Imaging of structure at and near the core mantle boundary using a generalized Radon transform

by

Ping Wang

Submitted to the Department of Earth, Atmospheric and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author .....
Department of Earth, Atmospheric and Planetary Sciences
April 26, 2007

Certified by .....
Robert D. van der Hilst
Cecil and Ida Green Professor of Earth and Planetary Sciences
Thesis Supervisor

Certified by .....
Maarten V. de Hoop
Professor, Purdue University
Thesis Supervisor

Accepted by .....
Maria Zuber
Head, Department of Earth, Atmospheric and Planetary Sciences

Imaging of structure at and near the core mantle boundary using a generalized Radon transform

by

Ping Wang

Submitted to the Department of Earth, Atmospheric and Planetary Sciences
on April 26, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this thesis, concepts from inverse scattering and modern statistics are combined into a powerful tool for imaging interfaces in Earth's deep interior. Specially, a generalized Radon transform (GRT) approach is developed to image heterogeneity at and near interfaces in Earth's lowermost mantle with broadband, three-component seismograms from Global Seismograph Networks (GSN). With this GRT method I transformed ~100,000 transverse-component *ScS* waveforms into image gathers of a core mantle boundary (CMB) patch beneath Central America and juxtaposition of stacks of these gathers produces a 2-D image profile. To enhance this image profile, I collaborated with statisticians and used mixed-effects statistical modeling to produce the best estimates of reflectivity along with their uncertainty. I demonstrate that the method outlined above works well and – thus – paves the way to large-scale seismic exploration of the lowermost mantle. With the new technology I mapped the structure at and near the CMB beneath Central and North America. Several interfaces are detected, and some of them are consistent with expectations from phase transformations in Magnesiüm perovskite. If we know which interface is associated with a particular phase transformation, and if we know the thermodynamic (P-T) relations of the stability fields of the phases, then we can estimate temperature from the pressure as inferred from the depth at which the transition occurs in the seismic sections. Here we associate a seismically observed wavespeed increase with the perovskite to post-perovskite transition and a wavespeed decrease with the back transformation to perovskite. Using P-T data from experimental and theoretical mineral physics we can then estimate the lateral temperature variations and radial (thermal) gradients near the CMB. In addition, the temperature of the CMB and global heat loss are estimated. To improve D'' imaging even further, I have constructed a generalized Radon transform approach, compensating for the liquid outer-core, which can be used to transform seismic signals passing trough the outer-core, such as *SKKS* and its precursors and coda. I apply this method to the same region as used in *ScS* studies. The image gathers computed from *SKKS* are in excellent agreement with the results (for the same image points) obtained from *ScS*. With this development

we now have a tool for detailed D'' imaging – on sub-global scale – with joint interpretation (by means of the GRT and mixed-method statistics) of the broadband ScS and $SKKS$ wavefields.

Thesis Supervisor: Robert D. van der Hilst
Title: Ceil and Ida Green Professor of Earth and Planetary Sciences

Thesis Supervisor: Maarten V. de Hoop
Title: Professor, Purdue University

Dedicated to my parents, who never pushed me too hard and never allowed me to give up too easily, to my mother-in-law, who has been supportive since the day I met her. Thank you for everything. And in memory of Changlu Li, my father-in-law, who passed away too young to attend his beloved daughter's wedding, to see his adorable grandsons. You are missed.

Acknowledgments

I am very lucky to be under the supervision of Rob van der Hilst and Martijn de Hoop, one of the best combinations ever. As one of my thesis advisors, Rob provided me not only with invaluable scientific guidance and insight, but also with career advice and encouragement. He deserves the highest praise for the close guidance which a student can dream of. I am extremely indebted to my other thesis advisor, Martijn, for his key help with the theory, his hospitality during my visits to Colorado and Indiana, and his endless patience with me in difficult moments.

I also thank all of my collaborators. I had the pleasure of working with Ping Ma when he was a post-doc at Harvard and a professor at UIUC. I enjoyed the face-to-face discussion with Luis Tenorio during my visits to Colorado School of Mines and Purdue University. Along with the scientific collaboration, Dan Shim provided a lot of help with Latex when I wrote my thesis.

I worked closely with Brad Hager and Eric Hetland for my second general project. I thank Brad for all of his help and encouragement. I thank Eric for all of his help with the MATLAB codes and geodynamical models.

I'd like to thank the rest of the professors in my general and thesis committees, Stéphane Rondenay, Nafi Töksoz, Tom Herring, and Vernon Cormier. I thank you all for the constructive suggestions for improvement and ideas for future work.

I thank Youshun Sun and his wife Naxin Zhou for their huge help at the beginning of our life in the US and lots of help and guidance afterwards. The friendship with their extended family of daughter (Abbigail Sun) made the life in the Boston area unforgettable.

The fifth floor of the Green Building has been a terrific place where I have spent six years. Frederik Simons, Keli Kárason, and Maureen Long have been great sources of answers and suggestions. Change Li, my classmate since 1992 at Peking University, has offered me various kinds of help and constructive discussion. Thank you! Huajian Yao is not only a good person to discuss the science but also a great fellow to make friends with. Einat Lev provided great ideas to improve my presentation. Thanks also to Jeffrey

Andrews-Hanna, Scott Burdick, Qin Cao, Krystle Catalli, Chin-wu Chen, Hugh Cox, James Dennedy-Frank, Kang Hyeun Ji, Sarah Stewart Johnson, Rosalee Lamm, Lisa Lassner, Sergei Lebedev, Andrea Llenos, Jiangling Lü, Erwan Mazarico, Emily Van Ark, Min Xu, and many, many others who have made the 5th floor a pleasant place to come to work.

