly developed imaging condition is adapted to maintain a consistent polarity of the converted phases for a given velocity contrast;

Step 5: Stack the images generated from different earthquake sources to suppress random noise and artifacts.

1.2.2 The forward and back-propagated wavefields

To explain the above procedure and illustrate the effects of our new approaches, we generate a 3D synthetic model, which contains two layers with different velocities as shown in Figure 1.1a. The velocity interface between the two layers is offset vertically and laterally, thus the model represents a simplified 3D geological structure in places with vertical and lateral Moho offsets. In later discussion, we refer to the interface at 30 km depth as the shallow Moho and that at 50 km depth as the deep Moho. The dimension of this model is 222 km $\times$ 222 km laterally and 100 km in depth. The finite-difference grid spacing is about 0.555 km, 0.555 km, and 0.5 km in the x (co-latitude), y (longitude), and z (depth) directions, respectively. The upper and lower layer S-wave velocities are 3.9 km/s and 4.4 km/s, respectively, while the P-wave velocities are proportional to the S wave velocities ($V_p/V_s=1.74$).

Since the model is created in the spherical coordinate, the horizontal grid spacing decreases slightly with depth (Zhang et al., 2012).

Synthetic seismograms are calculated using the elastic wave equations with a free-surface boundary condition:

$$\rho \frac{\partial^2 \mathbf{u}^s}{\partial t^2} = \nabla \cdot \boldsymbol{\sigma}^s + \mathbf{f}, \quad (1.1)$$

$$\boldsymbol{\sigma}^s = \mathbf{c} : \boldsymbol{\varepsilon}^s, \quad (1.2)$$

$$\boldsymbol{\varepsilon}^s = \frac{1}{2} [\nabla \mathbf{u}^s + (\nabla \mathbf{u}^s)^T], \quad (1.3)$$

$$\boldsymbol{\sigma}_{xz}^s(z = 0) = \boldsymbol{\sigma}_{yz}^s(z = 0) = \boldsymbol{\sigma}_{zz}^s(z = 0) = 0. \quad (1.4)$$

Here $\mathbf{u}^s$ is the source-side (forward), 3-component displacement wavefield, ρ is the density, $\mathbf{f}$ is the external force representing the earthquake source, $\mathbf{c}$ is the fourth-order stiffness tensor, $\boldsymbol{\sigma}^s$ is the stress tensor and $\boldsymbol{\varepsilon}^s$ is the strain tensor (see details in Zhang et al., 2012). The superscript s represents the source-side variables. The subscript x, y , and z represent south, east and up directions, respectively, in the spherical coordinate. We calculate the synthetic wavefield with a 3D non-staggered-grid, finite-difference solver (Zhang et al., 2012) in the spherical coordinate with 4th

order spatial accuracy and 2nd order temporal accuracy. The free surface condition is implemented with the stress-image method (Levander, 1988; Graves, 1996), which anti-symmetrically sets the value of the stress component at ghost points above the free surface (Zhang et al., 2012). For other boundaries, we use 12 perfectly matched layers as the absorbing boundaries (PML, Zhang and Shen, 2010).

Two types of seismic sources are used in this study: local/regional sources (Figure 1.2a) and teleseismic sources (Figure 1.2c). We save the synthetic 3-component velocity seismogram on the free surface of the model with a recording length of 40 s. This recorded synthetic wavefield is used as our “data” to image the subsurface structure. Figure 1.2b shows an example of 3-component seismograms for a regional source represented by the green star #3 in Figure 1.2a. Figure 1.2d shows an example of 3-components seismograms of a plane wave with a back azimuth of 0 degree (due north) and an incidence angle of 27 degrees (from the vertical, the black arrow in Figure 1.2c). In both examples, the main P arrival and its coda are color-coded and presented on the 3-component synthetic seismograms (Figures 1.2b and 1.2d). Due to the complexity of the velocity model (relative to 1D models with flat interfaces) and the differences in the wave paths, the converted S-wave arrivals have complex traveltimes-distance curvatures, so the move-out and stacking procedure in the CCP method will not stack the converted waves from the vertical interfaces and sharp corners effectively (Shang et al., 2012). Furthermore, since the CCP method usually uses only the radial component receiver function for imaging purposes, it does not utilize the information provided by other components.

To construct the receiver-side elastic wavefield, we use seismograms on the evenly distributed finite-difference grids on the free surface (340 by 340 grids with about 0.555-km station spacing), reverse the 3-component seismograms in time, set them as the boundary condition on the surface of the study region (equation 1.8), then use the elastic wave solver of Zhang et al. (2012) to back-propagate the ground motion on the surface into the model interior by solving elastic wave equations:

$$\rho \frac{\partial^2 \mathbf{u}^r}{\partial t^2} = \nabla \cdot \boldsymbol{\sigma}^r, \quad (1.5)$$

$$\boldsymbol{\sigma}^r = \mathbf{c} : \boldsymbol{\varepsilon}^r, \quad (1.6)$$

$$\boldsymbol{\varepsilon}^r = \frac{1}{2} [\nabla \mathbf{u}^r + (\nabla \mathbf{u}^r)^T], \quad (1.7)$$

$$\mathbf{u}^r(z = 0, t) = \mathbf{u}^s(z = 0, T - t). \quad (1.8)$$

These equations are similar to equations (1.1-1.4) except that the superscript r represents the receiver-side variables and the free surface boundary condition is replaced with the time-reversed ground motion boundary condition. The external force term is removed in equation (1.5) because we don't add any earthquake sources for back propagation. Different from forward propagation, in which the wavefield satisfies the causal solution of the wave equation, the back propagation

should satisfy the anti-causal solution. In this case, we add 12 PML layers above the free surface to absorb any waves propagate upwards and reflect from the free surface that contaminates the back-propagated wavefield.

The model used to back-propagate the synthetic seismograms is a simple two-layer model with an interface at 40 km depth and no vertical and lateral Moho offsets (Figure 1.1b). We purposefully simplify the model for back propagation from the true model but still keep a sharp velocity discontinuity in it, in contrast to the smoothed models used in Shang et al. (2012). The goal is to demonstrate the recovery of the Moho offsets, without the artifacts of the discontinuity in the model used for back propagation.

To illustrate the possible artifacts in the back-propagated wavefield and the reason for separating the main P arrival and its coda in our new approach, we first back-propagate the entire time-reversed wave train, including the main arrival (P) and its coda, as in Shang et al. (2012). Figure 1.3 shows the snapshots at time=12.3 s of the 3-component forward- and backward-propagated wavefields. The converted wave generated from the shallow Moho (at 30 km depth), deep Moho (at 50 km depth) and vertical interfaces are identified in the receiver-side wavefield and marked as thick black dashed lines 1, 2 and 3, respectively in Figure 1.3. By comparing the forward- and backward-propagated wavefields, we can also identify artifacts in the receiver-side wavefield that could contaminate the separation of the P and S modes and construction of the subsurface images.

These artifacts can be attributed to two causes. First when we numerically back-propagate P wave (or S wave), it also generates artificial, lower amplitude S wave (or P wave), because the corresponding stress field on the surface is not recorded (Ravasi and Curtis, 2013a and 2013b) and thus the boundary condition is incomplete. The artifact caused by incomplete boundary condition is small in amplitude compared to the direct P wave, but comparable to the converted S wave (see the thin dotted line in Figure 1.3d and the white dash line A in Figure 1.4b). Second, when we use a reference velocity model with discontinuity interfaces to back-propagate the recorded waves, the resulting waves generate undesirable reflected and converted waves at the interfaces, which do not exist in the forward wavefield (black dash line B in Figure 1.4b). This artifact, if not removed, results in discontinuity interfaces in the RTM images that are inherited from the reference velocity model (white dash line in Figure 1.4c). Finally, when the P coda contains multi-reflected P waves, such as Pp_mP (P wave reflected at the free surface and the Moho), these phases will be back-propagated together with the converted S wave, complicating the interpretation.

1.2.3 Isolating P and S modes (Steps 1-3)

To solve the problems identified in the above section, we propose several new computational approaches in the data and image domains. In the data domain, we separate the three-component seismograms into the main P arrival and its time-

windowed coda that contains the expected converted S-wave (Figures 1.2b and 1.2d). In the examples shown, the P-wave window is from the beginning of the seismograms to 2 s after the peak of the direct P arrival; the coda wave window is from the end of the P wave window to the end of the seismograms. We taper the P and coda windows at the both ends of the seismograms with a one side Hann window (10s width Hann window) before back propagate them separately.

We propose a modified Helmholtz decomposition operator to further isolate P and S waves in the back-propagated wavefields. In the image domain, several decomposition methods have been developed (Zhang and McMechan, 2010). The commonly used method to separate the P and S waves with Helmholtz decomposition is shown in equations (1.9) and (1.10) (Morse and Feshbach, 1954; Zhang and McMechan, 2010):

$$P(x, y, z, t) = \nabla \cdot \mathbf{u}^r(x, y, z, t), \quad (1.9)$$

$$\mathbf{S}(x, y, z, t) = \nabla \times \mathbf{u}^r(x, y, z, t), \quad (1.10)$$

where P and $\mathbf{S}$ are the P and S modes, respectively, $\mathbf{u}^r$ is the receiver-side displacement wavefield generated by integrating the velocity wavefield through time. Notice that the P mode is a scalar wavefield and S mode is a vector wavefield, and the vector direction information of P and S wave are lost.

In our modified Helmholtz decomposition, the P- and S-mode separation is achieved with equations (1.11) and (1.12),

$$\mathbf{P}(x, y, z, t) = -\nabla(\nabla \cdot \mathbf{u}^{pr}(x, y, z, t)), \quad (1.11)$$

$$\mathbf{S}(x, y, z, t) = \nabla \times (\nabla \times \mathbf{u}^{sr}(x, y, z, t)), \quad (1.12)$$

where $\mathbf{u}^{pr}$ is the displacement wavefield constructed from the back-propagated main P arrival, $\mathbf{u}^{sr}$ is the displacement wavefield constructed from the back-propagated windowed coda waves that contains the converted S wave. Here both the P and S modes are vector wavefields and the particle motions are consistent with the original P and S waves. More complex Helmholtz operators preserving the original physical units, phase, amplitude and vector characteristics (Zhang and McMechan, 2010; Brytik et al., 2011) are computationally more expensive and beyond the scope of this Chapter. Since we use modified Helmholtz decomposition to isolate the S mode in the back-propagated coda wavefield, we also suppress the multiple reflections of P wave in the coda (e.g., Pp_mP).

Figure 1.5 shows the P- and S-mode wavefields, obtained by separating the direct P wave and converted S wave in the data domain (Figure 1.2b and 1.2d), back-propagating them separately and extracting them with modified Helmholtz decomposition. Compared to the receiver-side displacement wavefield from back-

propagation of the entire wave train (Figures 1.3b, d and f), the artifacts are removed in the P and S modes (Figure 1.5). We note that the results of our modified Helmholtz decomposition (equations 1.11 and 1.12) resemble particle acceleration with side lobes on both sides of the main peak. For imaging purpose, it does not affect the location of the main peak, and thus it won't alter the location of the interface in the image.

1.2.4 Imaging (Steps 4 and 5)

With traditional Helmholtz decomposition (equations (1.9) and (1.10)), the conventional imaging condition (Duan and Sava, 2015) is:

$$I_{ni}(x, y, z) = \int_0^T P_n(x, y, z, t; n) \mathbf{S}_{ni}(x, y, z, t; n) dt, \quad (1.13)$$

$$I_i^{total}(x, y, z) = \sum_{n=1}^{n_{src}} I_{ni}(x, y, z). \quad (1.14)$$

Here n is the index of seismic sources, n_{src} is the total number of sources, and i is x, y or z component. P and $\mathbf{S}$ are the P mode and S mode obtained with traditional Helmholtz decomposition. I_{ni} is the i component image generated with an individual source n and I_i^{total} is the i component image after stacking images from all sources. For each source, the zero-offset cross correlations of different components of the P mode and different components of the S mode are calculated at

each point in the image domain (equation 1.13). Then the images of different sources are stacked to generate the final image (equation 1.14). Notice that the S mode is a vector wavefield (equation 1.10) and the P mode is a scalar wavefield, so equations (1.13) and (1.14) result in a vector image (3 image for 3 components), which may lead to difficulties in geological interpretation (Duan and Sava, 2015). Furthermore, this imaging condition can result in the change of the polarity of the S mode at the normal incident point, which may cause destructive contributions to the stacked migrated result from difference sources (Duan and Sava, 2015; Wang et al, 2015; Figure 1.6).

To overcome these issues, we use the modified Helmholtz decomposition operator (equations 1.11 and 1.12), which preserves the particle motion direction, and apply a new imaging condition:

$$I_n(x, y, z) = \int_0^T \text{sign}(\mathbf{P}_n(x, y, z, t; n) \cdot \mathbf{S}_n(x, y, z, t; n)) \cdot |\mathbf{P}_n(x, y, z, t; n)| \cdot |\mathbf{S}_n(x, y, z, t; n)| dt, \quad (1.15)$$

$$I^{total}(x, y, z) = \sum_{n=1}^{n_{src}} I_n(x, y, z) \cdot w_n \quad (1.16)$$

Here n is the index of seismic sources and n_{src} is the total number of sources. $\mathbf{P}$ and $\mathbf{S}$ are the P mode and S mode obtained with modified Helmholtz decomposition

(equations 1.11 and 1.12). I_n is the image generated with an individual source n and I^{total} is the image after stacking images from all sources. w_n is the weight of n th image which can be determined by the quality of the image (such as signal to noise ratio and maximum amplitude of the image).

