APPRAISING THE RELIABILITY OF SCATTERED WAVEFIELD IMAGING

Xin Liu

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Science
in the Department of Geological Sciences
Indiana University

Aug, 2011
Acknowledgement

As a graduate student here, the two-year experience of me is extraordinarily enjoyable and memorable. I’ve learned a lot both in earth sciences and allied sciences (math and computer sciences) from the professors and the students here.

Dr. Gary Pavlis is my advisor and is the first person I want to thank. Without him I could not have ever been here at IU. He is very nice and easygoing with students. He taught me scattered wavefield imaging, starting from the simplest point scatterer model. As my knowledge of seismology grew, he led me to the theoretical foundation of scattered wavefield imaging and helped me improving the point scatterer model by adding new features. He always has thousands of brilliant ideas and when I followed his ideas, I began to think like a scientist. With his help I also learned a lot in computing and programming. He not only educated me and brought to me the frontier of seismic imaging, but also offered me lots of opportunities to be exposed to the academic world and industry. He taught me that a scientist should be honest and never over-interpret the results, which greatly influenced my way of doing research.

I also want to thank Dr. Michael Hamburger and Dr. Kaj Johnson, who are my committee members. Dr Johnson revealed to me the secrets about mechanics and inverse problems and I learned the recondite maths without much difficulty. Dr. Hamburger gave me a grand view of the global tectonics, which is very important for me to understand the earth science questions when reading or writing a paper. I thank Dr. Bruce Douglas, who helped me whenever I had questions or didn’t catch his words on his class.

It’s my pleasure to acknowledge the RA funding from NSF and the Tudor Fellowship from Chevron and my department. The friendly atmosphere in our department helped me overcome many difficulties and made my work and study here a lot easier. I also thank my friends here who make my life so abundant.

Finally, I wish to express my gratitude to my parents and sister, Y.H. Liu, S.L. Guo and Sheng. They are my source of confidence and support at all times.
Abstract

Passive array wavefield imaging currently suffer from a lack of methods to appraise the reliability of the results. In this thesis I explore the use of point scattering models to produce a resolution test comparable to checkboard tests commonly used in seismic tomography. Passive array data differ strongly from seismic reflection data because the recording geometry is highly variable. As a result, I found it important for error appraisal to produce a simulated data set that exactly simulates the real data. For the point scattering simulation I use an asymptotic approximation based on the perturbation theory of single point scattering model. Point resolution is estimated by conducting numerical experiments with different point spacings. I also conducted tests with densely spaced point scatterers to approximate a subsurface and compared them with the simple layered simulation which is a complementary test. In addition, a probabilistic inversion which estimates the limits of a single station in resolving the subsurface discontinuity is performed. I apply this to appraise results from imaging the 410 discontinuity with USArray data. Our results suggest that the 410 discontinuity has scattering properties more like a rough surface than a mirror. I make this inference because the images have coherent structure at scales within the resolution limits revealed by the simulations. I cannot, however, rule out an alternative hypothesis that what I observe results from noise. Additional error analysis to objectively quantify noise levels in the migrated data is needed to resolve this issue.
# Table of Contents

Acknowledgement ........................................................................................................... iii

Abstract .......................................................................................................................... iv

1 Introduction .................................................................................................................... 1

2 A Forward Model based on Single Point Scattering Theory ...................................... 4

2.1 Single Point Scattering Model ................................................................................. 4

2.2 Algorithm and Implementation ................................................................................ 9

2.2.1 Polarization: Radial Component polarization and coordinate transformation. 10

2.2.2 Amplitude: Geometric Spreading and Scattering Pattern .............................. 13

2.2.3 Lag time in arrival time reference frame ......................................................... 16

2.2.4 Superposition of spike response functions and the choice of wavelet .......... 18

2.2.5 Parallel Algorithm .............................................................................................. 24

3 Resolution test and 410 discontinuity ....................................................................... 27

3.1 Widely spaced point scatterers ............................................................................... 28

3.2 Flat Subsurface Approximation Based on Point Scatterers or Simple Flat Layer Model ........................................................................................................................................ 35

3.2.1 Simple Layered Model ........................................................................................ 36

3.2.2 Superposition of densely spaced point scatterer responses based on Ricker wavelet 38

3.3 Simulating a Dipping Layer with Point Scatterers ............................................. 40

3.4 Checkboard test ...................................................................................................... 42

3.5 Application to 410 km discontinuities .................................................................... 44

4 Estimating the uncertainty of Scattered Wavefield Imaging by Trans-dimensional Monte Carlo Inversion ................................................................................................. 47

4.1 Forward model ........................................................................................................ 48

4.2 Inverse model and results ...................................................................................... 49

4.3 Discussion and conclusion ..................................................................................... 56

5 Conclusion and Suggestion on Future Research .................................................... 59

References ...................................................................................................................... 62

Curriculum Vitae

v
List of Figures

Figure 2-1 Polarization of P-to-S scattered wave. ................................................................. 9
Figure 2-2 Scattering pattern for P-to-S single point scattering. .............................................. 14
Figure 2-3 Geometric spreading in 3D spherical earth. .......................................................... 15
Figure 2-4 Iterative Newton’s method for computing travel time. .......................................... 18
Figure 2-5 2D case: densely spaced point scatterers to approximate a flat interface. ............... 20
Figure 2-6 2D case: superposition of point scatterer responses with Gaussian wavelet. ....... 21
Figure 2-7 2D case: superposition of point scatterer responses with Ricker wavelet. ........... 22
Figure 2-8 3D case: densely spaced point scatterers to approximate a flat interface. * is source, triangles are receivers and white spheres are point scatterers .................................... 22
Figure 2-9 3D case: superposition of point scatterer responses with Ricker wavelet. 
Scattering pattern is enabled in the forward program......................................................... 23
Figure 2-10 3D case: superposition of point scatterer responses with Ricker wavelet. 
Scattering pattern is disabled in the forward program......................................................... 24
Figure 3-1 Model geometry of five widely spaced point scatterers........................................ 29
Figure 3-2 Image of five widely spaced point scatterers using Gaussian pulse. ................. 29
Figure 3-3 Image of five widely spaced point scatterers using Ricker wavelet. ................. 30
Figure 3-4 Image slice across 3 point scatterers. ................................................................. 31
Figure 3-5 Image slice across 3 point scatterers. ................................................................. 32
Figure 3-6 Map view of point scatterer resolution test at 410 km depth. ............................... 33
Figure 3-7 Cross-section of point scatterer resolution test at 410 km depth. ....................... 34
Figure 3-8 Cross-section of the reconstructed image using four flat interfaces and a Gaussian function. .................................................................................................................. 36
Figure 3-9 Cross-section of the image of four major interfaces using a Ricker wavelet. .... 37
Figure 3-10 Cross-section of the image of 410 km discontinuity simulated by a densely spaced point scatterer grid. A Ricker wavelet is applied. ............................................. 38
Figure 3-11 Image of the simulated 410 km discontinuity using Ricker wavelet. ........ 39
Figure 3-12 Image of a dipping layer: vertical slice view. .................................................. 40
Figure 3-13 Image of a dipping layer: contour view......................................................... 41
Figure 3-14 Map view of checkboard test at 410 km depth. ........................................... 43
Figure 3-15 Vertical slice of checkboard test...................................................................... 44
Figure 3-16 Vertical slice of the amplitude variations at 410 km depth based on USArray
data................................................................................................................................. 45
Figure 3-17 Map view of the amplitude variations at 410 km depth based on USArray
data................................................................................................................................ 46
Figure 4-1 Illustration of the forward model. ................................................................. 48
Figure 4-2 Waveform fitting based on Monte Carlo method. Normally distributed noise is
added to the synthetic waveform data. ........................................................................... 50
Figure 4-3 The 2D model geometry for the synthetic data. ........................................... 51
Figure 4-4 Marginal probability distribution of \( V_s \) perturbation on 10 point scatterers. 54
Figure 4-5 Distribution of the model space dimension...................................................... 54
Figure 4-6 Inverted \( V_s \) perturbation based on Reversible Jump MCMC method........ 55
Figure 4-7 Standard deviation on each point scatterer in model space......................... 56
Figure 4-8 Shortest path and tangent point of isocron.................................................... 57
1 Introduction

Passive array scattered wavefield imaging methods used to date suffer from a lack of methods to appraise the reliability of the results. It is important to know the limitation of the imaging algorithm with currently available data. In this thesis, I approach this problem by developing a general way to model scattered waves and use this to simulate the irregular pattern that characterizes real data. The implications of the resolution tests for scattered wavefield imaging are discussed as well as hypotheses for appraising the reliability of the image based on the results.

Scattered wavefield imaging is a recently developed technology for imaging the deep interior of the earth. Based on the USArray data, 3D seismic imaging is capable of revealing geological features from lower crust to the mantle transition zone. 3D imaging is most useful in identifying and reconstructing discontinuity subsurfaces such as the Moho, the subducting slab and the 410 discontinuity. The heterogeneities along a subsurface might also be linked with the amplitude variations in the imaging volume.

This imaging technique I used is based on the single-point forward scattering model that is derived from perturbation theory and the Born approximation (Beylkin and Burridge 1990; Weglein et al., 2003; Bostock et al., 2001). Four possible scattering modes are P-to-P, P-to-S, S-to-P and S-to-S, each of which has a scattering pattern and a polarization term. P-to-S wavefield imaging has a broader application in seismological
studies than the other three scattering modes (Bostock et al., 2001; Poppeliers and Pavlis, 2003; Pavlis, 2011).