Big thank you also goes to the great Chinese community in Green Building and ERL. Xiaojun Huang is the one who interviewed me through phone before I came to MIT. Rongrong Lu is the one who always provides unreserved help. Shichun Huang and Guangping Xu are great great people from 12th floor. Yunpeng Wang is the one who taught me how to take care of a baby. Zhengya Zhu is the one who is ready to help any time. Thanks also to Shihong Chi, Xu Li, Lingfang Ma, Xue Xiao, Xin Zhan, Yang Zhang, Yibing Zheng who are great to have parties with.

My deep gratitude goes to various support staff in the 9th, 5th, and 2nd floor. Vicki McKenna and Carol Sprague in the Education Office have helped me out with all the problems I had over years. I am not sure if my computer is the one with most problems in the 5th floor or not. But Linda Meinke and Scott Blomquist keep it running all the time. Roberta Allard and Jacqui Taylor on the 9th floor have been ready to provide any help over years. A big thank you also goes to Beth MacEachran, who made the 5th floor run smoothly. I owe a huge debt of gratitude to Joe Hankins. Thanks.

People who deserve the most credit to make this thesis possible are my family members. Great thanks to my parents, Xiangang and Qiongfang, for their continuous support and encouragement. My deep thanks also go to my two brothers Xukui and Xuliao, my best friends in the world. I also thank my in-laws for allowing me to marry their daughter. It would have been impossible for me to finish my thesis without the help from my mother-in-law to take care of my two sons.

Finally, and most importantly, let me thank the support and the love of Weiwei Li, my wife, who has sacrificed a lot during the last several years to give me the opportunity to work toward a Ph.D. There is no way I could have accomplished this task without her unwavering support. Along with my Ph.D degree, our greatest harvests over these years are our two lovely sons James Wang and Kenneth Wang. It's they who have maintained

my perspective during the last couple of years as I performed and completed the research, wrote and defended this thesis. Thanks, James and Kenneth.

This study would not have been possible without the efficient data storage at and retrieval from the Data Management Center of the Incorporated Research Institutions for Seismology (IRIS). The research presented here was funded by the Collaborative Mathematics and Geosciences program of the US National Science Foundation (grant EAR-0417891).

Contents

Chapter 1

Introduction

Earth's large scale radial stratification was a prime focus of seismology in the first half of the 20th century. In the past decades research emphasis has shifted to mapping the lateral variations in seismic propagation speed, changes in the depth to and character of deep mantle interfaces, and the boundary layers associated with thermo-chemical convection. The remote sensing of deep mantle discontinuities and, in particular, the lowermost mantle (that is, the core mantle boundary – or CMB – and D''), is a challenge because the seismic waves used to probe them propagate through Earth's heterogeneous shallow mantle before they are observed.

The rapidly increasing availability of broad-band global network data and data from dense receiver arrays, such as the *USArray* component of *EarthScope*, defines a need for the development of methodologies for automated extraction of structural signal from large data sets. To enable the detection, imaging, and characterization of singularities (including interfaces) using large data volumes, I combine concepts from inverse scattering and modern statistics into a two-step strategy. I first develop a generalized Radon transform (GRT) of broad-band global seismic network data to produce so-called 'common image-

point gathers', which reveal and characterize multi-scale variations in elastic properties at and near interfaces. Subsequently, in collaboration with my colleagues, we develop a theory for statistical inference of singularities (discontinuities).

In the rest of this chapter, I first discuss the relationship between inverse scattering and tomography, between inverse scattering and forward modeling, and between the GRT approach and other inverse scattering methods. Second, I argue that GRT is inseparable from the statistical analysis. Third, I introduce the wavefields used in this thesis. Finally, I give the outline of the thesis.

1.1 Inverse scattering

(Seismic) inverse scattering refers to a class of inverse theories used to characterize Earth's structure on scale lengths comparable to or smaller than the wavelengths of the seismic waves used as data.

1.1.1 Inverse scattering vs tomography

In current inverse scattering approaches, linearization of discontinuous perturbations (fine scale) about smooth variations (coarse scale) is required. In such a framework, smooth variations are determined with tomographic methods.

Seismic (transmission and normal mode) tomography has been successful in revealing long-period (smooth) changes in wavespeed (*Dziewonski* [1984]; *Van der Hilst et al.* [1997]; *Romanowicz* [2003]). Structures at length scales far smaller than can be resolved by tomography cause wavefield scattering, including reflections and phase conversions. The scattered wavefield has been used in many studies, for instance, to estimate stochas-

tic properties of deep mantle heterogeneity (e.g., *Hedlin et al.* [1997]; *Margerin and Nolet* [2003]), to determine variations in depth to and reflectivity of known mantle discontinuities (e.g., *Paulssen* [1988]; *Van der Lee et al.* [1996]; *Shearer and Flanagan* [1999]; *Shearer et al.* [1999]; *Gu and Dziewonski* [2002]; *Deuss and Woodhouse* [2002]; *Chambers et al.* [2005]), to explore the lowermost mantle (*Garnero* [2000]; *Castle and Van der Hilst* [2000], and many others), and to search for previously unknown interfaces (e.g., *Lay and Helmberger* [1983b]; *Revenaugh and Jordan* [1991]; *Kawakatsu and Niu* [1994]; *Vinnik et al.* [2001]; *Castle and van der Hilst* [2003]).