Following Wang et al. (2015), we use the sign of the dot product of the P and S modes to correct the polarity of the image, as it maintains the polarity on both sides of the normal incident point for a given velocity contrast (from low-to-high velocities or high-to-low velocities). However, for incoming P waves with small incident angles, such as teleseismic arrivals, the dot product of the P and S modes is close to zero. So unlike in Wang et al. (2015), we use the product of the absolute P and S modes to preserve the magnitude of the converted phase. This imaging condition creates one image instead of three, because product of absolute P and S modes is a scalar instead of a vector (Wang et al., 2015). After that we normalize individual images with different weights based on the quality of the images, and stack the normalized images of different sources to generate our final image with equation (1.16). In the synthetic tests, we use the reciprocal of the maximum amplitude of the image as the weight. For real applications, the weights depend on the noise level and illumination of the images.

1.3 Numerical examples

We use the same velocity model introduced in the Method section (Figure 1.1a) to demonstrate the results of the 3D PS-RTM method. In this example, we run two synthetic experiments using two types of sources. In the first experiment (Case 1), we use deep regional sources in the model (Figure 1.2a), which may represent deep earthquakes in a subduction zone setting, to illuminate the study region. For simplicity, we use evenly distributed explosive sources with a Gaussian source time function having a central frequency of ~ 1 Hz. The sources are located at 90 km depth and have the same magnitude. In the second experiment (Case 2), we use plane waves with the incidence angle ranging from 12 to 27 degrees, the back-azimuth angle from 0 to 360 degrees (Figure 1.2c), and a central frequency of ~ 1 Hz. Since a plane wave is an approximation for teleseismic arrivals, this experiment represents the case of using teleseismic sources with different epicentral distances and back-azimuths. The simplification of the sources is justifiable as a source-equalization and de-convolution step can be applied to real earthquake data to remove the source effects.

In Case 1, the results obtained with the new imaging condition (equations 1.15 and 1.16) and the background model in Figure 1.1b for back-propagation for 6 different sources (green stars in Figure 1.2a) are presented in Figure 1.7, showing that different parts of the model are illuminated by the different sources. For example, sources 1 and 6, which are located at the southeast and northwest corners, illuminate the southeast and northwestern parts of our model, respectively. Sources 2, 3, 4 and 5, which are located at the middle of the model, illuminate the middle

part of the model. We note that the polarity of the same interfaces remains the same in all the images. After stacking all the images created with 36 different sources, the stacked image shows nearly uniform amplitude in the whole model (Figure 1.8a). Notice that in the horizontal slice (Figure 1.8b), the vertical interface at the eastern end (positive Y direction) is not well imaged, this is because the 36 sources are located near the center, so they don't provide full illumination for the vertical structure near the edge of the model, which is a limitation caused by source distribution.

In Case 2, we use 126 plane waves with different incident angles (12-27 degrees with a 2.5-degree interval) and back-azimuths (0-360 degree with a 20-degree interval) to represent teleseismic P arrivals to generate synthetic seismograms. This case more closely resembles receiver function studies in terms of the source-station geometry. Compared to Case 1, the whole volume of our model is well illuminated (Figure 1.8c) due to a larger and more complete source aperture.

We note that the interfaces in the horizontal slice (Figure 1.8d) have different polarities. For the horizontal interfaces the polarity reflects whether P wave travels from high- to low-velocity media (positive), or from low- to high-velocity media (negative). For the vertical interfaces in this model, the polarity of the interface is more complicated: the plane wave travel either from the high- to the low-velocity medium or vice versa, depending on back-azimuths, so the polarity of the interface

in the final stacked image depends on the majority of the converted wave record by the stations.

For the vertical interface near the model edge ($Y = 150$ km in Figure 1a), a P wave traveling in the negative Y direction enters the low-velocity layer first before it passes the vertical interface with a low-to-high velocity contrast. The converted phase from the interface is then fully recorded by the stations. For a P wave traveling in the opposite (positive Y) direction, it passes the vertical interface with a high-to-low velocity contrast, but some of the converted phase travels out of the model boundary. So the converted waves from P waves from the negative Y direction likely dominate in the stack, leading to a negative-polarity interface in the migrated image (blue line near $Y=150$ km in Figure 1.8d).

For the vertical interface near the center of the model ($Y = 111$ km), the positive polarity (Figure 1.8d) requires a different explanation, which we attribute to the weakening of the P-wave coming from the negative Y direction at the deep Moho by reflection and conversion and then the weakening of the converted S wave at the shallow Moho by reflection and conversion. So the polarity of this interface is dominated by the P wave coming from positive Y direction, which is from high- to low-velocity medium (positive polarity).

This example shows that although the new image condition (equation 1.15) can account for the polarity change on both sides of the normal incident point, care

should be taken in interpreting high-dipping-angle and vertical interfaces as the sense of velocity contrast may change with back azimuths of incoming P waves and it may be necessary to stack the images by different back azimuth groups.

The separation of the main P arrival and its coda in the data domain and their separate back-propagation to isolate the P and S modes allow us to use realistic velocity models with sharp discontinuities (Figures 6), though the tradeoff is doubling of the memory requirement used in back-propagation, which leads to about four times longer computation than back-propagating the whole wave train and separating the P and S modes only in the image domain.

The computational cost of 3D PS-RTM is generally proportional to the number of earthquakes used in imaging. In our study, calculation for each earthquake needs about 2 hours wall-clock time with 40 CPU cores on a high-performance cluster at the University of Rhode Island. To calculate the wavefields of 126 teleseismic sources, we need a total of about 24 hours with 400 CPU cores. Bases on this experience, we anticipate that application of 3D PS-RTM will be feasible at regional (a few hundred kilometers to a thousand kilometers) scales in the near future, given an adequate seismic array and moderate (100s to 1000s CPU cores) high-performance computational resources.

1.4. Conclusions

We have developed a 3D PS-RTM method to image velocity discontinuities and scatters. Compared with the CCP method, this method is better suited to image complex structures (e.g., Moho offsets, sutures, subducting slabs). The new method builds on the 2D PS-RTM method of Shang et al. (2012) with several significant differences and improvements. First, we extend the method to 3D using a finite-difference wave equation solver in the spherical coordinate (Zhang et al., 2012), so it can be applied to 3D geological structures at regional and global scales. Second, we separate the main P arrival and its coda containing the converted S wave in the data domain and back propagate them separately. This new procedure doubles the memory requirement, but suppresses the artificial converted waves generated by the discontinuities in the reference velocity model in the back-propagated wavefields, allowing us to use realistic velocity models with discontinuities. The new procedure also suppresses the multiple P reflections (e.g., P_{pmp}) and artifacts caused by the incomplete boundary condition, though converted P-to-S phases from shallow reverberations could be falsely mapped as a deep structure, as in P wave receiver functions. We apply a modified Helmholtz decomposition operator to isolate P and S wave energy in the separate, back-propagated wavefields, then apply a new imaging condition to the P- and S-mode wavefields to generate images. The new image condition maintains the polarity on both sides of the normal incident point and thus yields a consistent polarity of the converted phases for a given velocity contrast (from low-to-high or high-to-low velocities) for sources with different azimuths and incidence angles, making it more straightforward to stack images obtained from earthquake sources from different azimuths and incidence

angles. With these new data processing and computational procedures, the 3D PS-RTM can use regional seismic arrays to image 3D structures with a realistic background velocity model. Though care should be taken for vertical and high-dipping angle interfaces, for which the sense of the velocity contrast may change with the back azimuths of incoming waves.

Bibliography:

- Abe, Y., T. Ohkura, K. Hirahara, and T. Shibutani (2011), Common-conversion-point stacking of receiver functions for estimating the geometry of dipping interfaces, *Geophysical Journal International*, 185(3), 1305-1311, doi:10.1111/j.1365-246X.2011.05001.x.
- Ai, Y., Q.-f. Chen, F. Zeng, X. Hong, and W. Ye (2007), The crust and upper mantle structure beneath southeastern China, *Earth and Planetary Science Letters*, 260(3), 549-563, doi:<https://doi.org/10.1016/j.epsl.2007.06.009>.
- Anderson, K. R., et al. (2016), Full wavefield recordings of Oklahoma seismicity from an IRIS-led community experiment, *American Geophysical Union Fall Meeting, 2016*, @AbstractS11D-2493.
- Baysal, E., D. D. Kosloff, and J. W. C. Sherwood (1983), Reverse time migration, *Geophysics*, 48(11), 1514-1524, doi:10.1190/1.1441434.
- Bostock, M. G., S. Rondenay, and J. Shragge (2001), Multiparameter two-dimensional inversion of scattered teleseismic body waves 1. Theory for oblique incidence, *Journal of Geophysical Research: Solid Earth*, 106(B12), 30771-30782, doi:10.1029/2001JB000330.
- Brytik, V., M. V. de Hoop, H. F. Smith, and G. Uhlmann (2011), Decoupling of Modes for the Elastic Wave Equation in Media of Limited Smoothness, *Communications in Partial Differential Equations*, 36(10), 1683-1693, doi:10.1080/03605302.2011.558554.

- Chang, W. F., and G. A. McMechan (1987), Elastic reverse - time migration, *GEOPHYSICS*, 52(10), 1365-1375, doi:10.1190/1.1442249.
- Chang, W. F., and G. A. McMechan (1994), 3-D elastic prestack, reverse - time depth migration, *GEOPHYSICS*, 59(4), 597-609, doi:10.1190/1.1443620.
- Chen, L., L. Wen, and T. Zheng (2005a), A wave equation migration method for receiver function imaging: 1. Theory, *Journal of Geophysical Research: Solid Earth*, 110(B11), n/a-n/a, doi:10.1029/2005JB003665.
- Chen, L., L. Wen, and T. Zheng (2005b), A wave equation migration method for receiver function imaging: 2. Application to the Japan subduction zone, *Journal of Geophysical Research: Solid Earth*, 110(B11), n/a-n/a, doi:10.1029/2005JB003666.
- Cheng, C., T. Bodin, and R. M. Allen (2016), Three-dimensional pre-stack depth migration of receiver functions with the fast marching method: a Kirchhoff approach, *Geophysical Journal International*, 205(2), 819-829, doi:10.1093/gji/ggw062.
- Cheng, C., T. Bodin, B. Tauzin, and R. M. Allen (2017), Cascadia subduction slab heterogeneity revealed by three-dimensional receiver function Kirchhoff migration, *Geophysical Research Letters*, 44(2), 694-701, doi:10.1002/2016GL072142.
- Duan, Y., and P. Sava (2015), Scalar imaging condition for elastic reverse time migration, *GEOPHYSICS*, 80(4), S127-S136, doi:10.1190/geo2014-0453.1.
- Dueker, K. G., and A. F. Sheehan (1997), Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track,

- Journal of Geophysical Research: Solid Earth*, 102(B4), 8313-8327,
doi:10.1029/96JB03857.
- Gao, S. S., and K. H. Liu (2014), Mantle transition zone discontinuities beneath the contiguous United States, *Journal of Geophysical Research: Solid Earth*, 119(8), 6452-6468, doi:10.1002/2014JB011253.
- Graves, R. W. (1996), Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bulletin of the Seismological Society of America*, 86(4), 1091-1106.
- Kind, R., X. Yuan, and P. Kumar (2012), Seismic receiver functions and the lithosphere–asthenosphere boundary, *Tectonophysics*, 536, 25-43,
doi:<http://dx.doi.org/10.1016/j.tecto.2012.03.005>.
- Levander, A., F. Niu, and W. W. Symes (2013), Imaging Teleseismic P to S Scattered Waves Using the Kirchhoff Integral, in *Seismic Earth: Array Analysis of Broadband Seismograms*, edited, pp. 149-169, American Geophysical Union,
doi:10.1029/157GM10.
- Levander, A. R. (1988), Fourth - order finite - difference P-SV seismograms, *GEOPHYSICS*, 53(11), 1425-1436, doi:10.1190/1.1442422.
- Liu, K., and A. Levander (2013), Three-dimensional Kirchhoff-approximate generalized Radon transform imaging using teleseismic P-to-S scattered waves, *Geophysical Journal International*, 192(3), 1196-1216, doi:10.1093/gji/ggs073.
- McMechan, G. A. (1983), Migration by extrapolation of time-dependent boundary values, *Geophysical Prospecting*, 31(3), 413-420, doi:10.1111/j.1365-2478.1983.tb01060.x.

- Morse, P. M., and H. Feshbach (1954), Methods of theoretical physics, *American Journal of Physics*, 22(6), 410-413.
- Poppeliers, C., and G. L. Pavlis (2003a), Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 1. Theory, *Journal of Geophysical Research: Solid Earth*, 108(B2), n/a-n/a, doi:10.1029/2001JB000216.
- Poppeliers, C., and G. L. Pavlis (2003b), Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 2. Stacking multiple events, *Journal of Geophysical Research: Solid Earth*, 108(B5), n/a-n/a, doi:10.1029/2001JB001583.
- Ravasi, M., and A. Curtis (2013a), Elastic imaging with exact wavefield extrapolation for application to ocean-bottom 4C seismic data, *GEOPHYSICS*, 78(6), S265-S284, doi:10.1190/geo2013-0152.1.
- Ravasi, M., and A. Curtis (2013b), Nonlinear scattering based imaging in elastic media: Theory, theorems, and imaging conditions, *GEOPHYSICS*, 78(3), S137-S155, doi:10.1190/geo2012-0286.1.
- Revenaugh, J. (1995), A Scattered-Wave Image of Subduction Beneath the Transverse Ranges, *Science*, 268(5219), 1888-1892.
- Rondenay, S., M. G. Bostock, T. M. Hearn, D. J. White, and R. M. Ellis (2000), Lithospheric assembly and modification of the SE Canadian Shield: Abitibi-Grenville teleseismic experiment, *Journal of Geophysical Research: Solid Earth*, 105(B6), 13735-13754, doi:10.1029/2000JB900022.