The 3D prestack migration imaging technique (Pavlis, 2011) is a P-to-S wavefield imaging method and the goal here is to estimate how well it can recover the predefined scatterers. The current imaging results are based on the Earthscope Automated Receiver Survey (EARS) data (Crotwell and Owens, 2005) and the synthetic data should share a common receiver geometry and frequency characteristics with the real data.

In scattered wavefield imaging, the forward problem is that the scattered wave propagating from scatterer (discontinuity) to receiver and the output are seismograms. The inverse problem is to image the subsurface discontinuities based on the seismograms by back-projecting scattered phases on the seismogram to imaging point.

The strategy we apply is to define a forward model geometry (receivers, hypocenter and point scatterers) and compute the synthetic data based on the geometry. Then the synthetic data are processed with the imaging program to see how well we recover the original point scatterer geometry. By comparing the image and the original point scatterer geometry, I developed some systematic steps to appraise the reliability of point or planar discontinuities recovered by the scattered wavefield imaging program.

Similar tests in reflection seismology have been implemented by approximating the diffraction pattern of an interface using Kirchoff integral for acoustic waves (Trorey 1970; Hilterman 1982). However, they assume only P-to-P backscattering (conventional
seismic reflection) and the angular dependant scattering pattern is ignored.

In chapter 2, I discussed the 3D scattering pattern, polarization and amplitude based on the single point scattering model and perturbation theory. I derive the fundamental equations used for the actual implementation. This chapter only deals with the forward problem.

In chapter 3, I apply the prestack migration imaging program to reconstruct the point scatterer geometry using the synthetic seismogram data. The discontinuity can be point scatterers with large spacing or densely spaced point scatterers. Horizontal discontinuity interface and dipping interface are approximated by densely spaced point scatterers. The amplitude variation and topography in 410 km discontinuity is discussed based on USArray data.

In chapter 4, the uncertainty in scattered wavefield imaging is estimated using Reversible-Jump Monte Carlo method. This experiment aims to investigate the ability of a single receiver in resolving subsurface point discontinuities. The results show that the scatterers in the small neighborhood of the tangent point to the isocron can not be distinguished by the waveform because they generate scattered wave arrivals at nearly the same time. A hypothesis based on this result and Fermat’s principle of least time is provided.
2 A Forward Model based on Single Point Scattering Theory

2.1 Single Point Scattering Model

Scattered wave field imaging is based on the single-point scattering model which is sometimes referred to as the Born approximation solution to the wave equation in perturbed media. The reference media is assumed to be isotropic with perturbation in a very small volume (not zero).

Beylkin and Burridge (1990) derived the 3D scattered wavefield Green’s function in time domain and obtained the scattering pattern. Weglein et al. (2003) and Bostock et al. (2001) solutions are restricted in 2D cases but they applied frequency domain methods. Weglein et al. adopted matrix arithmetic while Bostock et al. introduced complex functions, both of which are involved with the displacement of scattered wavefield in frequency domain. Here we mainly follow Beylkin and Burridge’s time-domain method and we also combine Weglein et al. and Bostock et al. notations to simplify the derivation and clarify key steps.

We assume we know a reference media that is isotropic and normally treated as relatively simple. We consider a very small volume $D$ in which the elastic parameters and density are perturbed. On the boundary $\partial D$ which bounds $D$ the perturbation values are assumed to be zero.
In the reference media, the wave field propagating from source \( s \) to scatterer \( x \) satisfies

\[
L^0_{lp} \tilde{G}^0_{jp} = \delta_{jl} \delta(t) \delta(x - s),
\]  

(2.1)

where \( \tilde{G}^0_{jp} \) is the Green’s function, and \( L^0_{lp} \) is a linear differential operator

\[
L^0_{lp} = \rho^0 \frac{\partial^2}{\partial t^2} \delta_{lp} - \frac{\partial}{\partial x_m} c^0_{mpq} \frac{\partial}{\partial x_q}
\]  

(2.2)

\( \delta_{jl} \) is the Kronecker delta function.

For isotropic reference media, the elastic parameters \( c^0_{mpq} \) satisfy

\[
c^0_{mpq} = \lambda^0 \delta_{lm} \delta_{pq} + \mu^0 (\delta_{lp} \delta_{mq} + \delta_{lq} \delta_{mp})
\]  

(2.3)

where \( \lambda^0 \) and \( \mu^0 \) are Lame constants in the reference media.

In the perturbed media, the wave equation becomes,

\[
L_{lp} \tilde{G}_{jp} = \delta_{jl} \delta(t) \delta(x - s)
\]  

(2.4)

where \( \tilde{G}_{jp} \) is the Green’s function, and \( L_{lp} \) is a linear differential operator

\[
L_{lp} = \rho \frac{\partial^2}{\partial t^2} \delta_{lp} - \frac{\partial}{\partial x_m} c_{mpq} \frac{\partial}{\partial x_q}
\]  

(2.5)

The elastic parameters \( c_{mpq} \) satisfy

\[
c_{mpq} = \lambda \delta_{lm} \delta_{pq} + \mu (\delta_{lp} \delta_{mq} + \delta_{lq} \delta_{mp})
\]  

(2.6)

where \( \lambda \) and \( \mu \) are Lame constants in the perturbed media.

We define the perturbation operator

\[
\Delta L_{lp} = L_{lp} - L^0_{lp} = \Delta \rho \frac{\partial^2}{\partial t^2} \delta_{lp} - \frac{\partial}{\partial x_m} \Delta c_{mpq} \frac{\partial}{\partial x_q}
\]  

(2.7)

where \( \Delta \rho = \rho - \rho^0 \), \( \Delta c_{mpq} = c_{mpq} - c^0_{mpq} \).

The scattered wave field equation can be derived from (2.1) and (2.4)
\[ L_{ip}^0 \Delta G_{jp} = - \Delta L_{ip} \bar{G}_{jp} \]  \hfill (2.8)

where \( \Delta G_{jp} = \bar{G}_{jp} - \bar{G}_{jp}^0 \) is the scattered wave field.

The right hand side of (2.8) can be regarded as an effective source for the scattered wave (Bostock et al., 2001; Weglein et al., 2003).

Let \( \bar{G}_{kl}^0 \) be the Green’s function of seismic wave propagating from point scatterer \( x \) to receiver \( r \) in the reference media. Then we can derive the Lipmann-Schwinger equation

\[ \Delta G_{jk} = \int_D \left( - \Delta L_{ip} \bar{G}_{jp} \right) *_t \bar{G}_{kl}^0 dx \]  \hfill (2.9)

where \( *_t \) denotes convolution.

The first order approximation, or Born Approximation to equation (2.7) is (Beylkin and Burridge 1990)

\[ \Delta G_{jk} = \int_D \left( - \Delta L_{ip} \bar{G}_{jp}^0 \right) *_t \bar{G}_{kl}^0 dx. \]  \hfill (2.10)

Here we only derive the P-to-S scattering because it’s the mode used in the forward and inverse models. Based on simple ray theory, the Green’s functions for incident P wave and scattered S wave are, respectively (Beylkin and Burridge 1990)

\( \bar{G}_{jl}^0 = \bar{G}_{jl}^P = \bar{A}_{jl}^P \delta (t - \bar{\phi}^P) \) \hfill (2.11)

\( \bar{G}_{kl}^0 = \bar{G}_{kl}^S = \bar{A}_{kl}^S \delta (t - \bar{\phi}^S) \), \hfill (2.12)

where \( \bar{\phi}^P \) is the P wave travel time from the source to the domain \( D \) (scatterer), and \( \bar{\phi}^S \) is the S wave travel time from the scatterer to the receiver.

Substitute \( \Delta L_{ip} \) (2.7) into (2.10), then integrate the second term by parts (Beylkin and Burridge 1990)
\[ \Delta G_{jk}^{PS} = - \int_D \left[ \Delta \rho \frac{\partial^2}{\partial t^2} \left( \tilde{G}_{ji}^p \right) \ast_t \tilde{G}_{kl}^S + \Delta c_{impq} \frac{\partial \tilde{g}_{ji}^p}{\partial x_q} \ast_t \frac{\partial \tilde{g}_{kl}^S}{\partial x_m} \right] dx \]  

(2.13)

Combining (2.11), (2.12) and (2.13) we have

\[ \Delta G_{jk}^{PS} = - \frac{\partial^2}{\partial t^2} \int_D \left[ \Delta \rho \delta_{lp} + \Delta c_{impq} \frac{\partial \tilde{g}_{ji}^p}{\partial x_q} \frac{\partial \tilde{g}_{kl}^S}{\partial x_m} \right] \tilde{A}_{jp}^{\delta} \tilde{A}_{kl}^{\delta} \delta(t - \tilde{\phi}^p - \tilde{\phi}^S) dx. \]  

(2.14)

The amplitude \( \tilde{A}_{jp}^{\delta} \) can be written as \( \tilde{A}_{jp}^{\delta} = \tilde{\alpha}_{jp}^p \tilde{A}_j^p \), where \( \tilde{\alpha}_{jp}^p \) is the direction of propagation of the incident P wave. Quasi-plane wave incidence is assumed due to a small point scatterer and the amplitude term is computed by ray theory.