1.1.2 Inverse scattering vs forward modeling

Forward modeling of judiciously selected seismological phases has shown convincingly that at and near the CMB strong heterogeneity exists on a wide spectrum of length scales (e.g., *Garnero* [2000]; *Rost and Revenaugh* [2004]; *Helmberger and Ni* [2005]). However, forward modeling of complex waveforms is, as yet, only practical for relatively simple structural geometries, and it requires waveforms in which the signal of interest has sufficient amplitude. The latter often implies a restriction to fairly small epicentral distance ranges, which – in turn – restricts the CMB regions that can be studied. Moreover, one needs to have prior knowledge of the target structures. As a complementary technique, imaging by inversion of seismic data can overcome some of these limitations and can be used to explore 'Terra Incognita'. Traditional seismic inverse theory, however, allows for the correct interpretation of only a fraction of the information contained in multi-component broad-band waveform data. Moreover, noise from various sources can mask weak signals in the seismic data. Signal can be enhanced by stacking, and in recent years several exciting applications to lowermost mantle imaging have been published (e.g.,

Thomas et al. [2004]; *Hutko et al.* [2006]; *Avants et al.* [2006b]). However, these studies require laborious data analysis (including visual inspection and forward modeling at some stage of the analysis), which complicates application to large data sets and geographical regions.

1.1.3 GRT vs other inverse scattering approaches

We investigate scatterers (e.g., interfaces) in Earth's lowermost mantle (that is, the core-mantle boundary and the so called D'' region above it) with a generalized Radon transform (GRT) adapted from application to near surface (hydrocarbon reservoir) imaging. A GRT maps singly scattered waves to multiple images (or 'common image-point gathers') –for different opening or scatter angles (source-receiver distances)– of the same target structure. The theoretical analysis dates back to *Guillemin* [1985], but the first application to seismic waves is credited to *Beylkin* [1985]. This early work was done in the context of hydrocarbon reservoir imaging with acoustic waves and in the absence of caustics. Later extensions included anisotropic media (*Burridge et al.* [1998]), resolution analysis (*De Hoop et al.* [1999]), and generic elasticity with caustics (*Stolk and de Hoop* [2002]).

The GRT is a comprehensive theory/framework. Kirchhoff migration can be viewed as a special case, typically, assuming the absence of caustics; in Kirchhoff migration one often uses surface offset as the redundant variable, which leads to fundamental artifacts in image gathers in the presence of caustics. [*Revenaugh*, 1995] presents an example of classical Kirchhoff migration using all the data but loosing their sign information. Kichhoff time migration is a special case of Kirchhoff migration – it assumes straight rays in an effective medium that changes with the depth of image point (*Simon et al.* [1996]). Double beam-forming (*Scherbaum et al.* [1997]) provides the input to so-called map migration (*Hedlin*

et al. [1991]) revealing, geometrically, the propagation of singularities by the imaging operators. Applying the GRT to a single data point generates an image distribution with as its singular support an isochrone. If one has only a few isolated data points, one can overlay the corresponding isochrones to localize the scatter point(*Lay and Young* [1996]).

1.2 GRT and statistics

The generalized Radon transform (GRT) of global seismic network data in heterogeneous, anisotropic elastic media to map tens of thousands of seismograms results in a set of multiple images of the same target structure. These 'common image-point gathers' reveal multi-scale variations in elastic properties. Presumably, any (local) reflector should show up at least at (or close to) the same radius for any processed angle. However, due to the difference in coverage, quality of data for different epicentral distances and inaccuracy in background wavespeed model, this is not always the case. Therefore, instead of simply stacking the image gathers (short for 'common image-point gathers') over different scattering angles (and azimuths) into a single reflectivity profile, we use statistical methods to estimate and enhance the GRT images (**Chapter III**, *Ma et al.* [2006]). A key notion of our approach is that noise in the data and the image gathers is allowed to have mixed (that is, white and coherent) components. The correlated components can be due to, for instance, uneven source-receiver distribution, conflicting phases, multiple scattering, and the use of an inaccurate reference model for (3D) mantle wavespeed, and the parameters that control them can be estimated from the image gathers through prediction error minimization (also known as generalized cross-validation).

Along with improved estimates of the reflectivity profile (the image), it also produces

rigorous Bayesian confidence bands. These confidence bands replace forward modeling as the initial model validation tool. This allows (routine) application to very large data sets and can be used to focus the interpretation to structures that are imaged at a particular level of confidence.

1.3 About the data

To investigate the lowermost mantle, core-related phases, such as *ScS* (*PcP*) and *SKKS* (*PKKP*), and their precursors and coda are pertinent. *Rost and Revenaugh* [2004] found a strong arrival in the early coda of major-arc *PKKPab* and interpreted it as an underside reflection from D''.With this arrival, they found a D'' at 280 km above the CMB by converting the traveltime to depth. Energy stack of seismograms (Figures 2-6B and 4-3) and theoretic traveltime calculation show that the *ScS* and *SKKS* phase are rather clean from other major phases for large distance ranges (*ScS*: $0-75°$; *SKKS*: $90-180°$). In this thesis, we use *ScS* wavefield to scan the core mantle boundary (CMB) area from the topside and *SKKS* wavefield to scan from the underside [1]. If we indeed see the same structures from two totally different data sets, it is probably the best way to validate our results. Or even better, we can do a joint inversion of *ScS* and *SKKS* data to scan the CMB area from both the topside and underside. In principle, *ScS* and *SKKS* can be combined (through compensation of the outer core) to form a GRT integration over migration dip directions covering entire (unit) sphere. However, one should do the integration with caution: The integration will be sensitive to anisotropy. Conversely, we could use the integration to detect the presence of anisotropy. The use of *SKKS* has several advantages. First, with S_dS (we use