- Ryberg, T., and M. Weber (2000), Receiver function arrays: a reflection seismic approach, *Geophysical Journal International*, 141(1), 1-11, doi:10.1046/j.1365-246X.2000.00077.x.
- Shang, X. (2014), Inverse scattering: theory and application to the imaging of the Earth's seismic discontinuities, Massachusetts Institute of Technology.
- Shang, X., M. V. de Hoop, and R. D. van der Hilst (2012), Beyond receiver functions: Passive source reverse time migration and inverse scattering of converted waves, *Geophysical Research Letters*, 39(15), n/a-n/a, doi:10.1029/2012GL052289.
- Sheehan, A. F., P. M. Shearer, H. J. Gilbert, and K. G. Dueker (2000), Seismic migration processing of P-SV converted phases for mantle discontinuity structure beneath the Snake River Plain, western United States, *Journal of Geophysical Research: Solid Earth*, 105(B8), 19055-19065, doi:10.1029/2000JB900112.
- Shen, Y., S. C. Solomon, I. T. Bjarnason, and C. J. Wolfe (1998), Seismic evidence for a lower-mantle origin of the Iceland plume, *Nature*, 395(6697), 62-65, doi:http://www.nature.com/nature/journal/v395/n6697/supinfo/395062a0_S1.html.
- Tauzin, B., R. D. van der Hilst, G. Wittlinger, and Y. Ricard (2013), Multiple transition zone seismic discontinuities and low velocity layers below western United States, *Journal of Geophysical Research: Solid Earth*, 118(5), 2307-2322, doi:10.1002/jgrb.50182.

- Wang, C., J. Cheng, and B. Arntsen (2015), Imaging Condition for Converted Waves Based on Decoupled Elastic Wave Modes, in *SEG Technical Program Expanded Abstracts 2015*, edited, pp. 4385-4390, doi:10.1190/segam2015-5920952.1.
- Zhang, J., and T. Zheng (2015), Receiver Function Imaging with Reconstructed Wavefields from Sparsely Scattered Stations, *Seismological Research Letters*, 86(1), 165-172, doi:10.1785/0220140028.
- Zhang, Q., and G. A. McMechan (2010), 2D and 3D elastic wavefield vector decomposition in the wavenumber domain for VTI media, *GEOPHYSICS*, 75(3), D13-D26, doi:10.1190/1.3431045.
- Zhang, W., and Y. Shen (2010), Unsplit complex frequency-shifted PML implementation using auxiliary differential equations for seismic wave modeling, *GEOPHYSICS*, 75(4), T141-T154, doi:10.1190/1.3463431.
- Zhang, W., Y. Shen, and L. Zhao (2012), Three-dimensional anisotropic seismic wave modelling in spherical coordinates by a collocated-grid finite-difference method, *Geophysical Journal International*, 188(3), 1359-1381, doi:10.1111/j.1365-246X.2011.05331.x.

Figure 1.1. (a) The velocity model used to generate synthetic seismograms. It contains two layers with different P- and S-wave velocities as described in the text. The grey surface represents the velocity discontinuity interface between the two layers. The thin white lines outline three slices of the model at $X = 138$ km, $Y = 138$ km and $Z = -40$ km. (b) The simplified reference velocity model used to back-propagate the recorded seismograms and image the discontinuity interfaces. The two layers have the same P- and S-wave velocities as in (a).

Figure 1.2. (a) Geometry of the sources (red and green stars) and selected stations (blue triangles) for the regional-scale synthetic experiment; (b) Synthetic seismograms of source 3 recorded by the selected stations in (a). The P arrival and its coda are color coded as red and blue, respectively, on the three components of the seismograms. (c) Geometry of the selected stations (blue triangles) and the propagation direction of a plane wave with a back-azimuth angle of 0° and incidence angle of 27° (black arrow); (d) Synthetic seismograms for the plane wave and stations in (c). The P arrival and its coda are color coded as in (c).

Figure 1.3. A snapshot at time =12.3 s on a vertical slice across the middle of the model (at $x=111$ km) of the forward (left panels: a, c, e) and backward (right panels: b, d, f) propagated wavefields for source 3 in Case 1 (see Figure 2). The reconstructed wavefield is obtained by back propagating the entire recorded seismic wave. The rows represent the different components of the displacement wavefield (from the top to bottom, the x , y , z component, respectively). The converted wave from the shallow Moho (at 30 km depth) is marked as number 1; the converted wave from the deep Moho (at 50 km depth) is marked as number 2; the converted wave from the vertical interfaces is marked as number 3. The artifact caused by the incomplete boundary condition is marked with a thin dash line.

Figure 1.4. Comparison of the snapshot of the Y component of the S model at time = 12.3 s, obtained from separately back-propagated coda wavefield (a) and back-propagation of the P and coda together (b). The white dashed line A in (b) is the artifact caused by the incomplete surface boundary condition. The thin black dashed line B is the artifact caused by spurious reflected and converted waves at the interface of the reference model (at depth of 40 km, Figure 1b). The thick black dashed lines represent converted S waves. (c) Image obtained when we back-propagate the entire P and P coda together and separate the P and S modes only in the image domain. The dashed line represents the artifact caused by spurious reflected and converted waves. The two vertical slices are at $X = 138$ km and $Y = 138$ km as outlined in Figure 1a.

Figure 1.5. A snapshot at time =12.3 s on a vertical slice across the middle of the model (at $x=111$ km) of the P (left panels) and S (right panels) mode wavefields for source 3 in Case 1, obtained by back-propagating the direct P wave and P coda separately. Different rows from the top to bottom represent the x, y and z components, respectively. The color scales for the two columns are normalized. The converted waves from the shallow and deep Moho interfaces (marked as 1 and 2, respectively) and the vertical interfaces (marked as 3) can be clearly identified in the S mode wavefield. Notice the differences in the curvatures of the converted phases from the flat Moho (1 and 2) and the vertical interfaces (3).

Figure 1.6. Results obtained with the conventional imaging condition (Equations 1.13 and 1.14). (a), (c) and (e) show the images on the vertical slices $X=138\text{km}$ and $Y=138\text{km}$, while (b), (d) and (f) show the map view of the images on the horizontal slice $Z=40\text{km}$. (a) and (b) are the result from the P mode and the x component of the S mode; (c) and (d) are the result from the P mode and the y component of the S mode; (e) and (f) are the result from the P mode and the z component of the S mode. Notice the change in the polarity of the image at different locations of the discontinuity.

Figure 1.7. Results obtained with the modified imaging condition (Equation 1.15) for the 6 difference sources marked in Figure 2a. The two vertical slices are at $X = 138$ km and $Y = 138$ km as outlined in Figure 1a.

Figure 1.8. (a) and (b) are stacked images with the imaging condition of Equation 1.16 using the sources in Case 1. (a) is the image result on the two vertical slices $X = 138$ km and $Y = 138$ km as outlined in Figure 1a; (b) is the image result on the horizontal slice $Z = -40$ km as outlined in Figure 1a. (c) and (d) are the same as (a) and (b), but obtained with teleseismic sources.

CHAPTER 2

Numerical tests of 3D passive-source reverse-time migration of interpolated wavefields recorded by sparse seismic arrays

By

Jiahang Li¹ and Yang Shen¹

Part of the materials presented in this chapter was accepted as supporting information in Li et al. (2018) by *Journal of Geophysical Research: Solid Earth*

---

1. *Graduate school of Oceanography, University of Rhode Island, 215 S Ferry Rd, Narragansett, RI 02882, USA*

2.1 Introduction

In Chapter 1, we demonstrate the ability of 3D passive-source reverse time migration (PS-RTM) in imaging complex subsurface structures such as Moho sutures using converted wave with numerical grids on the surface as seismic stations (about 500 m station spacing). However, due to geography, deployment cost, local policy and other logistic difficulties, seismic arrays are often sparsely and unevenly distributed (e.g., Zhang and Zheng, 2015). Moreover, seismic stations are usually not located exactly at the grids in numerical simulations. Therefore, interpolation of the wavefield recorded by sparsely and unevenly distributed seismic stations onto dense and evenly distributed numerical grids on the surface is necessary to make use of this wave-equation-based method.

Interpolation methods have been widely used in active-source reverse time migration, though in those exploration seismic applications seismometers are often densely and evenly distributed. Thus, the traditional methods (Abma and Kabir, 2005) used in active-source studies are not directly applicable in the PS-RTM method. For passive-source wavefield interpolation, Neal and Pavlis (1999) developed a pseudo-station stacking method based on an aerial smoothing function; Poppeliers and Pavlis (2003a, 2003b) applied pseudo-station stacking with a plane wave migration method to image the lithosphere scale discontinuities under northwestern Colorado; Zhang and Zheng (2015) presented the interpolation result of a cubic spline method with a dense seismic network in Northeastern China; Song

et al. (2017) introduced the Radial Basis Function method to reconstruct teleseismic wavefield with a dense seismic network in South China to image the Moho structures. The advantages and disadvantages of the above methods are discussed in Song et al. (2017). In addition, Shang et al. (2017) proposed a curvelet regularization to reconstruct the surface wavefield for a two-dimensional PS-RTM method. Among these methods, the cubic spline method is a simple and low-cost method that can be easily implemented. In this chapter, we test the efficacy of this interpolation method using synthetic data and then use the interpolated seismograms as the input data for the PS-RTM method.

In the following sections, we first introduce the methodology and procedures of the cubic spline method. Then we use synthetic seismic arrays with different station spacing to investigate the station spacing requirement for the interpolation method, and add random noises of different amplitudes to the synthetic data to investigate the level of tolerance of the method to random noise.

2.2 Method

In this study, we use cubic spline interpolation to interpolate seismograms recorded by sparse seismic stations onto dense numerical grids on the free surface (Zhang and Zheng 2015; Press et al. 2007). First, we align the seismograms along the direct P wave. In this procedure, seismic phases with time-distance move-out curvatures similar to that of the direct P wave are also aligned or nearly aligned. For synthetic

data, the direct P arrival can be easily picked by choosing the time of maximum amplitude in the time series of vertical-component displacement seismograms. Alternatively, this can be achieved with waveform cross-correlation. We align the seismograms by shifting them by the relative arrival times of the P wave at different stations, which can be expressed as in equation (2.1)

$$u_{ij}'(t) = u_{ij}(t - t_i), \quad (2.1)$$

where u_{ij}' is the aligned seismogram, u_{ij} is the original displacement seismogram, t is time, and t_i is the time difference of the direct P arrival between station i and the reference station. The subscript i represents the index of the station while subscript j represents one of the three components (vertical, north and east). For observed data, we use the STA/LTA method (Allen 1978, 1982) to pick the initial direct P-wave arrival times and then the multi-channel cross-correlation method to align the seismograms (VanDecar and Crosson 1990; Rondenay and Fischer 2003).

Based on the time shifts at the stations, a time-shift map can be constructed with the cubic spline method as in equation (2.2)

$$t_k = \sum_{i=1}^N w_{ik} t_i, \quad (2.2)$$

where w_{ik} is the weight station i contribute to numerical grid k , which is calculated with the Matlab function “griddata” using the cubic spline method. The

implementation of the cubic spline method in Matlab is based on triangulation, so w_{ik} is non-zero only at the grids within the three vertexes of the triangle (Press et al. 2007); t_k is the interpolated relative arrival time of the direct P arrival at numerical grid k . N is the total number of seismic stations used.

The cubic spline method is used again to interpolate the seismograms at each time step of the time-shifted seismograms.

$$u_{jk}''(t) = \sum_{i=1}^N w_{ik} u_{ij}'(t), \quad (2.3)$$

Here w_{ik} is the same weight number as equation (2.2). u_{jk}'' is the j component seismogram at grid k after interpolation.

The interpolated seismograms are shifted to the absolute arrival time using the relative arrival time map (equation 2.2) with the following equation:

$$u_{jk}^0(t) = u_{jk}''(t + t_k), \quad (2.4)$$

The shifted seismograms u_{jk}^0 serve as our input data for the 3D PS-RTM package to image the subsurface structure.

2.3 Numerical tests

To test the cubic spline interpolation method, we create a model contains a homogeneous crust and a slab in an otherwise uniform upper mantle (Figure 2.1a). The crust and the background mantle S-wave velocities are 3.900 km/s and 4.480 km/s, respectively; the corresponding P-wave velocities are 6.800 km/s and 8.036 km/s, respectively. The slab structure has a +8% velocity perturbation compared with the background velocity, subducting into the mantle with a 10-degree dipping angle. For simplicity, we only show a vertical profile perpendicular to the slab, which cross the model along the longitudinal direction (Figure 2.1a). The reference velocity model used for back propagation has the same crust and background mantle, but no slab structure (Figure 2.1b). The synthetic seismograms are calculated with the finite difference method (Zhang et al. 2012) using 126 plane waves with different incidence angles (12~27 degrees with a 2.5-degree increment) and back-azimuths (0~360 degrees with a 20-degree increment) propagating upward from the bottom of the model.

2.3.1 Ideal station coverage

We first record the seismogram on each numerical grid on the free surface except the grids near the edge of the model (about 9km wide) and use the recorded seismogram to image the slab structure. In this case, the stations are densely and evenly distributed on the numerical grids with a station spacing of 0.555 km. This test provides an ideal station coverage that serves as a benchmark to compare with the results of more realistic seismic arrays.