Let \( \tilde{\alpha}_S^S \) be the unit vector in the direction of propagation of scattered S wave. To decompose the polarization of converted S wave, we define the basis in the transverse direction

\[ \gamma^{PS} = \frac{\tilde{\alpha}_S^p \times \tilde{\alpha}_S^S}{\| \tilde{\alpha}_S^p \times \tilde{\alpha}_S^S \|} \]  

(2.15)

and the basis in the radial direction

\[ \beta^{PS} = \frac{\tilde{\alpha}_S^S \times \gamma^{PS}}{\| \tilde{\alpha}_S^S \times \gamma^{PS} \|}. \]  

(2.16)

Therefore, \( \gamma^{PS}, \beta^{PS} \) and \( \tilde{\alpha}_S^S \) are three orthogonal basis vectors and we have

\[ \tilde{\alpha}_j^P \beta_j^{PS} = \sin \theta^{PS} \]  

(2.17)

\[ \tilde{\alpha}_j^P \tilde{\alpha}_j^S = \cos \theta^{PS} \]  

(2.18)

where \( \theta^{PS} \) is the P-to-S scattering angle which is defined as the angle between the incident P ray and the converted S ray.

Because the S wave polarization must be perpendicular to the direction of propagation \( \tilde{\alpha}_S^S \), it can be written as a linear combination of \( \gamma^{PS} \) and \( \beta^{PS} \). We have

\[ \tilde{A}_{kl}^S = \tilde{A}_{kl}^{PS} (\beta_j^{PS} a_j^{PS} + \gamma_i^{PS} a_i^{PS}). \]  

(2.19)
The scalar coefficients of converted S wave polarization \( a_1^{PS} \) and \( a_2^{PS} \) satisfy

\[
(a_1^{PS})^2 + (a_2^{PS})^2 = 1. \tag{2.20}
\]

Let \( a_1^{PS} = \cos \phi, \ a_2^{PS} = \sin \phi \), then (2.18) can be satisfied implicitly.

Substituting (2.19) (2.17) (2.18) and \( \vec{A}_j^P \) into (2.14), and knowing that \( \frac{\partial \phi_P}{\partial x_q} = \frac{\dot{\alpha}_q^P}{c_P} \)

\[
\text{and} \quad \frac{\partial \phi_S}{\partial x_m} = \frac{\dot{\alpha}_m^S}{c_S}, \quad \text{we have (Beylkin and Burridge 1990)}
\]

\[
\Delta G_{jk}^{PS} = -\frac{\partial^2}{\partial t^2} \int_D \rho_0 \left[ \frac{\Delta \rho}{\rho_0} \sin \theta_P^{PS} + \frac{\Delta \mu}{\mu_0 \ c_P} \sin 2\theta_P^{PS} \right] \vec{A}_j^P \vec{A}_k^{PS} \delta(t - \phi_P - \phi_S) \ dx. \tag{2.21}
\]

Beylkin and Burridge (1990) show that the 3D scattering pattern is

\[
f^{PS} (\theta^{PS}) = -\left[ \frac{\Delta \rho}{\rho_0} \sin \theta_P^{PS} + \frac{\Delta \mu}{\mu_0 \ c_P} \sin 2\theta_P^{PS} \right] \tag{2.22}
\]

where \( \mu_0 \) is the shear modulus and \( \rho_0 \) is the density in the reference media.

The minus sign is important and it goes into the scattering pattern term in the implementation of the forward scattering simulation program.

The scattered wave displacement is

\[
\Delta u_k^{PS} = \Delta G_{jk}^{PS} \ast_t S(t) \dot{\alpha}_j^P \tag{2.23}
\]

where \( S(t) \) is the source-time function and \( \dot{\alpha}_j^P \) is a constant vector.

Using (2.23) (2.21) and (2.22), and noting that \( \vec{A}_j^P \dot{\alpha}_j^P = \vec{A}_j^P \), we have

\[
\Delta u_k^{PS} = \frac{\partial^2}{\partial t^2} \int_D \rho_0 f^{PS} (\theta^{PS}) \vec{A}_P \vec{A}_k^{PS} \cos \phi S(t - \phi_P - \phi_S) \ dx. \tag{2.24}
\]

The displacement \( \Delta u_k^{PS} \) is determined by the scattering pattern \( f^{PS} (\theta^{PS}) \) and the factor \( \vec{A}_k^{PS} \cos \phi \). The polarization of P-to-S converted waves is parallel to \( \vec{A}_k^{PS} \) and is scaled by \( \cos \phi \), which is defined in (2.17) as \( a_1^{PS} \). This result suggests that the polarization has a symmetric pattern about the radial direction \( \beta^{PS} \) with the maximum
value in radial direction and zero value in transverse direction $\gamma^{PS}$ (Figure 2-1). In addition, it indicates that our radial-component method, which assumes that only the converted S polarization parallel to $\beta^{PS}$ is due to the scattering of incident P wave, is a good approximation. This is also the reason we will see later that this method is asymptotic to layered model interface scattering when we use a very dense array of point scattering bodies to approximate a flat interface.

![Figure 2-1 Polarization of P-to-S scattered wave. The longitudinal axis is the direction of propagation $\hat{a}^s$ of converted S wave. The polarization is scaled by $\cos\phi$.](Image)

**2.2 Algorithm and Implementation**

We implemented a forward program for 3D spherically symmetric geometry based on a radially symmetric velocity model and the single point scattering model described in section 2.1.
This program reads in station coordinates, synthetic earthquake origin and other necessary metadata from an Antelope database. Then it computes the polarization, amplitude and lag time (relative to the first P arrival time) of the spike response for P-to-S scattered wave. Finally, it convolves the spike responses with a Ricker wavelet to produce three-component synthetic seismograms.

2.2.1 Polarization: Radial Component polarization and coordinate transformation.

The radial-component polarization for P-to-S converted waves is proved a good approximation to the analytical results (equation 2.24) in Chapter 2.1.

Here I follow the notation used by Pavlis (2010) and apply the radial-component polarization approximation and some standard results from differential geometry. Three coordinate systems are defined: the World Coordinate System (WCS), the Scatterer Coordinate System (SCS) and the Local Coordinate System (LCS). The origin of WCS is defined at the center of the earth with the z axis pointing to the north pole, the x axis pointing to 0° E, 0° N and the y axis pointing to 90° E, 0° N. The SCS is defined at the point scatterer, where the z axis is the local vertical, the x axis points to east horizontally and y axis points to north. The LCS is defined at the receiver with the z axis vertically up, the x axis east and y axis north.

The first step is to evaluate the unit vectors $L_p$ (local tangent to the converted S wave) and $P_{in}$ (the direction of propagation for the incident P wave) at the point scatterer.
In the SCS at the point scatterer, the azimuth angles of $L_p$ and $P_{in}$ can be calculated from great circle path, while the elevation angles of $L_p$ and $P_{in}$ are computed from Snell’s law for a spherical earth. Once we have $L_p$ and $P_{in}$ in the SCS, we apply a coordinate transformation to convert them to the WCS. The transformation matrix is,

$$A = [E_{sc} \ N_{sc} \ Z_{sc}] = \begin{bmatrix} -\sin\varphi_{sc} & -\sin\theta_{sc}\cos\varphi_{sc} & \cos\theta_{sc}\cos\varphi_{sc} \\ \cos\varphi_{sc} & -\sin\theta_{sc}\sin\varphi_{sc} & \cos\theta_{sc}\sin\varphi_{sc} \\ 0 & \cos\theta_{sc} & \sin\theta_{sc} \end{bmatrix}, \quad (2.25)$$

where $E_{sc}$, $N_{sc}$ and $Z_{sc}$ are basis vectors (column vectors) of the SCS. $\varphi_{sc}$ and $\theta_{sc}$ are the longitude and latitude of the point scatterer, respectively. $L_p$ and $P_{in}$ can be converted to WCS coordinates $L_p'$ and $P_{in}'$ by left-multiplying the transformation matrix $A$,

$$L_p' = AL_p \quad (2.26)$$

$$P_{in}' = AP_{in} \quad (2.27)$$

$L_p'$ and $P_{in}'$ are 3 by 1 vectors.

According to radial-component polarization approximation, the polarization of P-to-S converted wave is in the plane spanned by $P_{in}'$ and $L_p'$ and is perpendicular to $L_p'$. Therefore, it can be written as,

$$R_{sc} = \frac{L_p' \times (P_{in}' \times L_p')}{\|L_p' \times (P_{in}' \times L_p')\|}, \quad (2.28)$$

where $\times$ denotes cross product.

Because of the curvature of seismic ray path, $R_{sc}$ is rotated along the ray path. To account for this rotation from the point scatterer to the receiver, we introduce the Ray
Coordinate System (RCS) which is comparable to the Frenet–Serret frame in differential geometry. RCS is defined by 3 basis vectors \( L'_p \), \( T'_p \) and \( R'_p \). Since \( L'_p \) is already defined, we have

\[
T'_p = \frac{Z_p \times L'_p}{\|Z_p \times L'_p\|}, \tag{2.29}
\]

\[
R'_p = \frac{L'_p \times T'_p}{\|L'_p \times T'_p\|}, \tag{2.30}
\]

where \( Z_p \) is the local vertical vector pointing up. As a result, \( T'_p \) remains constant as the RCS rotates along the ray path, assuming a spherically symmetric earth velocity model.

The RCS at the receiver can be defined by \( L''_p \), \( T''_p \) and \( R''_p \), where \( T''_p = T'_p \). One important fact is that the components of vector \( R_{sc} \) in RCS do not change as the frame rotates. Therefore, the polarization at the receiver can be computed if \( L''_p \) and \( R''_p \) are known. Furthermore, if the basis vectors of the LCS are known, the three components \( E \), \( N \) and \( Z \) can be computed by applying another coordinate transformation to the polarization.