[1]The GRT approach developed in this thesis is ready for application to *PcP* and *PKKP* wavefields

the subscript d for topside reflection and superscript d for underside reflection) phase, the detection of lower interface of double crossing is relatively hard (*Flores and Lay* [2005]), whereas the amplitudes of phase SKS^dSKS are almost the same for a wavespeed increase and decrease (see Figure 4-5). Furthermore, $SKKS$ provides excellent data coverage. Since the maximum epicentral distance for ScS data is about 80°, and because there are almost no receivers and events in large intraplate regions, such as oceans, the ScS middle point coverage is very sparse there. Indeed, Central America and Eastern Eurasia are among the few regions where ScS data coverage is likely to be sufficient for successful application of the GRT with ScS data. On the other hand, the $SKKS$ middle point coverage is very good in most regions (see Figure 4-2). The main reason is that the epicentral distances used for $SKKS$ data are from 100-180°. All these features make the $SKKS$ phase an excellent complementary phase of ScS. On the other hand, to use $SKKS$ data to image the CMB area from the underside has its challenges. i) the lower limit of the earthquake magnitude which can be used to image with $SKKS$ data is higher; ii) $SKKS$ is a 'mini-max' phase and its waveform is distorted by a $\pi/2$ phase shift; iii) $SKKS$ propagates not only in the solid mantle but the liquid outer core, so that one has to compensate for the liquid outer-core; iv) For $SKKS$ imaging one has to deal with the coupled *P-SV* system. The potential mixture of SKS^dSKS and SKP^dPKS seems to make this method almost infeasible. Fortunately, one can select an epicentral distance range (100-180°) in which there is no SKP^dPKS energy because the incidence angle of K at the CMB is beyond the post-critical angle for the mantle *P*-wave (see Figure 4-1).

1.4 Outline of thesis

In **Chapter II** (published in *Journal of Geophysical Research*, (11), B12034, 2006), I construct a generalized Radon transform (for heterogeneous, anisotropic elastic media) to map broadband seismogram windows –comprising main arrivals with their coda and precursors– into multiple images of a target structure. The method is applied to the CMB area beneath Central America. In **Chapter III** (published in *Journal of Geophysical Research, in press*), we develop a theory for statistical inference of singularities (discontinuities). Several "mixed-effect" models are introduced to enhance the GRT images and provide uncertainty estimates. In **Chapter IV** (*Geophysical Journal International, in preparation*), an extended GRT approach is developed to deal with both solid-solid and liquid-liquid medium perturbations. I compare the results by this method with those by the method in **Chapter II**. In **Chapter V** (adapted from a published paper in *Science*, (315), 1813-1817, 2007), I interpret our images in terms of mineral physics and geodynamics. The discussion of this thesis is given in **Chapter VI**, where the key results of this research work are summarized and future work related to this research is discussed.

Chapter 2

Imaging of structure at and near the core mantle boundary using a generalized Radon transform: I- construction of image gathers[†]

Abstract

We introduce a new method for imaging heterogeneity at and near interfaces in Earth's lowermost mantle with broadband, three-component seismograms from global seismograph networks. Our approach is based on inverse scattering and allows the extraction of pertinent signal from large data sets and requires few *a priori* assumptions about the heterogeneity under study, which makes it complementary to the forward modeling of selected waveforms. Here (Paper I) we construct a generalized Radon transform (for heteroge-

[†]Published as: Imaging of structure at and near the core mantle boundary using a generalized Radon transform: I- construction of image gathers, *J. Geophys. Res.*, 111, B1230, doi:10.1029/2005JB004241, 2006.

neous, anisotropic elastic media) to map broadband seismogram windows – comprising main arrivals with their coda and precursors – into multiple images of a target structure. The 'common image-point gathers' thus produced reveal multi-scale variations in elastic properties near deep interfaces. The GRT can be applied to narrow and wide angle data, and the (automated) extraction of signal from data over a wide range of epicentral distances enables exploration of CMB regions that cannot – with present-day data coverage – be imaged with the triplicated waveforms used in forward modeling studies. Tests with synthetic data, produced both with idealized and actual source-receiver distributions, illustrate pertinent aspects of the theory and show that (multiple) weak interfaces can be detected and located correctly, even in the presence of (random) noise that would prohibit visual inspection and modeling of the subtle signals. We transformed ~100,000 transverse-component *ScS* waveforms into image gathers of a core mantle boundary (CMB) patch beneath Central America. Juxtaposition of stacks of these gathers produces a 2-D image profile revealing contrasts in elasticity near the target depth of the CMB and ~280 km above it. The latter may mark the top of the so called D'' region. The images also reveal a richness of structures in between these depths. Combined with a statistical analysis of the significance of these singularities (described in Paper II, **Chapter III**), the approach to imaging presented here paves the way to large-scale seismic exploration of the lowermost mantle.