The result shows that the PS-RTM method yields good constraints on the boundary of the slab structure (Figure 2.1c): the main peaks in the image match well with the velocity interfaces in the slab model. We note that the lower boundary and the edge of the slab have a negative polarity, which is consistent with the polarity of the converted waves from P-wave traveling from the low-velocity background mantle to a high-velocity slab. In contrast, the top interface of the slab has a positive polarity, like that of the Moho, as expected for the converted waves from P wave traveling from high to low velocities. Below the slab, the horizontal structure at depth of about 115 km is the multiple reflected and converted arrival PpP_{ms} (P reflected at the free surface and converted at the Moho). This example shows that the PS-RTM images based on P wave and its coda also suffer from the interference of shallow, multiple converted phases, as in P receiver functions. By using SdP (S wave converted to P wave at discontinuities) to image the discontinuities with the same method we can compare the results and identify the multiples.

2.3.2 Effect of station spacing

In the above synthetic test, we use the numerical grids on the surface as seismic stations to record the ground motion and back-propagate it to reconstruct the receiver side wavefield. Although it is unrealistic to have such a dense station coverage over a ~ 200 km by 200 km area, there have been deployments designed to record spatially unaliased wavefields over smaller areas (e.g. Anderson et al., 2016).

Deployment of large-N seismic array is a new trend in seismology and we expect to see more of this kind of deployments in the future. To make better use of Large-N arrays, it is worth to investigate the effect of station spacing on the image result.

To illustrate the effect of station spacing on the results of 3D PS-RTM, we design three seismic arrays with an average station spacing of 5 km, 10 km, and 15 km (Figure 2.2 a, c and e respectively). For each individual seismic array, we first design an evenly distributed array and then we change the seismic station location by a random value from 0 to 2.5 km in both latitudinal and longitudinal directions, so the stations are not evenly distributed as in likely real deployments. We note that wavefield spatial aliasing on the surface is a function of horizontal wavelengths (and thus wave frequencies and ray parameters). According to the Nyquist-Shannon sampling theorem, the wavefield can be fully reconstructed from the recorded seismograms at the stations, when the station spacing is smaller than the half apparent wavelength on the surface. Based on the S-wave velocity of the crustal layer in our model (3.9 km/s) and the range of incident angles of plane waves (12 to 27 degrees, which is the general incident angle of teleseismic arrivals that avoiding upper mantle triplication and core interferences), we calculate that the ray parameter is from 0.053 to 0.116 s/km. So for the 1-Hz plane waves used in our numerical experiments, the apparent horizontal wavelength on the surface is from 8.6 to 18.8 km. Thus, the station coverage in these three experiments ranges from unaliased and slightly aliased (5 km station spacing, depending on ray parameter) to severely aliased (15 km station spacing).

We first apply the PS-RTM method to the seismograms recorded by the seismic array with the average station spacing of 5 km without interpolation (Figure 2.2a). The result shows strong artifacts near the surface (Figure 2.3), which suggests that even for a slightly aliased seismic array the artifacts severely contaminate the RTM image. The reason is that the grids without stations have zero values, resulting in a discrete and discontinuous surface boundary condition for back-propagation. So, the interpolation method is necessary to reconstruct a smooth receiver side wavefield to create a continuous image of discontinuity interfaces.

Then we apply the cubic spline method (equation 2.1-2.4) to interpolate the recorded seismograms onto numerical grids on the free surface and use the interpolated seismograms as the input data to apply the PS-RTM method. The interpolated seismograms using seismic arrays with 5 km station spacing and the original synthetic seismograms are shown in Figure 2.4. In this example, we use a plane wave travelling with a 27-degrees incident angle and a back azimuth of about 280 degrees (nearly from the west). The misfits (the interpolated minus the synthetic seismograms) between the synthetic and interpolated seismograms are small. The misfit of the direct P arrival on the z component (up direction) is less than 10% of the amplitude of the direct P arrival, indicating that the interpolation has reconstructed the wavefield recorded on the free surface fairly well for the unaliased and slightly aliased wavefield. Furthermore, for the x-component (co-latitudinal direction, near transverse component), the amplitude is comparable with

the y-component (longitudinal direction, near radial component)(Figure 2.4a). This suggests that for a tele-seismic event that is located off the great circle path of the source and a 2D linear array by even a small angle (10 degrees in this case), the seismic energy on the transverse component cannot be ignored. This limits the application of the 2D PS-RTM method since it can only use limited sources located near the great circle of the 2D profile within a small azimuth range (Shang et al. 2017).

The results of PS-RTM using the interpolated seismogram are shown in Figure 2.5. For unaliased and weakly aliased wavefields (5 km average station spacing), the resulting image (Figure 2.5a) is nearly identical to the image created with the ideal station coverage (Figure 2.1c). For aliased and strongly aliased wavefields (10 km and 15 km average station intervals), the steep interface at the truncated deep end of the slab deteriorates, though the horizontal interface (the Moho) and the gently dipping upper and lower interfaces of the slab remain well imaged (Figure 2.5b and 2.5c). We attribute the differences to the relative moveout (time-distance curve/surface) between the converted phases and the P arrivals. The relative moveout is small for the converted waves from the flat and gently dipping interfaces, so waveforms at adjacent stations aligned on the P arrival are similar enough in time for interpolation to reconstruct the converted phases even at 10-15 km station spacing. On the other hand, the relative moveout of the converted phase from the steep interface at the truncated end of the slab is much larger, so the reconstruction of the converted phase by interpolation works well only at small station spacing

(Figure 2.5a). Other more sophisticated interpolation methods that can identify and link arrivals with various moveout curves/surfaces at adjacent stations are worth exploring to improve the wavefield reconstruction in applications (Jin et al., 2017; Shang et al., 2017; Ainiwaer and Gurrola, 2018).

2.3.3 Effect of random noise

To show the effect of random noise, we present another experiment, in which white noise is added to the 3 component seismograms, resulting in a signal to noise ratio (SNR) of 3, 10 and 20 (Figure 2.6). The SNR is defined as the peak amplitude of the direct P arrival divided by the standard deviation of the amplitude of random noise. In the SNR=20 case, the amplitude of noise is at about the same level as the amplitude of converted waves.

We use the same interpolation method as discussed above to interpolate the seismograms recorded by 5-km-spacing seismic stations onto the numerical grids on the free surface. The RTM result using seismograms with SNR=20 (Figure 2.7c) shows that even with a noise level comparable with the converted wave signal, we can still image the deep structure in the upper mantle. However, in the shallow crust there are some artifacts due to the random noise and the lack of overlapping back-propagated waves from multiple stations to suppress the noise. Since we are interested in the upper mantle structure in this example and the slab structures are well imaged, the artifact in the crust is not an issue. For the image result using

seismograms with $\text{SNR}=10$, in which the amplitude of white noise is about 2 times larger than that of the converted wave, the slab structure is still well imaged with more random noise in the background. The shallow structures are contaminated more by the artifacts than in the case with $\text{SNR}=20$. For the image result using seismograms with $\text{SNR}=3$, the image is totally contaminated by random noise. This suggests that we should be able to successfully image the deep structure with converted waves that have similar or slightly lower amplitudes than those of random noise. We attribute this to two factors: First, the converted waves from multiple stations converge coherently at depth, while random noise does not; and second, stacking of the images for teleseismic sources with difference back azimuths and incidence angles (or ray parameters) suppresses random noise.

2.4 Summary

In this chapter, we tested the cubic spline method in interpolating the surface wavefield to make use of sparsely and unevenly deployed seismic arrays with the 3D PS-RTM method. We first showed the RTM result without wavefield interpolation for a seismic array with about 5 km station spacing. The image is contaminated by strong artifact due to a discontinuous wavefield used in back-propagation simulation. By applying the cubic spline method to seismograms recorded by stations with different spacing, we showed that the interpolation method recovers the receiver-side wavefield as long as the station spacing is smaller than the half apparent wavelength of the converted waves we are interested in. At

larger station spacing, the cubic spline method fails to recover the converted wave generated at steep interfaces due to the differences in the move-out curves of the direct P arrival and the converted S wave. Though the most straightforward solution to this issue is to deploy denser seismic arrays, with limited resources, it is worth developing more sophisticated interpolation methods that account for the move-out curvatures of difference phases. The 3D PS-RTM method using data interpolated with the cubic spline method has a good tolerance of random noise. As long as the noise level is comparable with or even slightly larger than the converted wave we are interested in, we can image the discontinuities with seismic sources from different back-azimuths and ray parameters.

Bibliography:

- Abma, R., and N. Kabir (2005), Comparisons of interpolation methods, *The Leading Edge*, 24(10), 984-989, doi:10.1190/1.2112371.
- Ainiwaer, A., and H. Gurrola (2018), Wavefield iterative deconvolution to remove multiples and produce phase specific Ps receiver functions, *Geophysical Journal International*, 212(3), 2016-2030, doi:10.1093/gji/ggx454.
- Allen, R. (1982), Automatic phase pickers: Their present use and future prospects, *Bulletin of the Seismological Society of America*, 72(6B), S225-S242.
- Allen, R. V. (1978), Automatic earthquake recognition and timing from single traces, *Bulletin of the Seismological Society of America*, 68(5), 1521-1532.
- Anderson, K. R., et al. (2016), Full wavefield recordings of Oklahoma seismicity from an IRIS-led community experiment, *American Geophysical Union Fall Meeting*, 2016, @AbstractS11D-2493.
- Cheng, C., T. Bodin, B. Tauzin, and R. M. Allen (2017), Cascadia subduction slab heterogeneity revealed by three-dimensional receiver function Kirchhoff migration, *Geophysical Research Letters*, 44(2), 694-701, doi:10.1002/2016GL072142.
- Crosson, J. C. V. R. S. (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bulletin of the Seismological Society of America*, 80(1), 150-169.

- Jin, S., S. Chen, J. Wei, and X.-Y. Li (2017), Automatic seismic event tracking using a dynamic time warping algorithm, *J. Geophys. Eng.*, 14, 1138-1149, <https://doi.org/10.1088/1742-2140/aa7309>.
- Li, J., Y. Shen, and W. Zhang (2018), Three-Dimensional Passive-Source Reverse-Time Migration of Converted Waves: The Method, *Journal of Geophysical Research: Solid Earth*, n/a-n/a, doi:10.1002/2017JB014817.
- Neal, S. L., and G. L. Pavlis (1999), Imaging P-to-S conversions with multichannel receiver functions, *Geophysical Research Letters*, 26(16), 2581-2584, doi:10.1029/1999GL900566.
- Poppeliers, C., and G. L. Pavlis (2003a), Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 1. Theory, *Journal of Geophysical Research: Solid Earth*, 108(B2), n/a-n/a, doi:10.1029/2001JB000216.
- Poppeliers, C., and G. L. Pavlis (2003b), Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 2. Stacking multiple events, *Journal of Geophysical Research: Solid Earth*, 108(B5), n/a-n/a, doi:10.1029/2001JB001583.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992), *Numerical recipes in C (2nd ed.): the art of scientific computing*, 994 pp., Cambridge University Press.
- Press, W. H. T., Saul A; Vetterling, William T; Flannery, Brian P. (2007), *Numerical recipes 3rd edition: The art of scientific computing*, Cambridge university press.

- Rondenay, S., and K. M. Fischer (2003), Constraints on localized core-mantle boundary structure from multichannel, broadband SKS coda analysis, *Journal of Geophysical Research: Solid Earth*, 108(B11), n/a-n/a, doi:10.1029/2003JB002518.
- Shang, X., M. V. de Hoop, and R. D. van der Hilst (2017), Common conversion point stacking of receiver functions versus passive-source reverse time migration and wavefield regularization, *Geophysical Journal International*, 209(2), 923-934, doi:10.1093/gji/ggx069.
- Song, P., X. Zhang, Y. Liu, and J. Teng (2017), Moho imaging based on receiver function analysis with teleseismic wavefield reconstruction: Application to South China, *Tectonophysics*, 718, 118-131, doi:https://doi.org/10.1016/j.tecto.2017.05.031.
- Trujillo, M. W. R. A. C. Y. J. B. M. H. S. M. J. (1998), A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bulletin of the Seismological Society of America*, 88(1), 95-106.
- Zhang, J., and T. Zheng (2015), Receiver Function Imaging with Reconstructed Wavefields from Sparsely Scattered Stations, *Seismological Research Letters*, 86(1), 165-172, doi:10.1785/0220140028.
- Zhang, W., Y. Shen, and L. Zhao (2012), Three-dimensional anisotropic seismic wave modelling in spherical coordinates by a collocated-grid finite-difference method, *Geophysical Journal International*, 188(3), 1359-1381, doi:10.1111/j.1365-246X.2011.05331.x.

Figure 2.1. (a) The slab model used to calculate the synthetic seismograms; (b) The reference model used to back-propagate the recorded wavefield and image the slab structure; (c) The RTM result using an ideal station coverage.

Figure 2.2. Station distributions with the average station spacing of 5km, 10km and 15km respectively in (a), (b) and (c). The red triangles are seismic stations and the black triangles are numerical grids used to validate the interpolation result.

Figure 2.3. RTM result using the 5-km-spacing seismic array without interpolation of seismograms on the surface.

Figure 2.4. (a) Comparison of three-component synthetic seismograms (blue lines) and interpolated seismograms (red lines) at numerical grid marked in black triangle in Figure 2.2a. The interpolated seismograms are calculated with the 5-km-spacing seismic array in Figure 2.2a. (b) Misfit between the synthetic and interpolated seismograms in (a).