Let \( W \) be the coordinates of \( R_{sc} \) in the RCS at the point scatterer,

\[
W = [R_{sc} \cdot T'_p \ R_{sc} \cdot R'_p \ 0]^T. \tag{2.31}
\]

Then the predicted polarization \( U \) recorded at the receiver can be computed from:

\[
U = BW \tag{2.32}
\]

where \( B \) is the transformation matrix,

\[
B = [E_r \ N_r \ Z_r]^T [T''_p \ R''_p \ L''_p] = \begin{bmatrix}
-sin\varphi_r & -sin\theta_r cos\varphi_r & cos\theta_r cos\varphi_r \\
 cos\varphi_r & -sin\theta_r sin\varphi_r & cos\theta_r sin\varphi_r \\
 0 & cos\theta_r & sin\theta_r
\end{bmatrix}. \tag{2.33}
\]

\( E_r \ N_r \) and \( Z_r \) are basis vectors of LCS. \( \varphi_r \) is the complimentary angle of the back
azimuth of converted S wave and $\theta_r$ is defined as the elevation angle of $L_p$.

### 2.2.2 Amplitude: Geometric Spreading and Scattering Pattern

We have generalized the scattering pattern to 3D in chapter 2.1. The scattering pattern scales the amplitude of P-to-S scattered wave at the point scatterer. It also implies rotational symmetry in isotropic background media.

Here we rewrite the scattering pattern (2.20) as a function of the scattering angle, the density perturbation and the velocity perturbation,

$$ f^{PS}(\theta^{PS}) = -\left[ \frac{\Delta \rho}{\rho_0} \sin \theta^{PS} + \left( \frac{\Delta \rho}{\rho_0} + \frac{2 \Delta c_S}{c_{S0}} \right) \frac{c_{S0}}{c_{P0}} \sin 2\theta^{PS} \right] $$

(2.34)

where $\frac{\Delta c_S}{c_{S0}}$ denotes the S wave velocity perturbation and $\frac{\Delta \rho}{\rho_0}$ denotes the density perturbation. $c_{P0}$ and $c_{S0}$ are P and S wave velocities in the reference media. The scattering angle is $\theta^{PS}$.

Equation (2.34) shows that there is angular dependence in the P-to-S scattering pattern. The maximum amplitude corresponds to an scattering angle of $\sim 45^\circ$ and $\sim 135^\circ$. The zero amplitude occurs when scattered P-to-S wave propagates in the direction parallel to the incident P wave.

Apart from the angular variance, the 3D scattering pattern (Figure 2-2) shows a linear dependence on the perturbation values. When the scattering angle $\theta^{PS}$ is fixed, the amplitude varies linearly with density and velocity perturbation values.
Figure 2-2 Scattering pattern for P-to-S single point scattering. Incident P wave direction is vertical up. Red denotes high amplitude; blue denotes low amplitude.

Generally, the geometric spreading term scales the amplitude by distance. The ray theory amplitude is a high-frequency approximation (Shearer 1999), while the finite frequency method computes amplitude with better stability (Nolet et al 2005). Here we derive the standard form of the geometric spreading amplitude based on ray theory.

A ray path travels from the point scatterer to the surface of the earth with a distance of $\Delta$ (in radians, Figure 2-3). The take-off angle is $\theta$ at the scatterer. If we increase the take-off angle by a small amount $d\theta$, then we get another ray path which reaches the surface of earth $d\Delta$ away from the original one.
The derivation here follows Shearer (1999) except that we apply it to a spherical earth instead of a flat earth.

Now we estimate the energy in the circular belt bounded by rays with take-off angles of \( \theta \) and \( \theta + d\theta \). The area of the circular belt is \( 2\pi \sin \theta d\theta \). The circular belt is on a unit sphere, the area of which is \( 4\pi \). Assuming uniform radiation pattern at the point scatterer, the total energy in the belt is,

\[
E_{d\theta} = 0.5\sin \theta d\theta E_S,
\]

(2.35)

where \( E_S \) is the total energy of the radiation.

When the energy propagates to the receiver at the surface of earth, a new circular belt can be defined with a radius of \( R\sin \Delta \) and a width of \( Rd\Delta \cos \theta' \), where \( \theta' \) is the incident angle of the ray path at surface. As a result, we define the total energy in this belt as

\[
E_{d\Delta} = (2\pi R\sin \theta)(Rd\Delta \cos \theta')\tilde{E} = E_{d\theta},
\]

(2.36)
where $E$ is the energy density at the receiver.

At the scatterer, the spherical earth ray parameter $p_{\text{sph}}$ satisfy,

$$p_{\text{sph}} = r \sin \theta$$  \hspace{1cm} (2.37)

$$dp_{\text{sph}} = r \cos \theta d\theta$$ \hspace{1cm} (2.38)

where $u$ is the slowness at the point scatterer. Combining equation (2.35) and (2.36), and substituting $p_{\text{sph}}$ and $dp_{\text{sph}}$ we have,

$$E = \frac{p_{\text{sph}}}{4\pi u^2 r^2 \sin \Delta \cos \theta \cos \theta'} \left| \frac{dp_{\text{sph}}}{d\Delta} \right| E_S$$ \hspace{1cm} (2.39)

Because the amplitude is proportional to the square root of $E$, we can compute the geometric spreading amplitude term by

$$A_1 = A_0 \sqrt{\frac{p_{\text{sph}}}{4\pi u^2 r^2 \sin \Delta \cos \theta \cos \theta'} \left| \frac{dp_{\text{sph}}}{d\Delta} \right|}$$ \hspace{1cm} (2.40)

where $A_0$ is the initial amplitude at the scatterer and $A_1$ is the amplitude of the ray path reaching the surface of earth.

### 2.2.3 Lag time in arrival time reference frame

Recall that in equation (2.19) the timing relationship is $\delta(t - \hat{\phi}^p - \hat{\phi}^s)$. For convenience, the lag time of a spike response is computed in the arrival time reference frame, which shifts the first P arrival to 0. As a result, the lag time $\tau$ is defined as,

$$\tau = \hat{\phi}^p + \hat{\phi}^s - t^p,$$ \hspace{1cm} (2.41)

where $\hat{\phi}^s$ is the travel time for converted S wave from the scatterer to the receiver, and $\hat{\phi}^p$ is the P wave travel time from event origin to the scatterer. $t^p$ is the travel time of first P arrival. To compute the P wave travel time from the hypocenter to the point
scatterer, we developed a ray-tracing technique based on the Newton’s method. First we define a function $\text{pslow}(\text{distance}, \text{source\_depth})$, which computes the P wave horizontal slowness at the receiver given the great circle distance and the depth of the source. The P wave horizontal slowness should be the same for the wave transmitting from hypocenter to receiver and the wave from scatterer to receiver. That is,

$$\text{pslow}(\Delta + \Delta_s, \text{hypocenter\_depth}) = \text{pslow}(\Delta_s, \text{scatter\_depth}).$$ \hspace{1cm} (2.42)

Let

$$g(\Delta_s) = \text{pslow}(\Delta + \Delta_s, \text{hypocenter\_depth}) - \text{pslow}(\Delta_s, \text{scatter\_depth}) = 0 \hspace{1cm} (2.43)$$

and apply Newton’s method, we have,

$$\Delta'_s = \Delta_s - g(\Delta_s) \frac{(\Delta_s + \varepsilon - \Delta_s)}{g(\Delta_s + \varepsilon) - g(\Delta_s)},$$ \hspace{1cm} (2.44)

where $\Delta_s$ is the current distance from scatterer to receiver, $\Delta'_s$ is a better estimation of the distance and $\varepsilon$ is a very small increment in $\Delta_s$ (since we replace the derivative term by finite difference). The initial distance $\Delta_s$ is set to a small positive real number. Then this iteration continues until the $g(\Delta_s)$ is very close to 0.

Once we have $\Delta_s$ which satisfy the relationship above, we can subtract the total travel time from hypocenter to receiver by the travel time from scatterer to receiver and then get the travel time from hypocenter to scatterer.
2.2.4 Superposition of spike response functions and the choice of wavelet

To this point we discussed the impulse response of a single point scatterer—that is, only the amplitude, polarization and time of an impulse response function is computed. It is important, however, that the synthetic seismic data should be comparable to the data it is intended to simulate. We use a standard approach and apply an analytic wavelet that has the expected bandwidth of the data after deconvolution of the incident P wavefield from the radial component.

Tests in my thesis used two standard wavelets: a Gaussian pulse and a Ricker wavelet. The former is best for deconvoluted functions while the latter is widely used in reflection seismology where it is convolved with reflection coefficients to produce 1D
synthetics. I emphasize, however, that alternative wavelets are an option in the software we developed for this work.

My experiments suggest that the Gaussian pulse is good for single point or widely spaced point scatterers while the Ricker wavelet is appropriate for closely spaced point scatterers where each time sample in the three-component time series is a superposition of point responses from many point scatterers.