2.1 Introduction

Earth's large scale radial stratification was a prime focus of seismology in the first half of the 20th century. In the past decades research emphasis has shifted to mapping the lateral variations in seismic propagation speed and changes in the depth to and character of deep mantle interfaces and boundary layers associated with thermo-chemical convection. The relatively smooth variations in seismic wavespeed can be delineated by transmission and normal mode tomography, e.g., *Dziewonski* [1984]; *Van der Hilst et al.* [1997]; *Romanowicz* [2003]. Structures at length scales far smaller than can be resolved by tomography cause wavefield scattering, including reflections and phase conversion. Scattering of seismic waves has been used, for instance, to estimate stochastic properties of deep mantle heterogeneity (e.g., *Hedlin et al.* [1997]; *Margerin and Nolet* [2003]), to determine variations in depth to and reflectivity of known mantle discontinuities (e.g., *Paulssen* [1988];

Van der Lee et al. [1996]; *Shearer and Flanagan* [1999]; *Shearer et al.* [1999]; *Gu and Dziewonski* [2002]; *Deuss and Woodhouse* [2002]; *Chambers et al.* [2005]), to explore the lowermost mantle (*Garnero* [2000]; *Castle and Van der Hilst* [2000], and many others), and to search for previously unknown interfaces (e.g., *Lay and Helmberger* [1983b]; *Revenaugh and Jordan* [1991]; *Kawakatsu and Niu* [1994]; *Vinnik et al.* [2001]; *Castle and van der Hilst* [2003]).

The remote sensing of deep mantle discontinuities and, in particular, the lowermost mantle (that is, the core mantle boundary – or CMB – and D''), is a challenge because the seismic waves used to probe them propagate through Earth's heterogeneous shallow mantle before they are observed. Forward modeling of judiciously selected seismological phases has shown convincingly that at and near the CMB strong heterogeneity exists on a wide spectrum of length scales (e.g., *Garnero* [2000]; *Rost and Revenaugh* [2004]; *Helmberger and Ni* [2005]). However, forward modeling of complex waveforms is, as yet, only practical for relatively simple structural geometries, and it requires waveforms in which the signal of interest has sufficient amplitude. The latter often implies a restriction to fairly small epicentral distance ranges, which – in turn – restricts the CMB regions that can be studied. Moreover, one needs to have prior knowledge of the target structures. As a complementary technique, imaging by inversion of seismic data can overcome some of these limitations and can be used to explore 'Terra Incognita'. Traditional seismic inverse theory, however, allows for the correct interpretation of only a fraction of the information contained in multi-component broad-band waveform data. Moreover, noise from various sources can mask weak signals in the seismic data. Signal can be enhanced by stacking, and in recent years several exciting applications to lowermost mantle imaging have been published (e.g., *Thomas et al.* [2004]; *Hutko et al.* [2006]; *Avants et al.* [2006b]). However,

these studies require laborious data analysis (including visual inspection and forward modeling at some stage of the analysis), which complicates application to large data sets and geographical regions.

With the rapidly increasing availability of broad-band global network data and data from dense receiver arrays, such as the *USArray* component of *EarthScope*, there is a well recognized need for the development of methodologies for automated extraction of structural signal from large data sets. To enable the detection, imaging, and characterization of singularities (including interfaces) using large data volumes, we combine concepts from inverse scattering and modern statistics into a two-step strategy. The first step, presented here, is the development of a generalized Radon transform (GRT) of broad-band global seismic network data to produce so-called 'common image-point gathers', which reveal and characterize multi-scale variations in acousto-elastic properties at and near interfaces. In a companion paper, hereinafter referred to as Paper II, *Ma et al.* [2006] analyze these gathers using 'mixed-effects' statistical models. The statistical analysis is used to enhance the images and estimate formal (Bayesian) confidence levels. In our automated imaging of Earth's deep interior, the latter replaces forward modeling of (stacks of) waveforms as the (initial) means for model or image validation.

We develop a GRT for imaging of the lowermost mantle with the broad-band wavefield formed by direct *ScS* and its precursors and coda. In Section 2.2 we develop the theory underlying the GRT and describe how three-component broad-band data can be transformed into image gathers. The GRT was introduced to seismic imaging by *Beylkin* [1985] and *Miller et al.* [1987] but the development followed here builds on generalisations due to De Hoop and collaborators (e.g., *De Hoop et al.* [1994]; *De Hoop and Bleistein* [1997]; *Burridge et al.* [1998]). In Section 2.3 we use synthetic seimograms to illustrate key aspects

and test the performance of our methodology. We show that – in principle – the GRT can be used to detect (multiple) deep Earth interfaces and estimate their reflection coefficients. We also analyze how the radial resolution depends on the scattering angle. This dependence, which is shown to be related to what is called the dilation in a wavelet transform, is of key importance for the space-scale characterization of the interface. Furthermore, we demonstrate that the GRT is effective in suppressing (random) noise. Finally, in Section 2.4 we discuss pertinent aspects of the data processing, including the use of principal component analysis (PCA) for identifying and separating the direct and scattered wavefields, and we present preliminary results of our study of a CMB patch beneath Central America.