Figure 2.5. RTM results of the different seismic arrays with station spacing of 5 km, 10 km and 15 km (a, b and c, respectively) and the cubic spline interpolation of seismograms.

Figure 2.6. Radial component seismograms with the signal to noise ratio of 20, 10 and 3, respectively.

Figure 2.7. RTM results using seismograms with different SNR of 3, 10 and 20 (a, b and c, respectively) recorded by the 5-km-spacing seismic array. The cubic spline interpolation method is used in this example.

CHAPTER 3

3D passive-source reverse-time migration imaging of the mantle transition zone beneath the Yellowstone hotspot

By

Jiahang Li¹ and Yang Shen¹

---

3. *Graduate School of Oceanography, University of Rhode Island, 215 S Ferry Rd,
Narragansett, RI 02882, USA*

3.1 Introduction

In Chapters 1 and 2, we demonstrate the ability of the three-dimensional passive-source reverse-time migration (3D PS-RTM) method in imaging complex structures such as Moho sutures and subducting slabs with synthetic data. In this chapter, we present the application of the method to the Yellowstone hotspot and discuss data regularization in practice for an unevenly and sparsely deployed seismic array.

We select Yellowstone as our first application of the 3D PS-RTM for several reasons: First, the Yellowstone hotspot is one of the most studied intraplate hotspots, yet there is on-going debate regarding whether the source of the Yellowstone hotspot comes from a deep mantle plume (e.g., Humphreys et al., 2000; Christiansen et al., 2002; Yuan and Dueker, 2005; Waite et al, 2006; Smith et al., 2009; James et al., 2011; Fouch, 2012; Kincaid et al, 2013; Foulger et al., 2015; Leonard and Liu, 2016; Zhou et al., 2018; Zhou, 2018; Nelson and Grand, 2018); second, there have been previous studies of the mantle transition zone (MTZ) in the region using traditional receiver function methods (Schmandt et al., 2012; Gao and Liu, 2014), so we can compare our RTM results with the previous results; finally, the USArray provides arguably the best coverage of any known hotspot for imaging the MTZ over a broad region.

The depth of origin of hotspot volcanism is fundamental to the mantle plume hypothesis (Morgan, 1971) and has profound implications for plate tectonics (e.g.,

DeMets et al., 2010), mantle convection, and thermal and compositional evolution of the Earth's mantle (e.g., Olson et al, 1990). Shen et al. (1998) suggested using the geometry of the MTZ boundaries, the 410- and 660-km discontinuities in seismic velocities and density, to determine the depth of origin of mantle plumes. The 410- and 660-km discontinuities are generally thought to correspond to the two phase transformations from olivine (α -phase) to spinel (β -phase) crystal structure and from ringwoodite (γ -phase) to perovskite plus magnesiowüstite, with a positive (pressure increases with temperature) and a negative (pressure decreases with temperature) Clapeyron slope, respectively (e.g., Bina and Helffrich, 1994), though the mineral phase assemblage and associated Clapeyron slope change at very high temperature (e.g., Hirose, 2002; Ishii et al., 2018). The opposite Clapeyron slopes of the phase transformations predict that higher temperatures associated with a mantle thermal plume cause a deeper 410-km discontinuity and a shallower 660-km discontinuity within the plume conduit, and therefore a thinner MTZ thickness. Indeed, receiver functions containing P-to-S converted waves at mantle discontinuities reveal a thinner MTZ thickness beneath Iceland, which was suggested as evidence for the lower mantle origin of the Iceland hotspot (Shen et al., 1998). A thinner MTZ is also observed beneath Hawaii (Li et al., 2000; Shen et al., 2003; Agius et al., 2018), the Galapagos hotspot (Hooft et al., 2003), and some but not all other hotspots (e.g., Owens et al., 2000; Li et al., 2003; Tauzin et al., 2008; Vinnik et al., 2012; Reed et al., 2016; Wei and Chen, 2016). In addition, several studies using underside SS reflections off the MTZ boundaries have also observed a thinned MTZ beneath some hotspots (e.g., Niu et al., 2002; Deuss, 2007; Cao et al.,

2011; Yu et al., 2017). Because of the weak signals of the converted waves from the MTZ discontinuities, most of the receiver function studies used a ray-theory-based common-conversion point (CCP) stacking method (Dueker and Sheehan, 1997; Shen et al., 1998) to suppress noise and enhance signals.

Receiver functions have also been used to image the MTZ boundaries in the Yellowstone region with contradicting results. Using pre-USArray seismic data in the region, Fee and Dueker (2004) observed a thinner transition zone beneath Yellowstone, but since the discontinuity topographies are spatially uncorrelated, they concluded that their observation provided no evidence for a lower mantle plume currently beneath the hotspot. With data from the USArray plus long-term networks and other temporary arrays, Schmandt et al. (2012) showed a normal 410-km discontinuity and an anomalously shallow 660-km discontinuity that is uplifted by up to 18 km over an area of about 200 km wide. Hier-Majumder and Tauzin (2017) suggested a much broader (500-600 km in diameter) thinner transition zone beneath Yellowstone. Gao and Liu (2014), however, found no mantle transition zone anomaly beneath the Yellowstone hotspot. The latter three studies shared much of the same data in the Yellowstone region, so the different results are likely attributable to differences in data selection, processing, and the method of CCP stacking. The width of receiver function binning in the CCP method, for example, is a subjective choice that may affect the observation of the transition zone anomaly. Furthermore, the complex finite-frequency wave phenomena such as multi-pathing and scattering, which may be significant in the Yellowstone region because of its

highly heterogeneous three-dimensional velocity variations (e.g., Porritt et al., 2014; Schmandt and Lin, 2014), are not accounted for in previous receiver function studies.

By fully accounting for finite-frequency wave propagation in a 3D heterogeneous structure, our PS-RTM method holds the promise of improved accuracy and resolution of the MTZ discontinuities. The results will help address the question of the origin of the Yellowstone hotspot. Furthermore, details of the MTZ discontinuities provide important constraints on mantle dynamics if the Yellowstone hotspot is fed by a mantle plume. For example, the lateral offset between the MTZ anomaly and the surface hotspot reflects the relative motion between the surface plate and the mantle driven by slab subduction and global mantle density distributions (Steinberger, 2000; Shen et al., 2002; Smith et al., 2009), and the magnitude and lateral dimension of the MTZ discontinuity topography and thickness anomaly reflect temperature and heat flux in the transition zone, which may be related to the plume temperature and heat flux at the core-mantle boundary (Zhong, 2006; Lin and van Keken, 2006; Leng and Zhong, 2008).

In this chapter, we first discuss the data processing procedures for selecting quality earthquakes among 4114 teleseismic earthquakes occurred in 5 years based on the signal-to-noise ratio (SNR), cross-correlation coefficient of recorded seismograms and station distribution. The iterative deconvolution method (Ligorria and Ammon,

1999) and principle component analysis (Rondenay et al. 2005; Shang 2014) are used to remove the source and instrument effects from seismic signals for each event and suppress incoherent noise. We then use the cubic spline interpolation discussed in Chapter 2 to interpolate the sparsely and unevenly recorded seismic signal onto evenly distributed numerical grids on the free surface. For each earthquake, an illumination coefficient based on the first Fresnel zone of each station is determined as the weight of each grid point. The RTM volumes of individual earthquakes are stacked to obtain the final RTM model, with the illumination coefficient as the weight in stacking to enhance the signals of mantle discontinuities. The imaged MTZ boundaries show a depressed 410-km discontinuity and an uplifted 660-km discontinuity in the Yellowstone region, which provide evidence supporting a hot plume conduit across the MTZ. We discuss the effects of different parameters used in the 3D PS-RTM, such as the earthquake distribution, reference velocity model, dominant wave frequencies and principle component analysis, on the RTM results.

3.2 Earthquake Data and Wavefield Regularization

To apply the 3D PS-RTM method (Li et al. 2018), it is necessary to select quality earthquake arrivals, remove the source and instrument effects on the waveforms, suppress uncorrelated noise, and interpolate receiver functions at unevenly and often sparsely distributed seismic stations to evenly distributed numerical grids.

3.2.1 Data Selection

We use broadband earthquake signals recorded by the Transportable Array (TA) of the USArray and archived at the Incorporated Research Institutions for Seismology (IRIS) Data Management Center. The stations used are within 38°N to 50°N, and -117°E to -105°E, a 12° by 12° area roughly centered on Yellowstone (Figure 3.1).

We requested earthquake records between May 2nd, 2005 and Sep 5th, 2010 with the moment magnitude (M_w) larger than 5.0 and the epicentral distance between 28° to 98°, based on the source information in the global CMT catalog (Dziewonski et al., 1981; Ekström et al., 2012) (Figure 3.2). The time window of the downloaded seismograms is between 100 s before and 150 s after the direct P arrival, while the theoretical direct P arrival time is calculated using the TauP Toolkit (Crotwell et al., 1999) with the one-dimensional reference earth model iasp91 (Kennett, 1991). We use the obspy toolbox (Beyreuther et al. 2010) to preprocess seismograms with the following steps:

- 1, Delete the blank traces;
- 2, Remove the instrument response and de-trend the seismic signal;
- 3, Discard three-component seismic traces with SNR lower than 3.0. In this study, the SNR is calculated using the vertical-component seismogram and is defined as the ratio between the maximum amplitude of the direct P arrival and the standard

deviation of the trace in the time window 10 to 100 seconds before the direct P arrival;

4, Calculate the maximum amplitude of each seismogram and compare it to the median of the maximum amplitudes of all stations for each component and for each earthquake, and then remove traces with abnormally large (larger than 2 times of the median) or small maximum amplitude (smaller than $1/10^{\text{th}}$ of the median);

5, Pick the direct P arrival with the classic STA/LTA method (Withers et al., 1998) and then use the multi-channel cross correlation method (VanDecar and Crosson, 1990) to further align the direct P arrival on the vertical-component seismograms;

6, Delete traces with the P-arrival cross correlation coefficients lower than 0.5 and skip earthquakes recorded by less than 10 stations remained after the screening in the previous steps;

7, Filter the traces using a band-pass filter with a frequency band between 0.02 Hz to 0.1 Hz;

8, Trim the seismic traces using a time window of 50 s before and 100 s after the direct P arrival and taper the traces at both ends with a 5-s time window.

3.2.2 Deconvolution

In traditional receiver function studies, the receiver functions are calculated by deconvolving the vertical-component seismogram from the radial- and transverse-component seismograms for each station (e.g., Langston, 1979) and in most studies, the transverse-component receiver functions are not calculated and used. In contrast, we deconvolve an effective source time function, $s(t)$, from all three components. We obtain the effective source time function for an earthquake by aligning and stacking vertical-component seismograms recorded by all seismic stations (Figure 3.3). Thus our effective source time function is different from the source time function of earthquake rupture. It defines the component of the waveform common to the entire array, which is dominated by processes far from the array including the waveform effects of the source, source-side surface reflections and scattering and propagation along the path.

We use a cross-correlation-based iterative deconvolution method (Ligorria and Ammon, 1999) to calculate three-component receiver functions. The procedure can be summarized briefly as following:

1. Finding the largest spike in the data series using cross correlation of the effective source time function and data signal in each iteration (Equations 3.1 and 3.2, where $a(t)$ is the cross correlation between the source time function, $s(t)$, and remaining data time series, $r_{n-1}(t)$, in iteration n).

2. Subtracting the convolution of the largest spike, a_k , and the source time function, $s(t)$, from the data (Equation 3.3), and adding the spikes to the current receiver function, $rf_{n-1}(t)$ (Equation 3.4).
3. If the remaining energy ratio, ε (the L2-norm of the remaining signal divide by the L2-norm of the data Equation 3.5), and the convergence rate, σ (Equation 3.6), are higher than critical values, and the iteration number is less than a maximum value, the calculation goes to the next iteration (Equation 3.1). Otherwise it exits the iteration and filters the receiver function with a Gaussian filter (Equation 3.7). We calculate the parameter α using the signal width T_w , which we choose as the half-length of the dominant period (5 s for a 10-50 s period band) of the data, as suggested by Oldenburg (1981).

$$a(t) = s(t) \star r_{n-1}(t), \quad (3.1)$$

$$a_k = \max(|a(t)|), \text{ where } a_k = a(k\Delta t), \quad (3.2)$$

$$r_n(t) = r_{n-1}(t) - a_k s(t) \star \delta(t - t_0 - k\Delta t), \quad (3.3)$$

$$rf_n(t) = rf_{n-1}(t) + a_k \delta(t - t_0 - k\Delta t), \quad (3.4)$$

$$\varepsilon = \frac{\|r_n(t)\|}{\|r_0(t)\|}, \quad (3.5)$$

$$\sigma = \frac{\|r_n(t)\| - \|r_{n-1}(t)\|}{\|r_0(t)\|}, \quad (3.6)$$

$$rf(t) = rf_n(t) \star \sqrt{\frac{\alpha}{\pi}} e^{-\alpha t^2}, \text{ where } \alpha = \left(1.66/T_w\right)^2. \quad (3.7)$$

In this study, the critical values for ε (Equation 3.5) and σ (Equation 3.6) are 0.05 and 0.0005, respectively, and the maximum iteration number is 500. Figure 3.3 shows an example of seismograms before and after deconvolution for an earthquake on July 16th, 2007 near the west coast of Honshu at depth of 15.1 km and with a moment magnitude of 5.74.