Equation (2.21) can be written as,

\[
\Delta u_k^{PS} = \left[ (\frac{\partial^2}{\partial t^2})^{-1} \Delta G_{jk}^{PS} \right] * t \frac{\partial^2 S(t)}{\partial t^2} \tau_j^P
\]

where \( (\frac{\partial^2}{\partial t^2})^{-1} \) is the inverse of \( \frac{\partial^2}{\partial t^2} \) operator and \( \left[ (\frac{\partial^2}{\partial t^2})^{-1} \Delta G_{jk}^{PS} \right] \) represents the spike response. The right hand side term \( \frac{\partial^2 S(t)}{\partial t^2} \) of the convolution operation is the 2\(^{nd}\) derivative of the source time function \( S(t) \). As a result, the scattered wave field displacement \( \Delta u_k^{PS} \) can be computed by convolving the spike response with the 2\(^{nd}\) derivative of source time function, which coincides with the idea of using a Ricker wavelet because it is the 2\(^{nd}\) derivative of the Gaussian function.

From the perspective of receiver function technique, however, the Gaussian pulse is a natural choice. The reason is that the data we are simulating here were computed by the iterative deconvolution method of Ligorria and Ammon (1999), which shape the data to a Gaussian response. This is a subtle theoretical point we return to below.

To demonstrate the subtleties of the wavelet definition in simulations from arrays of point sources we first consider a densely spaced point scatterer grid and the simple source
receiver geometry illustrated in Figure 2.5.

Figure 2-5 2D case: densely spaced point scatterers to approximate a flat interface. * represent the earthquake source, triangles are receivers and white spheres are point scatterers.

For simplicity this example is a 2D line of scatterers immediately beneath a linear array of receivers (Figure 2-5). That is, the source, point scatterers and receivers are all in the same equatorial plane, and we are looking at the southern hemisphere. The point scatterers are all at a common depth of 410 km with an even spacing of 0.05°. The receivers have full coverage of the point scatterers and the receiver spacing is 0.5°.

The iterative deconvolution method used to produce the EARS data we are simulating is described by Crotwell (2007). His implementation of the iterative method of Ligorria and Ammon (1999) applies a Gaussian filter to the radial and vertical component data after deconvolution to a train of impulses computed in the iterative algorithm.
Figure 2-6 2D case: superposition of point scatterer responses with Gaussian wavelet.

When the spike response time series is convolved with a Gaussian wavelet, however, the result of the convolution features small slope net displacement after the first major impulse (Figure 2-6). This happens because the Gaussian pulse has a nonzero DC component and therefore introduces net displacement in the output unless the impulse response has no DC. This is not the case for the simulation result shown in Figures 2.5 and 2.6.

The Ricker wavelet result is appropriate because the Ricker wavelet is the 2\textsuperscript{nd} derivative of Gaussian function and has zero DC-component. In Figure 2-7, the E component is the radial component and Z is the vertical because the ray path travels from west to east in the equatorial plane.
The stacked impulse responses on a single station convolved with a Ricker wavelet are shown in Figure 2-7. This result is comparable to the response of a simple flat interface, which suggests that the closely spaced point scatters interfere with each other and they behave as a discontinuity subsurface as a whole.

We then generalized the 2D test to 3D. Like the 2D test, the source and receivers are still in the equatorial plane, except that the point scatterers are expanded in north and
south to approximate a subsurface at 410 km depth (Figure 2-8). The subsurface is 8° by 8° bounded by meridians and parallels, and is symmetric about the equator.

The superposition of the responses of this array of 3D point scatterers is also equivalent to that of a simple interface. The pulses (Figure 2-9) only have minor difference from the stacked impulse responses in 2D (Figure 2-7), which might come from the point scatterer responses out of the equatorial plane. The radial direction is defined by the first P arrival, not the scattered waves from in and out of the equatorial plane because each scattered ray path has a different radial direction.

Figure 2-9 3D case: superposition of point scatterer responses with Ricker wavelet. Scattering pattern is enabled in the forward program.
Figure 2-10 3D case: superposition of point scatterer responses with Ricker wavelet. Scattering pattern is disabled in the forward program.

The 3D test also suggests that the scattering pattern helps the impulse responses stack better. The synthetic data without a scattering pattern is shown in Figure 2-10, where some wiggles appear after the major impulse response. These wiggles should not exist according to the theory of ray transmission and conversion at a simple flat interface. For comparison, the 3D test with the scattering pattern (Figure 2-9) doesn’t have these wiggles. Therefore the scattering pattern is important in simulating scattered wavefield imaging because it may better recover the original scattering potential without introducing additional errors.

2.2.5 Parallel Algorithm

MPI stands for Message Passing Interface, which is a standard interface in parallel computing (Snir et al., 1998). The memory model of MPI is based on the distributed memory machines, but it also works on machine with multi-cores where each core is
assigned to an independent memory space. Because each CPU has its own memory space, there are communications among different CPUs, either the status of current job or the data, which help synchronizing working processes on all CPUs.

This program can be parallelized over 1) stations or 2) point scatterers. Suppose \( p \) cores/CPUs are used in the computation. In the first case the stations are equally partitioned into \( p \) subsets, each of which is assigned to a core. This parallel implementation is relatively easier but its parallel efficiency is less than 50\% when the number of point scatterers is more than 50,000. In the second case, the point scatterer grid is partitioned equally among \( p \) cores and the synthetic seismogram on each station is a superposition of waveforms computed on all \( p \) cores. Although the second implementation is more complicated and requires additional communication, it greatly improved the parallel efficiency to \(~80\%\) when the number of point scatterers is more than 100,000.

The strategy we employ here is a data-parallel architecture, in which each processor reads in the metadata (origin, receiver locations and time series definitions) for the 1\(^{st}\) event from a database and then sends the results back to processor 0 which not only computes its own task, but also assembles all results received from other processors and writes them to the output database. After the 1\(^{st}\) event is processed, all processors move on to read and process the data for the next event until all events are processed.

The message passing technique is implemented by two important MPI functions,
MPI_Send and MPI_Recv. They are blocking send and receive functions because both functions do not return until the data are sent or received. This feature synchronizes processor 0 with other processors because other processors cannot read the next event data until their current results are sent to processor 0, although processor 0 does some extra work on I/O.

The parallelized algorithm for the forward synthetics is very effective in reducing the time required to complete a simulation, especially when the point scatterers are very dense. The parallel efficiency is ~0.9 for 8 processes running on 8 cores of the same node on Quarry (a high performance cluster at Indiana University), when the number of point scatterers is small. It might be slower when running on multiple nodes due to a longer communication overhead for data transmitted via inter-node network, but running on more nodes will save a considerable amount of time for the forward program, which makes the inverse problem solvable based on fast forward program.
3 Resolution test and 410 discontinuity

The 410 discontinuity, marked by the $\alpha$-$\beta$ phase transition of olivine, is the upper boundary of the mantle transition zone (Akaogi et al., 1989; Anderson, 1989). In the earlier seismological studies, a single peak appears on all global seismograms at 410 km depth, therefore the 410 discontinuity was considered a flat, thin and uniform boundary. Nowadays, however, people argue about its topography, thickness and lateral variations in amplitudes. Fine details about 410 have been discovered by new methods with better data quality (Melbourne and Helmberger, 1998; Flanagan and Shearer, 1999; Houser and Williams, 2010). The topography and the thickness of 410 are related to the Clapeyron slope $dP/dT$ (Alex Song et al., 2004; Houser and Williams, 2010). Some experiments also suggest a low velocity layer on top of 410 under northwestern US due to high water content brought by subduction (van der Meijde et al., 2003; Alex Song et al., 2004).

This chapter firstly describes several tests on point resolution and the imaging of a horizontal or a dipping subsurface approximated by densely spaced point scatterers. Then it discusses the nature of 410 discontinuity in terms of its topography, laterally variation and thickness.

With the help of the forward scattering program, we are able to generate synthetic seismograms for any predefined model geometry with source, receiver and point scatterer locations. The goal here is to evaluate how well the original geometry of scatterers can be recovered by the plane-wave migration program described by Pavlis (2011). The
objective is to use the simulations to provide insights into the resolution of scattered wavefield imaging and the understanding of the 410 discontinuity.

The image geometry of western US is defined as 2450*2100 (km$^2$) horizontally and 800 km in depth and is the same geometry that we apply to our simulation tests. The synthetic data are generated from 60 events, the locations of which are the same as the 60 stacked events from USArray data.

### 3.1 Widely spaced point scatterers

Based on single point scattering model, we evaluated the widely spaced point scatterer model in terms of the imaging quality and point scatterer reconstruction. We apply both a Gaussian pulse and a Ricker wavelet to convolve with the point spike response and compare the results. Later we investigated the point resolution at 410 km depth by using a regularly spaced point scatterer grid with large point spacing.

The first model we constructed contains five point scatterers that are far apart, three of which are at 410 km depth and the other two are at 660 km (Figure 3-1).

When a Gaussian pulse is applied, image reconstruction shows that all five points can be recovered in the imaging volume (Figure 3-2). As described in section 2.2.5, the Gaussian pulse has 101 samples with a sample interval of 0.05 s and standard deviation of 0.5. Because the imaging program is based on EARS data which are computed by iterative deconvolution method with Gaussian filter applied, the synthetic data based on Gaussian pulses are comparable to the EARS data and produce disc-shaped high
scattering potential in the image volume. The contour surfaces are colored by gradient of scattering potential.

Figure 3-1 Model geometry of five widely spaced point scatterers. Yellow spheres are point scatterers.

Figure 3-2 Image of five widely spaced point scatterers using Gaussian pulse. Contour surfaces are computed for scattering potential equal to $4.00 \times 10^{-8}$. The contour surfaces are colored by gradient of scattering potential.