2.2 GRT to 'uniform' common image-point gathers

2.2.1 Historical perspective (anisotropic elastic case)

There have been many publications about high-frequency methods to invert seismic data in *acoustic* media. These methods date back to *Hagedoorn* [1954]; from a seismic perspective, it has taken thirty years to develop the basic analysis (*Schneider* [1978]; *Clayton and Stolt* [1981]; *Stolt and Weglein* [1985]; *Miller et al.* [1987]; *Schleicher et al.* [1993]). From a mathematical perspective, the analysis started with the reconstruction of the singular component of the medium coefficients in the Born approximation, in the absence of caustics, by *Beylkin* [1985] – using the framework of generalized Radon transforms. *Bleistein* [1987] discussed the case of a smooth jump using Beylkin's results. The simplest form of an asymptotic inversion procedure, however, can already be found in *Norton and Linzer* [1981].

Beylkin and Burridge [1990] discussed the asymptotic imaging of seismic data in the Born approximation in isotropic *elastic* media, under a no-caustics assumption. The generalized Radon transform in *anisotropic* elastic media was developed by De Hoop and coworkers (*De Hoop et al.* [1994]; *Burridge et al.* [1998]; *De Hoop et al.* [1999]). *De Hoop and Bleistein* [1997] introduced the imaging and inversion in general anisotropic elastic media, using a Kirchhoff-type approximation.

Guillemin [1985] discussed the so-called Bolker condition in the context of generalized Radon transforms, which ensures invertibility of the modeling or single scattering operator in the least-squares sense. *Stolk and de Hoop* [2002] made use of this result in the development and analysis of the generalized Radon transform in anisotropic elastic media allowing the presence of caustics; explicit expressions and algorithmic aspects in this case can be found in *De Hoop and Brandsberg-Dahl* [2000]. The foundations of the use of the Kirchhoff approximation in the generalized Radon transform, in the presence of caustics, were also given by *Stolk and de Hoop* [2002]. An implementation and application of these results to exploration seismic data can be found in *Brandsberg-Dahl et al.* [2003].

In global seismology, applications and adaptations of the no-caustic isotropic elastic generalized Radon transform to scattered teleseismic body waves can be found in *Bostock et al.* [2001] and *Poppeliers and Pavlis* [2003]. Poststack migration in the context of receiver functions was discussed by *Rydberg and Weber* [2000].

2.2.2 GRT imaging of deep mantle interfaces

In essence, the GRT enables the transformation of a large number of broad-band seismic waveform data into (multiple) images of a singularity in physical medium properties, for instance a deep mantle interface. In order to do so, however, one has to account carefully

for (smooth, possibly anisotropic) 3-D wavespeed variations in the background medium, geometrical aspects such as the focal depth and the radiation patterns of the earthquakes considered, and the uneven and sparse sampling (that is, acquisition geometry). Moreover, in applications to earthquake data (that is, passive seismics) one has to estimate and remove (for each earthquake) the source signature and (for each receiver) the instrument response.

In the subsections below we introduce the notation regarding the sources and receivers, define the ray-geometrical aspects, and develop the transform itself. We explain the multi-resolution aspects of 'common image-point gathers', and we discuss (anti-)aliasing and other issues related to uneven spatial and spectral sampling.

Sources and receivers – definitions

We consider waveform data from many sources (earthquakes), indicated by superscript s, recorded at many receivers (seismograph stations), indicated by superscript r. The stations are not required to be part of a contiguous, geographically restricted array.

The earthquake epicenters are denoted by x^s. We assume that the origin times t^s are reset to zero, and that the earthquake's time-rise function is deconvolved from the data (the related practicalities are discussed in Section 2.4.2). For each earthquake we write the equivalent body force f as

$$f_j(x,t) = -M_{ij}\partial_i\delta(x-x^s)\,H(t-t^s), \text{with } t^s \text{ set to } 0, \tag{2.1}$$

where M_{ij} is the moment tensor. Note that we use the subscript summation convention. The receivers, located at x^r, record three displacement components $u_p(x^s,x^r,t)$, $p=1,2,3$, which – after pre-processing – will be used as input in the GRT.

Scattering geometry – definitions

The geometry considered in transforming seismic data in a heterogeneous, anisotropic elastic medium is illustrated in Figure 2-1 (top), where the image point is denoted by $y = (y_1, y_2, y_3)$. The superscripts s and r indicate the association with a ray from a source and a receiver, respectively. The 'two-way' travel time for a particular diffraction branch associated with a ray path connecting x^s with x^r via y is denoted by $T = T(x^s, x^r, y)$.

The slowness vector of the ray connecting x^s with y (evaluated at y) is given by $p^s(y)$; in particular, $p^s(x^s)$ indicates the slowness along this ray evaluated at the source. The projection $\pi^s(x^s)$ of $p^s(x^s)$ onto Earth's surface is a horizontal slowness. Furthermore, we introduce the phase direction (that is, a unit vector normal to the wavefront) $\hat{\mathbf{k}}^s = p^s/|p^s|$ and, following the notation by *Červený* [2001], the phase velocity $V^s = 1/|p^s|$. A similar notation is used for the quantities along the ray connecting the image point with the receiver (that is, $p^r(y)$, $p^r(x^r)$, $\pi^r(x^r)$, $\hat{\mathbf{k}}^r$, and V^r). With ω the angular frequency, $\omega\pi^s$ and $\omega\pi^r$ are (horizontal components of) wave vectors. (We note that $(s, r, t, \omega\pi^s, \omega\pi^r, \omega)$ – that is, space, time, and their Fourier duals, wave vector and frequency – defines a point in data phase space.) Likewise, the polarization vector, h, associated with compressional- or shear waves, is defined at the source, receiver, and image point.