3.2.3 Principle Component Analysis

To suppress uncorrelated noise, we use a principle component analysis (PCA) method (Rondenay et al. 2005; Shang 2014), with the premise that the higher-order PCA components (the first, second, etc) account for correlated and partially correlated signals among the stations while the lower-order components correspond to random noise. The deconvolved seismograms are rotated from the north and east directions to the radial and transverse directions for each station (Figure 3.4a) of each earthquake using the back-azimuth angle calculated with the locations of the earthquake and the seismometer. Then we apply the PCA method to the vertical-, radial- and transverse-component receiver functions independently. We sort the singular values for all three components and calculate their sum as the total energy of the data. Then we keep the first few principle components corresponding to the singular values whose sum is larger than 95% of the total energy (Figure 3.4b). In the earthquake example shown in Figure 3.3, the 1st PCA component of the vertical and radial component contains more than 80% of the energy (Figure 3.4d) and has coherent converted wave signals on the radial

component (Figure 3.4c). On the transverse component, the 1st PCA component is less dominant, reflecting less coherent signals and random noise among the stations. Overall, it takes only the first few PCA components to reach the 95% of the energy on the vertical, radial and transverse components.

3.2.4 Interpolation

Following the procedures in Chapter 2, we then use the cubic spline interpolation method (Zhang and Zheng, 2015) to interpolate the three-component receiver functions from sparse stations to evenly distributed numerical grids on the surface.

We first interpolate the direct P arrival times at stations to construct an arrival time map as a function of latitude and longitude using cubic spline (Figure 3.5a), then interpolate the aligned three-component receiver functions at each time step, and shift the interpolated receiver functions back to the calculated arrival time for each surface grid using the arrival time map (Equation 2.1 to 2.4). In this procedure, the signals that have time-distance move-out curvatures similar to that of the direct arrival are well interpolated at numerical grids on the surface. Because of the uneven distribution of seismic stations, the wavefield sampling may be spatially aliased in regions that have sparser seismic stations, especially for waves with shorter periods (Figure 3.5b and c).

3.3 Reference Velocity Model, Back-Propagation and Stacking

We use the DNA13 velocity perturbation model (Porritt et al., 2014) added to the one-dimensional velocity model PREM (Dziewonski and Anderson, 1981) as our reference velocity model to back-propagate the signals and image the MTZ boundaries in the Yellowstone region. The main feature of this reference model in the Yellowstone hotspot region is a low velocity anomaly at the northeast end of the Snake River Plain throughout the upper 200km and a low velocity anomaly dipping from Yellowstone towards northwest at the transition zone depth, which may cause complex wave phenomena that are not accounted for in previous receiver function studies.

We use the non-staggered-grid finite-difference method of Zhang et al. (2012) to simulate wave propagation in a model that is 38°N to 50°N in the latitudinal direction, -117°E to -105°E in the longitudinal direction and 1000 km in depth. To minimize numerical dispersion, the grid spacing should be less than $1/7^{\text{th}}$ of the shear-wave wavelength. Based on the lowest wave period (10 s) and the reference velocity model we use, the shear-wave wavelength ranges from 32 km near the surface to 63.8 km at the bottom of our model. So we choose the horizontal grid spacing to be about 0.036 degree (about 4 km) in both the latitudinal and longitudinal directions, and the vertical grid spacing to be 3.65 km near the surface and gradually increasing to about 7.31 km at the bottom of the model.

For each selected earthquake, we back-propagate the P arrival and its coda from the free surface separately, use modified Helmholtz decomposition to further separate P and S waves in the image domain and apply a zero offset cross-correlation type image condition to calculate the correlation between P and S waves to image the mantle transition zone boundaries as described in Li et al. (2018).

Due to the distribution of seismic stations recording one specific earthquake and direction of wave propagation, this earthquake can illuminate specific part of the model. To combine the images created with different earthquakes, we develop a weighted stacking scheme, in which the weight is based on the first Fresnel zone of the P arrival with the dominant frequency for each earthquake and seismic station pair.

For each earthquake e and seismic station s , we calculate the travel time $t_e(i, j, k)$ from the earthquake e to the grid (i, j, k) in the model, the travel time $t_s(i, j, k)$ from the grid (i, j, k) to the seismometer s and the travel time t_{es} from the earthquake e to the seismometer s . If the total travel time from earthquake to grid to seismometer is larger than the travel time from earthquake to seismometer t_{es} plus half of the dominant period T (for grids outside of the first Fresnel zone), we set the weight to zero, otherwise we set the weight to one (Equation 3.8). The travel time $t_e(i, j, k)$, $t_s(i, j, k)$ and t_{es} are calculated with the Taup toolbox (Crotwell et al., 1999). The weights of different stations for each event are combined through Equation 3.9 and then used to obtain the stacked image with Equation 3.10.

$$w_{es}(i, j, k) = \begin{cases} 0, & \text{if } t_e(i, j, k) + t_s(i, j, k) - t_{es} > 0.5T, \\ 1, & \text{if } t_e(i, j, k) + t_s(i, j, k) - t_{es} \leq 0.5T, \end{cases} \quad (3.8)$$

$$w_e(i, j, k) = \begin{cases} 0, & \text{if none of } w_{es}(i, j, k) = 1, \\ 1, & \text{if any } w_{es}(i, j, k) = 1. \end{cases} \quad (3.9)$$

$$I(i, j, k) = \frac{\sum_{e=1}^N I_e(i, j, k) \cdot w_e(i, j, k)}{\sum_{e=1}^N w_e(i, j, k)}. \quad (3.10)$$

3.5 RTM Results

We obtain the RTM model for the Yellowstone hotspot by stacking 192 RTM volumes from 192 different earthquake sources with equation 3.10. Figure 3.6 show 6 different vertical slices of the final RTM model along the latitudinal and longitudinal directions (AA' to FF' in Figure 3.1). Compared with the previous study by Wang and Pavlis 2016 (Figure 3.7), the PSRTM method shows more clear images of the 410 and 660 discontinuities. The 410-km discontinuity can be clearly identified as a continuous feature on all the profiles, except south of 41.5°N on profile BB', where the image appears noisy. The 660-km discontinuity can also be identified and traced continuously along most of the profiles, though it appears as a slightly weaker feature compared to the 410-km discontinuity. In addition to the 410- and 660-km discontinuities, there is a feature at 500-520 km depth, which may correspond to the 520-km discontinuity and the β to γ phase change (e.g., Shearer, 1996), and a feature at 200-260 km depth, which may be associated with reverberation from shallower mid-lithosphere discontinuity (e.g., Abt et al., 2010; Lekic and Fischer,

2014; Liu and Gao, 2018). Because of the ~ 70 km station spacing of the USArray, the crust and lithosphere structures are not illuminated continuously and by multiple stations, which are essential for suppressing random noise. Thus we refrain from speculating whether the feature at 200-260 km depth is a reverberation from a shallow depth or not. Herein we focus our discussion on the more prominent features in the mantle transition zone, the 410- and 660-km discontinuities.

To determine the MTZ boundaries, we first choose one depth profile beneath Yellowstone (at latitude=44.667, longitude=-111.667) as the reference profile and interpolate the profile using an even-spacing depth interval of 0.5 km. We pick the depth of the 410-km discontinuity at the maximum amplitude in the profile between 390 km to 430 km and pick the depth of the 660-km discontinuity at the maximum amplitude between 640 km to 680 km. Then cross correlation is used to pick the depths of the two discontinuities at other locations within the deep range of 390-430 km and 640-680 km, respectively. We use a bootstrap method (Efron and Gong, 1983) to estimate the discontinuity depth and uncertainty. The bootstrap process involves 50 realizations of the RTM model. For each realization, we randomly select with replacement 192 RTM volumes from the original set of 192 volumes generated from different earthquakes, stack the bootstrapped volumes with equation 3.10 and measure the discontinuity depths as described above. Then we calculate the mean and standard deviation of the discontinuity depths of the 50 bootstrapped and stacked RTM models.

The resulting MTZ topography map shows a depression in the 410-km discontinuity and an upward deflection of the 660-km discontinuity near Yellowstone (Figure 3.8). The uncertainty of the discontinuity depth measurement is relatively high south of 42°N and along the eastern edge of the model. Both the 410- and 660-km discontinuities appear shallower along the northern edge of the model near the Canadian border. However, the transition zone thickness there is not abnormal, indicating that the apparent shallow discontinuities there are most likely due to velocity variations above the 410-km discontinuity. It is possible that the apparent shallower discontinuities near the northern edge of the RTM model is caused by an underestimation of high velocities of the craton in the north due to a lack of station coverage in the tomographic inversion (Porritt et al., 2014). Another possible reason is lacking of illumination near the boundary of our model. Combining seismic stations in a larger region would improve the image result can provide more constrain on the depth of the discontinuities.

In general, the MTZ thickness is a more accurate measurement than the depth of the 410- and 660-km discontinuities (Bina and Helffrich, 2014), because of uncertainties in shallow (<400 km depth) velocity variations affect the traveltimes of the converted waves from the 410- and 660-km discontinuities and thus their apparent depths, but do not affect the transition zone thickness due to nearly identical paths of the converted waves above the 410-km discontinuity. The MTZ thickness around the Yellowstone region is $\sim 230 \pm 7$ km (Figure 3.8), about 20 km thinner than the averages of the contiguous United States (250 km) and the western

United States (249 km) (Gao et al., 2014). For reference, the magnitude of the MTZ thinning is comparable to that of Iceland (~ 20 km, Shen et al., 1998) and the Galapagos hotspot (18 ± 8 km, Hooft et al., 2003). The uncertainty of the MTZ thickness from the bootstrap process is less than 7 km in the vicinity of the MTZ thickness anomaly, indicating that the 20 km thinning of the MTZ near Yellowstone is a robust observation. The thinning occurs over a region of approximately 200 km by 300 km, elongated in the east-west direction and shifted ~ 100 km north of the Yellowstone Park as outline by the 235-km contour.

Beneath the southwest corner of our model, previous studies (e.g., Schmandt et al. 2012) have shown the presence of subducted slabs, which have lower temperatures than the surrounding mantle. We found that both the 410 and 660 discontinuities in this region are depressed (Figure 3.9), with the 410 discontinuity depressed by about 20km and the 660 discontinuity depressed by about 30km. A possible explanation of the depressed 410 discontinuity is that highly depleted mantle with high magnesium composition, for which the phase transition from olivine to wadsleyite requires a relative high pressure (Frost 2003), is entrained by the cold down going slab. Meanwhile, the depressed 660 discontinuity is mainly caused by low temperature and possibly high water content.

3.5 Discussions

3.5.1 Effects of Earthquake Distribution

The earthquakes used in this study are unevenly distributed around the world (Figure 3.2) and can be roughly sorted into four groups based on their back-azimuths: The first group (back-azimuths between 90° to 200° , 72 earthquakes) contains earthquakes mostly located along the west coast of Central and South America and some earthquakes from the mid ocean ridge; the second group (back-azimuths between 200° to 270° , 42 earthquakes) contains earthquakes mostly located north of New Zealand and east of Papua New Guinea; the third group (back-azimuths between 270° to 0° , 71 earthquakes) contains earthquakes mostly located along the Aleutian islands, Japan islands and Mariana islands; the fourth groups (back-azimuths between 0° to 90°) contains a few earthquakes located sparsely in Eurasia.

To test the effects of earthquake distribution on the RTM result, we stack the RTM volumes for earthquakes from the four different groups (Figure 3.10). Groups 1 and 3 yield a more clearer image of the transition zone discontinuity structure than groups 2 and 4 due to more earthquakes in the former to depress noise and artifacts. Because of the back-azimuth differences among the groups, the illuminated region for each group is different, determined by the back azimuth. For example, the southeast parts of the model are well illuminated by group 1, while the northwest parts of the model are better illuminated by group 3. We note that the 660-km discontinuity is less well illuminated than the 410-km discontinuity, possibly due to

a larger lateral distance between the station and the P-to-S conversion point at the 660-km discontinuity, a relatively lower sampling density per unit area on the 660-km discontinuity than on the 410-km discontinuity, and geometric spreading of back-propagated waves that results in lower amplitudes of P and converted phases at the 660-km discontinuity than at shallower depths. By expanding the station coverage to a larger area in future studies, these differences can be reduced in the interior of the model sufficiently far away from the side boundaries of the model.

3.5.2 Effects of the Reference Velocity Model

The reference velocity model DNA13 we used in this study incorporates teleseismic P wave, independent SH and SV waves, and surface waves from both teleseismic earthquakes and ambient noise to constrain the relative wave-speed from the crust down into the lower mantle (Porritt et al., 2014). To understanding the sensitivity of the RTM results to the reference velocity model, we carried out RTM imaging for two additional models: PREM (Dziewonski and Anderson, 1981) and PREM perturbed with twice the velocity variation of DNA13.

The 410-km discontinuity has a depression in a 200 km by 300 km area around Yellowstone for all three reference velocity models (Figure 3.8a, 3.11a and d). However due to the differences in the velocity, the PREM model yields the largest average 410-km discontinuity, while the model with twice the DNA13 anomalies yields the shallowest average 410-km discontinuity depth. This is especially

obvious near the northern edge of the RTM models. This is caused by the low velocity anomaly in the northwest of DNA13 model. For the 660-km discontinuity, a 300km by 300km uplifted area is observed for all the image results. Similarly with 410km discontinuity, the more perturbation of DNA13 we add the shallower of the mean depth of 660km discontinuity (Figure 8b, 11b and 11e). The 660-km discontinuity depth is more sensitive of reference model than 410-km discontinuity. On the other hand, the MTZ thickness is less sensitive to the reference model (Figure 8c, 11c and 11f). The MTZ thickness around the Yellowstone region varies within 10km using different reference models.