The interferences among different point scatterers should be small since all points are sufficiently far apart. However, the scattering hyperbola of a shallower point may affect
the hyperbola of a deeper point because of the geometric spreading of the deeper point scatterer amplitudes. In addition, the shapes of those colored isosurfaces show that the vertical resolution of a single point is much better than the horizontal resolution.

Figure 3-3 Image of five widely spaced point scatterers using Ricker wavelet. Contour surfaces are computed for scattering potential equal to $3.10 \times 10^{-9}$. The contour surfaces are colored by gradient of scattering potential.

The image of the five point scatterers using Ricker wavelet for synthetics also recovers all the scatterers but it also generates more noise in the volume. The Ricker wavelet used in Figure 3-3 has a central frequency of 0.5 Hz and the same length as the Gaussian pulse.

Although the Ricker wavelet agrees with the theoretical derivation of scattered wavefield displacement, it contradicts the fact that the imaging program is based on back-projecting the polarizations of phases to the isocron surface. The Ricker wavelet is DC-balanced and has one central peak and two side lobes (trough) and the polarization
does change, thus making the back-projection of each data point hard to focus uniformly around the point scatterer like the data points of a Gaussian pulse. Despite of the noise that the Ricker wavelet method introduces, the Ricker wavelet still improves the imaging quality by bringing in the peak and trough and enhancing interference among point scatterer responses.

Figure 3-4 Image slice across 3 point scatterers. Synthetic data generated by convolving the spike responses with a Gaussian pulse. Yellow spheres show the actual position of point scatterers.
Figure 3-5 Image slice across 3 point scatterers. Synthetic data generated by convolving the spike responses with a Ricker wavelet. Yellow spheres show the actual position of point scatterers.

In Figure 3-4 and Figure 3-5, I display the image of 3 point scatterers on a 2D slice. For the Gaussian pulse method (Figure 3-4 and Figure 3-5), the high scattering potentials (red) smear out in the shape of a disc, while for the Ricker wavelet method, the red colors are more focused around the point scatterers with the characteristic alternating red-and-blue pattern. Given the same length for the both the Gaussian pulse and the Ricker wavelet, these results suggest that the Ricker wavelet (central frequency $f_c = 0.5$ Hz) could potentially better improve the image resolution of a single point scatterer than the Gaussian pulse ($\sigma = 0.5$), although it introduces noise.
Figure 3-6 Map view of point scatterer resolution test at 410 km depth. White points are point scatterers defined in the forward model and the red peaks are their reconstruction. A Ricker wavelet is used to convolve with the spike responses.

To better understand the horizontal resolution at a depth of 410 km, we build a regularly spaced point scatterer grid with a point spacing of 280 km. The S wave velocity perturbation $\frac{\Delta c_S}{c_{S0}}$ is -1%. The minus sign is needed because when the S velocity decreases it creates a negative perturbation. The yellow lines are geographic and state boundaries of western US.

The map view of the imaging result (Figure 3-6) shows that generally point scatterers that have good station coverage can be reconstructed by the imaging program with
different quality and the horizontal resolution at this depth is better than 280 km. The difference in illumination of recovered point scatterers is probably due to irregular station coverage and unbalanced azimuthal distribution of earthquake events as described by Pavlis (2011b).

![Figure 3-7 Cross-section of point scatterer resolution test at 410 km depth. The red and blue pairs are the reconstruction of point scatterers.](image)

The cross-section (Figure 3-7) of the image volume reveals subtle patterns of the responses of each point scatterer. Each point scatterer that crosses this vertical slice corresponds to a red-and-blue pair, which is partly due to the DC-balanced Ricker wavelet. It again shows that the vertical resolution (~10 km) of point scatterers at 410 km depth is much better than the horizontal resolution.

Noise is seen in both the vertical slice and the map view, especially at the west and southeast edges. The edge effect might be a result of lack of station coverage for some
point scatterers. Because of the pseudo-station stacking method that the imaging program applies (Pavlis 2011), some pseudo stations outside of the coverage could have nearly identical stacked waveforms, which might hinder the back-projecting and focusing of migration. The noise could also be related to the Ricker wavelet, as has been discussed before.

To summarize, the widely spaced point scatterer experiments are appropriate for their purpose as resolution tests, but in reality this infinitely small point scatterer doesn’t usually exist in the form of isolated scatterers individually. The point scatterers usually gather together in a very dense form and appear as discontinuities, which will be discussed in the next section.

3.2 Flat Subsurface Approximation Based on Point Scatterers or Simple Flat Layer Model

The classic 1D earth model is a layered model in which each layer is assigned a density, a P-wave velocity, an S-wave velocity, and in some cases a velocity gradient. The interface between two layers is treated as a mathematical boundary to which a boundary condition is applied. Therefore the P-to-S conversion at the boundary can be computed theoretically using the concept of reflection and transmission coefficients (Stein and Wysession, 2003).

Now with the help of the point scatterer forward model, we are able to approximate a flat subsurface with a densely spaced point scatterer grid. Each point scatterer in the grid
is regarded as a point source which is illuminated by the incident P wavefield and produces scattered S wavefield. Similar to the Kirchoff integral, this method applies the superposition of point source response functions to construct the new wavefront for P-to-S scattered wavefield. This scenario can be traced back to the Huygens–Fresnel principle, which explains plane wave refraction at an interface in terms of multiple closely spaced point sources at the interface.

3.2.1 Simple Layered Model

![Figure 3-8 Cross-section of the reconstructed image using four flat interfaces and a Gaussian function.](image)

The synthetic data for simple layered model are generated by the Simple Layer Simulation program used by Pavlis (2011). Here we convolved the spike responses of major interfaces with Gaussian function and Ricker wavelet and compared the results. The major interfaces are defined at depths of 30 km, 90 km, 410 km and 660 km with initial converted wave amplitudes of 0.5, -0.2, 0.3, and 0.3, respectively. Both the Gaussian pulse and the Ricker wavelet contain 101 points with 0.05 sec sample interval.
The variance of Gaussian pulse equals $0.5^2$ and the central frequency of Ricker wavelet is 0.5 Hz.

The slices are at the same position for both Figure 3-8 and Figure 3-9. Generally, the image of major interfaces computed from Gaussian pulses is characterized by higher amplitudes than the one computed from Ricker wavelet. In addition, the Ricker wavelet result features better vertical resolution (with alternating plus and minus values) than the Gaussian pulse result. That’s probably because the Ricker wavelet is DC-balanced and the troughs cancel out some amplitude of the peak, while the Gaussian pulse has only positive DC-component.

![Figure 3-9 Cross-section of the image of four major interfaces using a Ricker wavelet.](image)

The 410 and 660 km discontinuities are clear on both slices (Figure 3-8 and Figure 3-9). They take the shape of smooth horizontal boundaries, which agrees with the fact that the Simple Layer algorithm is based on the mathematical boundary condition of the interface between different elastic media. It also shows that the lateral amplitude
variations along the 410 and 660 km discontinuities are not significant with current station coverage, hypocenter locations and imaging volume geometry. This result suggests that if the 410 km discontinuity is uniform without lateral changes, the image of it should not suffer from great amplitude fluctuations.

We do not see a high noise level in the Ricker wavelet results, which is different from point scatterer images using Ricker wavelet. Therefore, the use of a Ricker wavelet is appropriate for the flat interface model.

3.2.2 Superposition of densely spaced point scatterer responses based on Ricker wavelet

A densely spaced point scatterer grid is implemented to approximate the 410 km discontinuity. The point spacing is 5 km and the grid contains 199479 point scatterer at 410 km depth for the western US geometry.

Figure 3-10 Cross-section of the image of 410 km discontinuity simulated by a densely spaced point scatterer grid.
scatterer grid. A Ricker wavelet is applied.

The cross-section image shows that the superposition of densely spaced point scatterer responses is comparable to the Simple Layered model result (Figure 3-9) which is a theoretical solution by matching boundary conditions of the wave equation. Despite of the sharpness of the interface at 410 km produced by superposition, however, the phase of the stacked impulse responses is changed by ±180° due to integration of point scatterer responses. The Simple Layered result is characterized by thinner interface image than the point scatterer model based on superposition because the pulse width is narrower for a single Ricker wavelet than a superposition of multiple ones.

![Image of the simulated 410 km discontinuity using Ricker wavelet. Contour surfaces (in gray) are computed for scattering potential equal to 4.52 x 10^-9. The white grid is computed from Delaunay triangulation of point scatterers.](image)

The contour view of the image of 410 km point scatterer grid (Figure 3-11) indicates
that the lack of station coverage outside US leads to insufficient point scatterer recovery outside the boundary. The white subsurface is a triangulation of point scatterers defined in the forward model. The good agreement between the contour and the white subsurface suggests that the superposition of densely spaced point scatterers provides accurate representation of a horizontal discontinuity interface, which is important to justify the use of this approach to model a curved subsurface.

### 3.3 Simulating a Dipping Layer with Point Scatterers

A dipping discontinuity is a good analog to a subducted slab because of different physical properties between the slab and the surrounding crustal or mantle material. The subducted slab, depending on its depth and dip angle, is generally characterized by lower temperature and seismic velocity than the background mantle material. Therefore it’s realistic to simulate it by a closely spaced point scatterer grid with negative S wave velocity perturbation values.

![Figure 3-12 Image of a dipping layer: vertical slice view. Red color corresponds to the dipping](image-url)
interface. A Ricker wavelet is applied.

The slab model assumes a linear relationship between horizontal distance (in radians) and depth (in km). The slab goes from 210 km to 410 km in depth and spans over 1000 km horizontally.