The quantity that controls image resolution is what we will call the migration slowness vector, $p^m(y) = p^s(y) + p^r(y)$, with a direction – known as the *migration dip* in the exploration literature – $\nu^m(y) = p^m(y)/|p^m(y)|$. (We note that $(y, p^m(y))$ defines a point in image phase space.) Together, the migration dip and the phase directions of incoming and scattered rays define the scattering vector

$$\psi = (\hat{\mathbf{k}}^s \times \hat{\mathbf{k}}^r) \times \nu^m \tag{2.2}$$

at image point y.

For a travel-time diffraction branch, and away from caustics at x^r or x^s, the scattering angle θ between incoming and scattered rays at y is related to the scattering vector as

$$\sin\theta = |\psi|, \quad \theta = \theta(x^s, x^r, y), \tag{2.3}$$

and the scattering azimuth ψ is the angular displacement of the scattering vector at y,

$$\psi = \psi(x^s, x^r, y), \tag{2.4}$$

normalized to one (that is, $\psi/|\psi|$); see, again, Figure 2-1(top).

2.2.3 Map (de)migration and isochrons

Map migration describes how the geometry of a (specular) reflection – defined as a combination of source and receiver coordinates, travel time, and horizontal slownesses – is mapped to the location and orientation of a reflector:

$$\Sigma : (x^s, x^r, t, \pi^s, \pi^r) \mapsto (y, p^m) \quad \text{at} \quad t = T(x^s, x^r, y), \tag{2.5}$$

see, for instance, *Kleyn* [1977]; *Douma and de Hoop* [2006].

For given (x^s, x^r, t), the set of equal times $T(x^s, x^r, y) = t$ defines an isochron. With slownesses π^s and π^r, which can be inferred from the data (for instance from travel time slopes, polarization analysis, or vespegrams, see *Rost and Thomas* [2002]), the mapping Σ locates y on an isochron. If such slowness information is not used, a data point (x^s, x^r, t) smears over an isochron in the interior of the Earth, see Figure 2-1 (bottom). In fact,

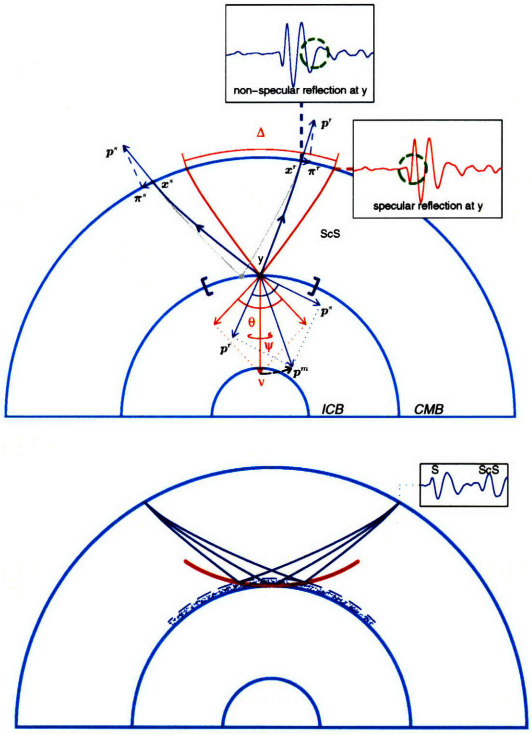

Figure 2-1: Schematic illustration of the path geometry (top) and the isochron concept (bottom) considered in the generalized Radon transform (GRT) of *ScS* data. Top: The source (x^s) and receiver (x^r) are separated by epicentral distance Δ. The image point at the CMB is denoted y. Slowness vectors are given by p, and π denote horizontal slownesses. The scattering angle is θ and scattering azimuth is ψ. The image is, essentially, created by integration over p^m. All other symbols are in the text. The two seismograms illustrate that information about a predescribed image point y is gleaned from different parts of data recorded at the different stations; for non-specular reflections part of the coda contributes to the stacks, whereas for specular reflections the information is retrieved from the main arrival. Bottom: For given (x^s, x^r, t), the set of points y constrained by $T(x^s, x^r, y) = t$ is identified as an isochron; p^m is normal to the isochron.

the impulse response of the kernel of the generalized Radon transform (considered as an integral operator) coincides precisely with an isochron.

The use of isochrons for deep Earth imaging is not new. For example, *Lay and Young* [1996] used them in their study of scattering in Earth's lowermost mantle. Isolating a proper time window of the coda wave (such as *S* to *ScS*), they applied a convolutional model approach to estimate for any scattered waves the arrival times t (at a station located at x^r for a given earthquake at x^s) and the amplitudes. Thus a combination $(x^s, x^r, t, \text{amplitude})$

is a data point, and multiple data points smear over a collection of isochrons. Instead of integrating over all isochrons at a particular image point, as done by the generalized Radon transform, *Lay and Young* [1996] considered a smooth mantle wavespeed model and kept track of how many isochrons associated with their data points hit a particular scattering (image) point in the lower mantle.