These tests show that the depths of the discontinuities are sensitive to the reference model used during the migration. Extra attention should be paid to the reference model to analysis the migration result and multiple velocity models should be used to test the image result if possible.

3.5.3 Effects of the Dominant Frequency

The dominant frequency band may affect the RTM result in term of image resolution and reliability. In this section, we compare the results using waves in two different frequency bands: 4-50 s period and 10-50 s period (hereinafter shorter- and longer-period data, respectively). We use the same data selection and wave field regularization method described in sections 3.2 and 3.3 except the frequency band used for filtering the data. We also cut the model to a maximum depth of 765 km to

reduce the computational cost.

The RTM results along the latitudinal (BB') and longitudinal (EE') directions created with PCA components containing 95% of energy are shown in Figure 3.12. Because the time windows of P and P coda waves in the data domain are chosen based on the dominant frequency (1 period after the direct P wave, Li et al. 2018). The P-wave coda separated from P using the shorter (4 s) time length contains more signals generated in the lithosphere. Compared to the images obtained with the longer-period waves (Figure 3.6), the shorter-period data and smaller grid spacing yield sharper and more detailed images. However, they also contain significant noise and artificial boundaries in the image that have shapes similar to the Fresnel zones beneath different stations. This is because the USArray with a station spacing of ~ 70 km has significant spatial aliasing in sampling for short-period waves (Figure 3.5). Furthermore, the Fresnel zones beneath the stations do not overlap sufficiently, particularly at shallow depths, so noise is not sufficiently suppressed and is reflected in the image result. In contrast, the RTM result using the longer-period data show smoother and clear result for the MTZ boundaries due to longer apparent wavelengths on the surface and less aliased wavefield interpretation.

Another practical matter of using the longer-period data is a much lower computational cost. To avoid numerical dispersion and instability in finite difference simulation, the grid spacing Δx cannot exceed $1/7^{\text{th}}$ of the minimum wavelength and the time step Δt cannot exceed $\Delta x/c$, where c is the apparent velocity at the

numerical grid (Zhang et al. 2012). For this reason, to use the 4-s period data, the total grid number is about 8 times more and the time step is about 2 time more than using the 10-s period data. With the same computational resources, the shorter-period results take about 16 times more cpu hours to obtain.

3.5.4 Effects of PCA

We use PCA regularization to reduce incoherent noise across the seismic array, however the regularization may also suppress signals that are incoherent across the array. To understand the effect of PCA regularization, we retain only the first principle component for both the vertical and radial components and none for the transverse component. Compared with the receiver functions that contain 95% energy, the seismograms of the first principle component contain only coherent signals. Consequently, the interpolation would interpolate the coherent signal across the seismic array. We then use the first PCA component constructed waves to image the MTZ (Figure 3.13). Signals generated at relatively flat discontinuities are enhanced, while those from dipping structures or scatters are smoothed out. The RTM results obtained with fewer principle components have relative smoother topography for both the shorter- and longer-period data.

3.6. Conclusions

In this chapter, we apply the 3D PS-RTM method with USArray data to image the MTZ boundaries in the Yellowstone region. The results demonstrate the ability of the PS-RTM method in imaging the mantle discontinuities. With growing deployment of dense seismic networks, there is an increasing opportunity to use the method to image complex and detailed crustal and mantle structures such as Moho topography, sutures, and subducting slabs. We caution that although the 3D PS-RTM method is a powerful technique compared to traditional receiver function imaging methods, care must be taken in data regularization procedures such as data selection, deconvolution and PCA analysis.

192 earthquakes out of 4114 earthquakes are selected for imaging the MTZ boundaries. Most earthquakes we selected are from the subduction zones, where there are deep sources with relative clean first P arrival. Then an iterative deconvolution method is used to remove the source effect on the three-component seismograms. Different from traditional receiver function studies, which often use only the radial component, we use all three components (radial, transverse and vertical) receiver functions.

As in the active-source RTM method used in exploration, the 3D PS-RTM requires a relatively dense seismic network to recover the wavefield on the surface. The distribution of seismic stations is always limited by costs and logistic difficulties and the recorded signals are always contaminated by natural or instrument noise. To minimize uncorrelated noise, we develop a data regularization procedure using the

PCA method, and interpolate the sparsely recorded signals onto evenly distributed numerical grids on the earth surface using cubic spline interpolation. However, when choosing the components of PCA, there is a tradeoff between reserving the signal and removing the noise. Using fewer PCA components would remove more uncorrelated noise, however it would also remove the uncorrelated signal that may be caused by scattering.

The earthquakes used to image the earth structure are located mostly at subduction zones and because of the uneven distribution of the sources, the RTM result in this study does not have a good illumination in the northeast part of the model.

Furthermore, since the lateral distance between the station and conversion point is larger for the deeper discontinuity, the 410 discontinuity is better illuminated than the 660 discontinuity. The 3D PS-RTM method is also sensitive to the reference velocity model used for migration. In this study, the RTM results using 1D and 3D reference models both show a depressed 410-km discontinuity and an uplifted 660-km discontinuity, though the average depths of the discontinuities changed when using different models.

In theory, higher frequency waves would yield more detailed images, though the maximum frequency suitable in the PS-RTM depends on the station spacing and the target depth of discontinuities. As the frequency increases, the apparent horizontal wavelength on the surface decreases. When the station spacing is greater than half of the apparent horizontal wavelength on the surface, spatial aliasing occurs,

causing artifacts in the interpolated wavefield. To image the MTZ with the USArray data, the (10-s) longer-period waves are a more appropriate choice than the shorter- (4 s) period waves.

3.7 Acknowledgement

Data collected by the TA network were made freely available as part of the EarthScope USArray facility, operated by Incorporated Research Institutions for Seismology (IRIS) and supported by the National Science Foundation, under Cooperative Agreements EAR-1261681. This research was supported by the AFRL contracts FA9453-13C-0362, FA9453-16C-0019 and FA9453-17C-0026.

Bibliography:

Abt, D.L., Fischer, K.M., French, S.W., Ford, H.A., Yuan, H., Romanowicz, B., 2010.

North American lithospheric discontinuity structure imaged by P_s and S_p receiver functions. *J. Geophys. Res.* 115 (B9). <https://doi.org/10.1029/2009JB006914>.

Agius, M.R., C.A. Rychert, N. Harmon, G. Laske (2018), Mapping the mantle transition zone beneath Hawaii from P_s receiver functions: Evidence for a hot plume and cold mantle downwellings, *Earth Planet. Sci. Lett.*, 474, 226-236, <http://dx.doi.org/10.1016/j.epsl.2017.06.033>

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010), ObsPy: A Python Toolbox for Seismology, *Seismological Research Letters*, 81(3), 530-533, doi:10.1785/gssrl.81.3.530.

Bina, C.R., and G. Helffrich (1994), Phase transition Clapeyron slopes and transition zone seismic discontinuity topography, *J. Geophys. Res.*, 99 (B8), 15,853-15860.

Bina, C.R., and G. Helffrich (2014), Geophysical constraints on mantle composition, in Holland H.D. and Turekian K.K. (eds), *Treaties on Geochemistry*, Second Edition, 3, 41-65. Oxford: Elsevier.

Cao, Q., R.D. van der Hilst, M.V. de Hoop, S.-H. Shim (2011), Seismic imaging of transition zone discontinuities suggests hot mantle west of Hawaii, *Science*, 332, 1068-1071.

- Christiansen, R.L., G.R. Foulger, and J.R. Evans (2002), Upper-mantle origin of the Yellowstone hotspot, *Geol. Soc. Am. Bull.*, 114, 1245-1256.
- Crotwell, H. P., and T. J. Owens (1999), The TauP Toolkit: Flexible seismic travel-time and ray-path utilities, *Seismological Research ...*, doi:10.1785/gssrl.70.2.154.
- DeMets, C., R.G. Gordon, D.F. Argus (2010), Geologically current plate motions, *Geophys. J. Int.*, 181, 1-80. <https://doi.org/10.1111/j.1365-246X.2009.04491.x>.
- Deuss, A. (2007), Seismic observations of transition-zone discontinuities beneath hotspot locations, in *Plates, Plumes and Planetary Processes*, G.R. Folger, D.M. Jurdy, eds., *GEO*[https://doi.org/10.1130/2007.2430\(07\)](https://doi.org/10.1130/2007.2430(07)).
- Dueker, K.G., and A.F. Sheehan (1997), Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track, *J. Geophys. Res.*, 102 (B4), 8313-8327.
- Dziewonski, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Physics of the Earth and Planetary Interiors*, 25(4), 297-356, doi:[https://doi.org/10.1016/0031-9201\(81\)90046-7](https://doi.org/10.1016/0031-9201(81)90046-7).
- Dziewonski, A. M., T. A. Chou, and J. H. Woodhouse (1981), Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *Journal of Geophysical ...*, doi:10.1029/JB086iB04p02825.
- Efron, B., and G. Gong (1983), A leisurely look at the bootstrap, the jackknife, and cross-validation, *The Am Statistician*, 37(1), 36-48.
- Ekström, G., M. Nettles, and A. M. Dziewoński (2012), The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Physics of the Earth and*

- Planetary Interiors*, 200-201, 1-9,
doi:<https://doi.org/10.1016/j.pepi.2012.04.002>.
- Fee, D., and K. Dueker (2004), Mantle transition zone topography and structure of the Yellowstone hotspot, *Geophys. Res. Lett.*, 31, L18603:
doi:10.1029/2004GL020636.
- Foulger, G.R., R.L. Christiansen, D.L. Anderson (2015), The Yellowstone “hot spot” track results from migrating basin-range extension, in Foulger, G.R., Lustrino, M., and King, S.D., eds., *Geological Soc. Am. Special Paper 514 and AGU Special Publication 71*, 215-238, doi:10.1130/2015.2514(14).
- Frost, D. J. (2003), The structure and sharpness of (Mg,Fe)₂SiO₄ phase transformations in the transition zone, *Earth and Planetary Science Letters*, 216(3), 313-328, doi:[https://doi.org/10.1016/S0012-821X\(03\)00533-8](https://doi.org/10.1016/S0012-821X(03)00533-8).
- Gao, S. S., and K. H. Liu (2014), Mantle transition zone discontinuities beneath the contiguous United States, *Journal of Geophysical Research: Solid Earth*, 119(8), 6452-6468, doi:10.1002/2014JB011253.
- Hier-Majumder, S., and B. Tazsin (2017), Pervasive upper mantle melting beneath the western U.S., *Earth Planet. Sci. Lett.*, 464, 25-35, doi:10.1016/j.epsl.2016.12.041.
- Hirose, K. (2002), Phase transitions in pyrolytic mantle around 670-km depth: Implications for upwelling of plumes from the lower mantle, *J. Geophys. Res.*, 107, B4, 2078, <https://doi.org/10.1029/2001JB000597>.
- Hooft, E.E.E., D.R. Toomey, S.C. Solomon (2003), Anomalously thin transition zone beneath the Galapagos hotspot, *Earth Planet. Sci. Lett.*, 216, 55-64.

- Humphreys, E.D., K.G. Dueker, D.L. Schutt, and R.B. Smith (2000), Beneath Yellowstone: Evaluating plume and nonplume models using teleseismic images of the upper mantle, *GSA Today*, 10(12), 1-7.
- Ishii, T., H. Kojitani, M. Akaogi (2018), Phase relations and mineral chemistry in pyrolytic mantle at 1600-2200 C under pressures up to the uppermost lower mantle: Phase transitions around the 660-km discontinuity and dynamics of upwelling hot plumes, *Phys. Earth Planet. Inter.*, 274, 127-137, <https://doi.org/10.1016/j.pepi.2017.10.005>.
- James, D.E., M.J. Fouch, R.W. Carlson, J.B. Roth (2011), Slab fragmentation, edge flow and the origin of the Yellowstone hotspot track, *Earth Planet. Sci. Lett.*, 311, 124-135, doi: 10.1016/j.epsl.2011.09.007.
- Kennett, B. L. N. (1991), IASPEI 1991 SEISMOLOGICAL TABLES, *Terra Nova*, 3(2), 122-122, doi:10.1111/j.1365-3121.1991.tb00863.x.
- Kincaid, C., K.A. Druken, R.W. Griffiths, and D.R. Stegman (2013), Bifurcation of the Yellowstone plume driven by subduction-induced mantle flow, *Nat. Geoscience*, 6, 395-399, doi: 10.1038/ngeo01774.
- Langston, C.A. (1979), Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *J. Geophys. Res.*, 84 (B9), 4749-4762.
- Lekic', V., Fischer, K.M., 2014. Contrasting lithospheric signatures across the western United States revealed by *S_p* receiver functions. *Earth Planet. Sci. Lett.* 402, 90–98. <https://doi.org/10.1016/j.epsl.2013.11.026>.