A vertical slice of the dipping layer image displays the dipping layer with thin, sharp red color (Figure 3-12). The edge of the discontinuity has shown some scattering phenomena but with low amplitudes. The amplitude along the dipping interface is generally uniform with minor variations.

![Figure 3-13 Image of a dipping layer: contour view. The green surface is the grid of point scatterers and the black contour surface is its reconstruction in image. Contour surfaces are computed for scattering potential equal to $5.00 \times 10^{-9}$. A Ricker wavelet is applied.](image)

The contour representation of the dipping discontinuity (Figure 3-13) shows good agreement between the original scatterer grid geometry and its image which is the output of the prestack migration imaging program. As is in 3.2.2, the point scatterer grid (green, dipping surface) is the triangulation of the scatterers. It also indicates that the resolution of edges at shallower depth is better than that at a deeper depth.

41
3.4 Checkboard test

We also performed a checkboard test to evaluate how well the imaging program can resolve point scatterers with alternating plus and minus perturbation values (polarities). This approach is commonly used in seismic tomography so we wanted to evaluate how well this might work for this imaging method.

Four sets of point scatterers at depths 30 km, 90 km, 410 km and 660 km, respectively, were defined in the forward model with horizontal point spacing of 140 km. A horizontal map (Figure 3-14) shows the recovered point scatterers at 410 km depth. The point scatterers within the seismic station coverage can be well reconstructed while those on the edge or outside of the coverage cannot be recovered.
Figure 3-14 Map view of checkboard test at 410 km depth. Point spacing is 140 km.

The vertical slice (Figure 3-15) of the checkboard test, however, implies that the point scatterers may affect the imaging quality of the point scatterers below. One possible explanation is that a shallower point scatterer produces scattering hyperbolae with higher amplitudes than a deeper point scatterer. As a result, the scattering hyperbolae produced by shallower point scatterer may erase the hyperbolae produced by deeper point scatterer. This example suggests that the checkboard test for point scatterers at different depths might not be appropriate for scattered wave field imaging. This is simply because this model is contrary to the basic objective of this imaging algorithm. The plane wave
migration algorithm fundamentally aims to focus scattered waves generated from continuous, curved surfaces. The checkboard model will tend to generate the looks more like the kind of noise the algorithm aims to suppress. Hence, we do not recommend this approach for appraising errors in this form of imaging.

Figure 3-15 Vertical slice of checkboard test at 30 km, 90 km, 410 km and 660 km depth. Point spacing is 140 km.

3.5 Application to 410 km discontinuities

At the end of this chapter, we analyze the amplitude variation along the 410 km discontinuity from western US imaged by the prestack migration program (Pavlis, 2011). Based on the discussions on widely or densely spaced point scatterer grids, we propose that the 410 discontinuity is comparable to the densely spaced point scatterer model with varying perturbation values and varying depth at 410 km.

The 410 discontinuity reconstructed in our imaging shows large amplitude variations
on a vertical slice (Figure 3-16), as well as significant topographic variations. These variations might be related to three possible explanations: 1) the perturbation in physical properties (shear wave velocity or density) of 410 km discontinuity varies laterally; 2) the seismic station coverage is irregular or 3) the Signal-to-Noise Ratio (SNR) varies by stations. In the numerical experiment of a horizontal, densely-spaced point scatterer grid (section 3.2.2), the irregular station coverage doesn’t introduce significant lateral variations in amplitude. The SNRs of the waveforms are carefully controlled and shouldn’t cause great amplitude variations either. As a result, the amplitude variation is most likely related to spatially varying physical properties.

Figure 3-16 Vertical slice of the amplitude variations at 410 km depth based on USArray data.

The map view (Figure 3-17) of the 410 km depth of the imaging volume also suggests that the amplitude varies laterally. However, the high amplitude anomalies at the edges of the seismic station coverage do not accurately reflect the amplitude variations,
which seem to be related to an as yet undetermined edge normalization problem in the migration programs (Pavlis 2011).

Figure 3-17 Map view of the amplitude variations at 410 km depth based on USArray data.
4 Estimating the uncertainty of Scattered Wavefield Imaging by Trans-dimensional Monte Carlo Inversion

Scattered wavefield imaging is based on the single point scattering model, which is derived from perturbation theory (Beylkin and Burridge 1990; Bostock et al 2001; Poppeliers and Pavlis 2003). For P-to-S converted wave imaging, only the S wave velocity perturbation and the density perturbation are related to the scattered wavefield. Discontinuity can be seen as a combination of these two perturbations. The question I bring forward here is: what is the horizontal resolution in identifying discontinuity at a certain depth for a single seismic station? Because of the nonlinear nature of the correspondent forward problem, it’s reasonable to conjecture that this horizontal resolution varies with its position relative to the station, as well as the incident P wave direction.

In this chapter, a forward program is created which calculates synthetic seismograms for predefined 2D/3D model geometry, 1D velocity model and distribution of point scatterers. This forward program connects the perturbations in physical properties (S velocity, density) with the impulse responses in the waveform. In the next step, a trans-dimensional Monte Carlo inversion based on the fast computation of the forward problem will calculate the most probable number of point scatterers and their perturbation value distributions at a common depth. A hypothesis on the high uncertainty zone is provided in the end.
4.1 Forward model

The forward program presumes a 2D flat earth geometry. A 1D velocity model AK135 is applied to the geometry. Assuming plane wave incidence at the bottom of the geometry for teleseismic P wave, the two-component synthetic waveforms can be computed from the scattered wavefield generated by plane wave incidence at point scatterers. Generally, this is a non-linear forward problem and can be formulated as,

\[ \mathbf{d} = G(\mathbf{m}) \]  

(4.1)

The model parameter \( \mathbf{m} = (\mathbf{\beta}, \mathbf{x}, \mathbf{y}, k) \), where vector \( \mathbf{\beta} \) contains S wave velocity \( (V_S) \) perturbations of point scatterers, vector \( \mathbf{x} \) contains coordinates of point scatterers, vector \( \mathbf{y} \) stores the coordinates of receivers and \( k \) is the number of point scatterers in the model space. \( \mathbf{d} \) represents the three-component seismograms on all receivers generated for a given \( \mathbf{m} \) in the simulation program.

![Figure 4-1 Illustration of the forward model.](image)

The forward program shares the same theoretical background as in chapter 2.1 except
that it presumes flat earth geometry and we have to write our own version of the travel
time calculator.

4.2 Inverse model and results

The inverse problem arises from inverting the optimal distributions of Vs
perturbation values by minimizing the least-square misfit of waveforms (Figure 4-2). We
assume only one receiver and the locations of receiver and point scatterers are fixed. As a
result, the model parameters we aim to estimate are $V_s$ perturbations $\beta$ and model
space dimension $k$.

First of all, we produce the synthetic data by computing the waveforms based on
known point scatterer $V_s$ perturbations and adding normally distributed noise. Then we
choose a maximum number of point scatterers $N$, fix their locations and sample the prior
information to give them initial $V_s$ perturbation values.
Figure 4-2 waveform fitting based on Monte Carlo method. Normally distributed noise is added to the synthetic waveform data.

In each iteration, the parameter $m_{\text{cand}}$ will be perturbed by either random walk in $V_S$ perturbation value of a randomly picked point scatterer, or trans-dimensional moves (birth or death of point scatterers). A new set of waveforms will be calculated using the forward program $d_{\text{cand}} = G(m_{\text{cand}})$. Then the likelihood function $L(m_{\text{cand}})$ is derived and a Metropolis rule is applied to decide the acceptance of the new parameter candidate as a new state in the Markov Chain. The iteration can be parallelized once the burn-in phase finishes.
Figure 4-3 The 2D model geometry for the synthetic data. The vertical axis is in depth. Seven point scatterers with alternating plus and minus perturbations are defined at 100 km depth with equal spacing of 30 km.

The 2D model geometry is defined as 200 km wide (X) and 210 km deep (Z). The receiver is located at 100 km in X (Figure 4-3). In the synthetic model, seven point scatterers with alternating -0.01 and 0.01 \( V_S \) perturbation values are located at 100 km depth with 30 km horizontal point spacing (Figure 4-3). The incident plane P wave has a ray parameter of 0.051298 sec/km and travels from the lower left corner to the upper right (positive X component). Three-component waveform data are generated but for the 2D problem we only need E and Z components. Normally distributed noise is added to
the waveform data: \( \varepsilon \sim N(0, \sigma^2) \), \( \varepsilon \) is a time sample in the waveform time series.

For the inverse problem, 37 point scatterers are defined at the same depth as in the synthetic model, except that their point spacing is 5 km, which is smaller. The goal of the inverse program is to evaluate how well the seven point scatterers in the synthetic model can be recovered.

The Bayesian formula is,

\[
p(m|d) \propto p(d|m)p(m)
\]

where \( p(m) \) is the prior Probability Density Function (PDF), \( p(d|m) \) is the likelihood and \( p(m|d) \) is the posterior PDF.

The likelihood function is defined as the least-square misfit of the data,

\[
L(m_{\text{cand}}) = p(d|m_{\text{cand}}) = \exp\left(-\frac{1}{2\sigma^2} e^T e\right)
\]

where \( e = d - G(m_{\text{cand}}) \) is the waveform misfit between the synthetic data and the data computed from the candidate model parameter \( m_{\text{cand}} \).

The Metropolis-Hasting rule we apply here to sampling the posterior probability is

\[
\alpha = \min\left[1, \frac{q(m|m_{\text{cand}})}{q(m_{\text{cand}}|m)} \frac{p(m_{\text{cand}})}{p(m)} \frac{L(m_{\text{cand}})}{L(m)} |J|\right],
\]

where \( q(m_{\text{cand}}|m) \) is the proposal distribution for moving from \( m \) to \( m_{\text{cand}} \), \( p(m) \) is the prior distribution, and \( J \) is the Jacobian of the transformation from \( m \) to \( m_{\text{cand}} \).

The determinant of Jacobian \( |J| \) is 1 if we assume the birth is the exact inverse of death and the sum of model space dimension and null space dimension is constant (equals 37 in this example) during trans-dimensional moves (Bodin and Sambridge 2009;
Agostinetti and Malinverno 2010).

It is also proved that the ratio \( \frac{q(m|m_{\text{cand}})}{q(m_{\text{cand}}|m)} \frac{p(m_{\text{cand}})}{p(m)} = 1 \) if the prior distribution equals the proposal distribution (Agostinetti and Malinverno 2010).

As a result, the Metropolis rule in equation (4.4) is reduced to

\[
\alpha = \min \left[ 1, \frac{L(m_{\text{cand}})}{L(m)} \right]
\]

for trans-dimensional moves.

Because of the added prior information, the Metropolis rule for the ordinary steps when the model dimension doesn’t change is,

\[
\alpha = \min \left[ 1, \frac{p(m_{\text{cand}})}{p(m)} \frac{L(m_{\text{cand}})}{L(m)} \right]
\]

which samples the posterior PDF according to the Bayesian formula. The ratio of proposal PDFs becomes 1 because of symmetry.

The advantage of this trans-dimensional Monte Carlo method is that it can automatically determine the best model space dimension (Bodin and Sambridge 2009).

Let the variance of data noise be \( \sigma^2 = (1.2 \times 10^{-6})^2 \), and the variance of proposal PDF and prior PDF equal to 0.005\(^2\). The total number of iterations is \( 1.25 \times 10^6 \). The Markov chain length is 100000 after the burn-in part is removed.

The inverted results contain the marginal probability distributions on 10 point scatterers (Figure 4-4) and the probability distribution of model dimension shows that the best estimation of model dimension is 9 or 10 (Figure 4-5). The marginal distribution of point 11 shows a mean value close to zero, which explains the difference between 9 and 10
point scatterers in model space.

Figure 4-4 Marginal probability distribution of $V_S$ perturbation on 10 point scatterers.

Figure 4-5 Distribution of the model space dimension (number of point scatterers k).

Based on the marginal probability distributions of $V_S$ perturbation, we calculated the average perturbation (Figure 4-6) and standard deviation (Figure 4-7) for each point scatterer in model space by statistical approaches.
Figure 4-6 Inverted $V_S$ perturbation based on Reversible Jump MCMC method. The total number of dimensions of the inverse model is 37 and 10 point scatterers out of 37 are in the model space of the inverted results.

The average $V_S$ perturbation value (Figure 4-6) shows that the point scatterer at 70 km in synthetic model is not well recovered, while all the other scatterers are fully recovered by this trans-dimensional MCMC method. Points 11-14 (Figure 4-4) that are not well recovered are associated with the large error bars in Figure 4-7, which indicate high uncertainty in inverting for the $V_S$ perturbation on those points.
4.3 Discussion and conclusion

The results suggest that for a 2D flat earth geometry, the single receiver waveform cannot resolve single point scatterer very well at certain location. The question is: how is this high uncertainty zone related to the relative positions of point scatterer, receiver and incident P wave?

Our conjecture here is based on the Fermat’s principle, which is comparable to the Snell’s law in seismology. I propose that the earliest P-to-S scattered wave arrival on the receiver is generated by a point scatterer which is on the shortest path. If we assume the point spacing of scatterers is small enough that the scatterers approximate a flat subsurface, the shortest path is the same as the one defined in Snell’s law for converted waves. If one point scatterer is on the shortest path, the point responses generated by its neighbors all arrive later than the earliest arrival. But what’s the time delay between the

Figure 4-7 Standard deviation on each point scatterer in model space.
earliest arrival and the later arrivals, and how does this nonlinear time relationship among different point scatterer responses contribute to the high uncertainty zone?

The time relationship for scattered wave recorded at the receiver is $\delta(t - \phi^p - \phi^s)$, where $\phi^p$ is the P wave travel time from origin to point scatterer, and $\phi^s$ is the converted S wave travel time from scatterer to receiver (Pavlis in press). When $t$ is constant, this equation defines a set of point scatterers that satisfy the time relationship and contribute to the same time sample on the receiver. These points construct a surface called isocron. Therefore, the isocron for the earliest P-to-S arrival has a tangent plane which is approximated by densely spaced horizontal point scatterers (Figure 4-8). The tangent point corresponds to the point scatterer on the shortest path.

Figure 4-8 Shortest path and tangent point of isocron.
As a result, the point scatterers in a small neighborhood of the tangent point generate point responses that arrive at nearly the same time, which makes them hard to distinguish. Or in other words, those point scatterers produce a single pulse that is comparable to a converted wave pulse of a flat interface. That’s why we couldn’t recover the individual point scatterers in the small neighborhood of the tangent point, which defines the high uncertainty zone.

Because the tangent point of isocron changes when the origin of earthquake changes, the resolution is likely to be enhanced if the images produced from multiple events are stacked together (Poppeliers and Pavlis 2003). On the other hand, if the edge of a subsurface falls into the high uncertainty zones of the most of the nearby stations, it might not be imaged very accurately by scattered wavefield imaging. Last but not least, the lack of spatial sampling (very large receiver spacing) of a flat subsurface might produce holes in the images because the image of a flat subsurface with no edges is based on the pulses generated in the high uncertainty zones.
5 Conclusion and Suggestion on Future Research

In this project I extended the derivation of P-to-S forward scattering formula for the single point scattering model (Beylkin and Burridge 1990) and adapted it to the development of a 3D forward program. Two important features of the theoretical results are discussed: 1) the 3D scattering pattern and 2) the radial-component approximation for the polarization of the converted S wave. I also tested the forward program by comparing the seismograms of densely spaced point scatterer responses on a linear seismic array using two different wavelets: a Gaussian pulse and a Ricker wavelet (2nd derivative of Gaussian pulse). The results show that the Ricker wavelet is a better choice for the superposition of densely spaced point scatterer responses, which agrees with the theory.

The results from the widely spaced point scatterer test prove that the single point scattering model is valid and the point scatterers in forward model, if not interfering with each other much, can be recovered by the prestack migration imaging program. In addition, the size and shape of point reconstruction in the image volume give some hints on the best possible point resolution at a certain depth with current USArray station coverage. The vertical resolution is much better than the horizontal resolution at a given depth. This is a direct consequence of the pseudostation stacking used in this prestack algorithm which allows the method to perform wavefield processing on spatially aliased data (e.g. USArray data; Pavlis, 2011b). These results confirm with numerical experiments an assertion made in Pavlis (2011a) the so called “migration impulse
response”.

The superposition of multiple point sources, on the other hand, demonstrates that a dense point scatterer grid can approximate a flat subsurface or a dipping subsurface. The stack of point scatterer responses from this dense flat grid produces a response comparable to that of a simple flat interface. Furthermore, a dense, dipping grid of discrete point scatterers can generate a dipping subsurface which is continuous. The very small spacing of point scatterers also leads to interference of different point scatterer responses such that they cancel out one another on the trailing part of seismogram. The choice of wavelet has a significant impact on the synthetic seismograms. A Gaussian pulse is not an ideal wavelet because it has a net DC component and generates long slope following the first pulse, while the Ricker wavelet works better for modeling a surface.

The message that these tests convey is that the unknowns we aim to estimate in scattered wavefield imaging are the perturbation values, and the 410 km discontinuity is comparable to closely spaced point scatterers but with varying perturbation values (amplitudes, scattering potential). In fact, the discontinuity itself can be defined by density and shear wave velocity perturbations. The widely spaced point scatterer test is good for its purpose as a resolution test but that situation does not exist in reality. In reality, the more common situation is a subsurface such as 410 km discontinuity with topography and varying scattering potential or the Farallon slab which is a dipping layer. The subsurface can be approximated by a grid of densely spaced point scatterers. Despite
of the edge effect which comes naturally in the synthetic seismograms that this point source model generates, the edge of subsurfaces (whether horizontal or dipping) can be recovered at an acceptable accuracy (resolution?).

Perhaps more enlightening, the variations in scattering potential (the perturbation values) along the 410 discontinuity can reveal detailed physical processes across the boundary because the perturbation values are directly related to the physical properties – density and shear wave velocity.

In chapter 4, the Monte Carlo inversion for perturbation values of point scatterers reveals the relationship between the tangent point of isocron and the low resolution zone which is caused by pulses received at nearly the same time. As a result, the information about perturbation values in the small neighborhood of the tangent point of isocron are lost, although this tangent point is important for the imaging of continuous subsurfaces.

For future work, I propose a simplified inverse method which constrains both the depth and the perturbation values of point scatterer in a densely spaced grid. A time window which contains primarily the scattered wave responses of the 410 discontinuity can be defined for the Receiver Function (RF) data. A Bayesian inversion can be implemented by minimizing the least-square residual between the synthetic data and the RF data in the time window. This method will be useful in reducing the computational complexity and localizing the image to a specific discontinuity subsurface.
References