- Leng, W., and S. Zhong (2008), Controls on plume heat flux and plume excess temperature, *J. Geophys. Res.*, 113, B04408, <https://doi.org/10.1029/2007JB005155>.
- Leonard, T., and L. Liu (2016), The role of a mantle plume in the formation of Yellowstone volcanism, *Geophys. Res. Lett.*, 43, 1132-1139, doi:10.1002/2015GL067131.
- Li, J., Y. Shen, and W. Zhang (2018), Three-Dimensional Passive-Source Reverse-Time Migration of Converted Waves: The Method, *Journal of Geophysical Research: Solid Earth*, 123(2), 1419-1434, doi:doi:10.1002/2017JB014817.
- Li, X.H., R. Kind, K. Priestley, S.V. Sobolev, F. Tilmann, X. Yuan, and M. Weber (2000), Mapping the Hawaiian plume conduit with converted seismic waves. *Nature*, 405, 938-941. <http://dx.doi.org/10.1038/35016054>.
- Li, X., R. Kind, X. Yuan (2003), Seismic study of upper mantle and transition zone beneath hotspots, *Phys. Earth Planet. Inter.*, 136, 79-92, doi: 10.1016/S0031-9201(03)00021-9.
- Ligorria, J. P., and C. J. Ammon (1999), Iterative deconvolution and receiver-function estimation, *Bulletin of the Seismological Society of America*, 89(5), 1395-1400.
- Lin, S.-C., and P.E. van Keken (2006), Dynamics of thermochemical plumes: 1. Plume formation and entrainment of a dense layer, *Geochem. Geophys. Geosys.*, 7, Q02006, <https://doi.org/10.1029/2005GC001072>.
- Liu, L., and S.S. Gao (2018), Lithospheric layering beneath the contiguous United States constrained by S-to-P receiver functions, *Earth Planet. Sci. Lett.*, 495, 79-86, <https://doi.org/10.1016/j.epsl.2018.05.012>.

- Morgan, W.J. (1971), Convection plumes in the lower mantle, *Nature*, 230, 42-43.
<https://doi.org/10.1038/230042a0>.
- Nelson, P.L., and S.P. Grand (2018), Lower-mantle plume beneath the Yellowstone revealed by core waves, *Nature Geoscience*, 11, 280-284, doi:10.1038/s41561-018-0075-y.
- Niu, F., S.C. Solomon, P.G. Silver, D. Suetsugu, H. Inoue (2002), Mantle transition zone structure beneath the South Pacific Superswell and evidence for a mantle plume underlying the Society hotspot, *Earth Planet. Sci. Lett.*, 198, 371-380.
- Oldenburg, D. W. (1981), A comprehensive solution to the linear deconvolution problem, *Geophysical Journal of the Royal Astronomical Society*, 65(2), 331-357, doi:doi:10.1111/j.1365-246X.1981.tb02716.x.
- Olson, P., P.G. Silver, and R.W. Carlson (1990), The large-scale structure of convection in the Earth's mantle, *Nature*, 344, 209-215.
<https://doi.org/10.1038/344209a0>.
- Owens, T.J., A.A. Nyblade, H. Gurrola, C.A. Langston (2000), Mantle transition zone structure beneath Tanzania, East Africa, *Geophys. Res. Lett.*, 27(6), 827-830.
- Poppeliers, C., and G. L. Pavlis (2003), Three - dimensional, prestack, plane wave migration of teleseismic P - to - S converted phases: 1. Theory, *Journal of Geophysical Research: Solid Earth*, 108(B2).
- Porritt, R. W., R. M. Allen, and F. F. Pollitz (2014), Seismic imaging east of the Rocky Mountains with USArray, *Earth and Planetary Science Letters*, 402, 16-25, doi:<https://doi.org/10.1016/j.epsl.2013.10.034>.

- Reed, C.A., S.S. Gao, K.H. Liu, and Y. Yu (2016), The mantle transition zone beneath the Afar depression and adjacent regions: implications for mantle plumes and hydration, *Geophys. J. Int.*, 205, 1756-1766, doi: 10.1093/gji/ggw116.
- Rondenay, S., M. G. Bostock, and K. M. Fischer (2013), Multichannel Inversion of Scattered Teleseismic Body Waves: Practical Considerations and Applicability, in *Seismic Earth: Array Analysis of Broadband Seismograms*, edited, doi:doi:10.1029/157GM12.
- Schmandt, B., K. Dueker, E. Humphreys, and S. Hansen (2012), Hot mantle upwelling across the 660 beneath Yellowstone, *Earth and Planetary Science Letters*, 331-332, 224-236, doi:https://doi.org/10.1016/j.epsl.2012.03.025.
- Schmandt, B., and F.-C. Lin (2014), P and S wave tomography of the mantle beneath the United States, *Geophys. Res. Lett.*, 41, 5342-6349, doi:10.1002/2014GL061231.
- Shearer, P. (1996), Transition zone velocity gradients and the 520-km discontinuity, *J. Geophys. Res.*, 101 (B2), 3053-3066.
- Shen, Y., S. C. Solomon, I. T. Bjarnason, and C. J. Wolfe (1998), Seismic evidence for a lower-mantle origin of the Iceland plume, *Nature*, 395(6697), 62-65, https://doi.org/10.1038/25714.
- Shen, Y., S.C. Solomon, I.T. Bjarnason, G. Nolet, W.J. Morgan, R.M. Allen, K. Vogfjord, S. Jakobsdottir, R. Stefansson, B.R. Julian, and G.R. Foulger (2002), Seismic evidence for a tilted mantle plume and north-south mantle flow beneath Iceland, *Earth Planet. Sci. Lett.*, 197, 261-272, https://doi.org/10.1016/S0012-821X(02)00494-6.

- Shen, Y., C.J. Wolfe, S.C. Solomon (2003), Seismological evidence for a mid-mantle discontinuity beneath Hawaii and Iceland, *Earth Planet. Sci. Lett.*, 214, 143-151, [http://dx.doi.org/10.1016/S0012-821X\(03\)00349-2](http://dx.doi.org/10.1016/S0012-821X(03)00349-2).
- Smith, R.B., M. Jordan, B. Steinberger, C.M. Puskas, J. Farrell, G.P. Waite, S. Husen, W.-L. Chang, and R. O'Connell (2009), Geodynamics of the Yellowstone hotspot and mantle plume: Seismic and GPS imaging, kinematics, and mantle flow, *J. Volcanology Geothermal Res.*, 188, 26-56, doi:10.1016/j.jvolgeores.2009.08.020.
- Steinberger, B. (2000), Plumes in a convecting mantle: Models and observations for individual hotspots, *J. Geophys. Res.*, 105, 11,127–11,152, <https://doi.org/10.1029/1999JB900398>.
- Tauzin, B., E. Debayle, G. Wittlinger (2008), The mantle transition zone as seen by global Pds phases: No clear evidence for a thin transition zone beneath hotspots, *J. Geophys. Res.*, 113, B08309, doi:10.1029/2007JB005364.
- VanDecar, J. C., and R. S. Crosson (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bulletin of the Seismological Society of America*, 80(1), 150-169.
- Vinnik, L., G. Silveira, S. Kiselev, V. Farra, M. Weber, E. Stutzmann (2012), Cape Verde hotspot from the upper crust to the top of the lower mantle, *Earth Planet. Sci. Lett.*, 319-320, 259-268, doi:10.1016/j.epsl.2011.12.017.
- Waite, G.P., R.B. Smith, and R.M. Allen (2006), Vp and Vs structure of the Yellowstone hot spot from teleseismic tomography: Evidence for an upper mantle plume, *J. Geophys. Res.*, 111, B04303, doi:10.1029/2005JB003867.

- Wang, Y., and G. L. Pavlis (2016), Roughness of the mantle transition zone discontinuities revealed by high-resolution wavefield imaging, *Journal of Geophysical Research: Solid Earth*, 121(9), 6757-6778, doi:10.1002/2016jb013205.
- Wei, S.S., and Y.J. Chen (2016), Seismic evidence of the Hainan mantle plume by receiver function analysis in southern China, *Geophys. Res. Lett.*, 43, 8978-8985, doi:10.1002/2016GL069513.
- Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo (1998), A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bulletin of the Seismological Society of America*, 88(1), 95-106.
- Yu, C., E.A. Day, M.V. de Hoop, M. Campillo, and R.D. van der Hilst (2017), Mapping mantle transition zone discontinuities beneath the Central Pacific with array processing of SS precursors, *J. Geophys. Res., Solid Earth*, 122, 10,364-10,378. <https://doi.org/10.1002/2017JB014327>.
- Yuan, H., and K.G. Dueker (2005), Teleseismic P-wave tomogram of the Yellowstone plume: *Geophys. Res. Lett.*, 32, L07304, doi:10.1029/2004GL022056.
- Zhang, W., Y. Shen, and L. Zhao (2012), Three-dimensional anisotropic seismic wave modelling in spherical coordinates by a collocated-grid finite-difference method, *Geophysical Journal International*, 188(3), 1359-1381, doi:10.1111/j.1365-246X.2011.05331.x.

- Zhong, S. (2006), Constraints on thermochemical convection of the mantle from plume heat flux, plume excess temperature, and upper mantle temperature, *J. Geophys. Res.*, 111, B04409, <https://doi.org/10.1029/2005JB003972>.
- Zhou, Q., L. Liu, and J. Hu (2018), Western US volcanism due to intruding oceanic mantle driven by ancient Farallon slabs, *Nature Geoscience*, 11, 70-76, doi:10.1038/s41561-017-0035-y.
- Zhou, Y. (2018), Anomalous mantle transition zone beneath the Yellowstone hotspot track, *Nat. Geo.*, 11, 449-453, doi:10.1038/s41561-018-0126-4.
- Zhu, L., and H. Kanamori (2000), Moho depth variation in southern California from teleseismic receiver functions, *Journal of Geophysical Research: Solid Earth*, 105(B2), 2969-2980, doi:doi:10.1029/1999JB900322.

Figure 3.1. The topography map of the study area. The red triangles represent the seismic stations. The dashed blue lines mark the locations of the vertical profiles in Figure 3.6. The dashed black line mark the location of the vertical profile along latitude 40N in Figure 3.9.

Figure 3.2. Distribution of the earthquakes used in the study. The blue circles represent the earthquakes we downloaded from IRIS and processed. The sizes of the circles represent the earthquake magnitude. The red, green, yellow and mauve circles represent the different azimuthal groups of earthquakes selected for PS-RTM imaging. The three concentric dashed circles centered on the Yellowstone mark the 30° , 60° and 90° epicentral distances, respectively.

Figure 3.3. Seismograms of an Mw 5.74 earthquake located on the west coast of Honshu on 2007 July 16th. (a) and (b) represent the three component seismograms before and after the deconvolution. The P arrival is arbitrarily placed at 50 s, while the arrival times of P410s and P660s phases are marked for reference. The traveltimes are calculated with the TauP Toolkit (Crotwell et al., 1999) using the iasp91 model (Kennett, 1991). (c) The color coded traces are aligned vertical component seismograms and the black trace is the stack of all the traces, which is used as the effective source time function.

Figure 3.4. (a) Deconvolved seismograms after the rotation to the transverse, radial and vertical components; (b) Seismograms after applying PCA to each component separately and using the first few components that contain over 95% of the total energy; (c) Seismograms after applying PCA and reserving only the first component for the radial and vertical components; (d) Energy distribution among the first 20 PCA components for transverse, radial and vertical seismograms. The first few PCA components that contain over 95% of the energy are colored in orange. Notice the dominant energy in the first PCA component of the vertical and radial seismograms.

Figure 3.5. (a) Relative arrival time map created by cubic spline interpolation of the relative arrival times at seismic stations; (b) A snapshot of wave amplitude of 10s dominant frequency wave on the free surface, warm color represents the amplitude of 10s direct P wave; (c) A snapshot of wave amplitude of 4s dominant frequency wave on the free surface, warm color represent the amplitude of 4s direct P wave.

Figure 3.6. Stacked RTM model obtained with 10-s period waves, the DNA13 velocity model and PCA components that contain 95% energy. (a), (b) and (c) represent three vertical slices of the RTM model along the latitudinal direction at longitude of -114.333° , -111.667° and -109° , respectively. (d), (e) and (f) represent three vertical slices of the RTM model along the longitudinal direction at latitude of 47.333° , 45.676° and 44.018° , respectively.

Figure 3.7. Image results from Wang and Pavlis (2016) (a), (b) and (c) represent three vertical slices of the their model along the latitudinal direction at longitude of -114.333°, -111.667° and -109°, respectively. (d), (e) and (f) represent three vertical slices of the model along the longitudinal direction at latitude of 47.333°, 45.676° and 44.018°, respectively.

Figure 3.8. (a) and (b) are topography map of the 410- and 660-km discontinuities respectively; (c) is the MTZ thickness, the white line mark the 235km depth contour; (d) and (e) are the uncertainty (one standard deviation) of the 410- and 660-km discontinuities respectively; (f) is the uncertainty (one standard deviation) of MTZ thickness.

Figure 3.9. Vertical slice along latitude 40N. The blue and red contour marks the +1% and -1% S-wave velocity perturbation in Schmandt and Lin (2014).

Figure 3.10. RTM results obtained by stacking different earthquake groups. (a) and (e) Vertical slices along the latitude and longitude directions of the RTM model of group 1. The locations of the BB' and EE' profiles can be found in Figure 3.1; (b) and (f) Results of group 2; (c) and (g) Results of group 3; (d) and (h) Results of group 4.

Figure 3.11. MTZ discontinuity topography map created using different velocity models. (a),(b) and (c) are topography map of the 410-, 660-km discontinuities and MTZ thickness using the 1D PREM model respectively; (c), (d) and (e) are topography map of the 410-, 660-km discontinuities and MTZ thickness using the 1D PREM plus twice the velocity perturbation in the DNA13 model respectively.

Figure 3.12. (a) and (b) Vertical slices along the latitude and longitude directions of the RTM model of 4-s period waves.

Figure 3.13. Image results obtained with only the 1st PCA component of the vertical and radial components and no transverse component. (a) and (b) Vertical slices along the latitudinal and longitudinal directions of the RTM model of longer-period waves; (c) and (d) Vertical slices along the latitude and longitude directions of the RTM model of shorter-period waves.