2.2.4 Generalized Radon transform inversion

With the geometrical concepts developed above we derive the basic form of the operators that transform the waveform data to a set of common image-point gathers in a heterogeneous, anisotropic medium. The elastic properties of the medium under consideration, here Earth's mantle and crust, are described by a stiffness tensor c_{ijkl} $(i,j,k,l \in \{1,..,3\})$ and mass density ρ. These parameters are decomposed as a sum of a smooth part (with superscript (0)) and a (non-smooth) perturbation (superscript (1)):

$$\rho(x) = \rho^{(0)}(x) + \rho^{(1)}(x), \quad c_{ijkl}(x) = c^{(0)}_{ijkl}(x) + c^{(1)}_{ijkl}(x). \tag{2.6}$$

Accordingly, the (singly scattered) part of the displacement field associated with the perturbed medium properties is denoted with superscript (1), so that $u = u^{(0)} + u^{(1)}$. For now we assume that an estimate of the smooth wavespeed variations (the background model) is available, for instance from tomography. For a given background model, the medium perturbations, which contain the discontinuities and other types of scatterer, are then found by imaging (or inverse scattering) through application of the GRT.

We further assume that the perturbations are (non-smooth) changes in elastic parameters across a local, laterally contiguous interface defined by a specific value of some func-

tion ϕ (*De Hoop and Bleistein* [1997])[1]. Such a function would describe, for example, the topography of the interface. The interface normal is given by $\nu_\phi = \nabla\phi/|\nabla\phi|$. Multiple interfaces are simply treated as a combination of such functions.

We can now formulate the migration of waveform data to uniform image gathers at a predescribed (common) image point y. These image gathers will be inferred from $\hat{S}_w(y;\theta,\psi)$, which in turn is obtained from the pre-processed waveform data $\tilde{u}^{(1)}$ through a GRT of the following form[2]:

$$\hat{S}_w(y;\theta,\psi) = \int_{E_{\nu^m}} \tilde{u}^{(1)}(x^s,x^r,y)\,|p^m(y)|^3 d\nu^m. \tag{2.7}$$

As illustrated in Figure 2-1, the source and receiver positions are here explicit functions of the image point, migration dip, scatter angle, and azimuth; that is, $x^s = x^s(y,\nu^m,\theta,\psi)$ and $x^r = x^r(y,\nu^m,\theta,\psi)$; these positions can be determined by ray tracing from image point y upward until the rays intersect Earth's surface at the source and receiver side. The integration over migration dip ν^m is restricted to E_{ν^m}, which depends on (θ,ψ) and reflects the effect of the acquisition imprint on the final image. In (2.7), $\tilde{u}^{(1)}$ represents the waveform data of the singly scattered constituent $\tilde{u}_p^{(1)}$ (here *ScS* and its precursors and coda) corrected for amplitude, polarization, phase, and travel time at y [*Burridge et al.*, 1998, (4.2)]:

$$\tilde{u}^{(1)}(x^s,x^r,y) = \overbrace{W(x^s,x^r,y)}^{\text{source}}\ \overbrace{h_p^r(x^r)}^{\text{polarization}}\ \overbrace{\partial_t \tilde{u}_p^{(1)}(x^s,x^r,T(x^s,x^r,y))}^{\text{waveform data}}$$
$$\cdot 2\underbrace{[\rho^{(0)}(x^r)V^r(x^r)V^r(y)V^s(y)\rho^{(0)}(x^s)V^s(x^s)]^{1/2}}_{\text{weights related to Green's functions}}\underbrace{[\det\mathbf{Q}_2(x^r,y)\det\mathbf{Q}_2(y,x^s)]^{1/2}}_{\text{geometrical spreading}} \tag{2.8}$$

[1]Mathematically, the jump function can be readily replaced by an element of a Zygmund class of order between 0 (i.e., a step function) and 1 (i.e., a ramp function), which also determines the local scaling properties of the singularity. This generalization is important, for example, in investigations of phase transitions.

[2]This expression is a stripped down version of equation (20) in *Brandsberg-Dahl et al.* [2003]; for instance, the radiation-pattern inversion has been removed to enable the direct estimation of a single reflection coefficient instead of multiple combinations of stiffnesses.

Here, the contraction with $h_p^r(x^r)$ represents rotations of the receiver components, and

$$W(x^s,x^r,y) = \tfrac{1}{2}M_{qr}(x^s)\,(h_q^s(x^s)p_r^s(x^s) + h_r^s(x^s)p_q^s(x^s)) \tag{2.9}$$

accounts for the different source mechanisms. Furthermore, $\tilde{u}_p^{(1)}$ represent the original, singly-scattered data $u_p^{(1)}$ corrected for possible phase shifts due to caustics:

$$\tilde{u}_p^{(1)}(x^s,x^r,T(x^s,x^r,y)) = \mathscr{H}^{\sigma(x^r,y,x^s)}u_p^{(1)}(x^s,x^r,T(x^s,x^r,y)), \tag{2.10}$$

with $\mathscr{H}$ the Hilbert transform and $\sigma(x^r,y,x^s) = \kappa(x^r,y) + \kappa(y,x^s)$ is the accumulated KMAH index (*Červený* [2001]) that keeps track of caustics that occur between x^r and y and between y and x^s. Such caustics readily appear in heterogeneous (but smooth) media. Finally, $\mathbf{Q}_2(x^r,y)$ and $\mathbf{Q}_2(y,x^s)$ are the relative geometrical spreading (*Červený* [2001]) for the receiver and source rays, respectively. All other parameters are as defined in Section 2.2.2.

Based on [*De Hoop and Bleistein*, 1997, equations (37)-(38)] the GRT in (2.7) is designed to reconstruct a combination of a singular function that characterizes some interface and a smooth (amplitude) function $S^{(1)}$ that represents the associated scattering coefficient,

$$\underbrace{S^{(1)}(y;\theta,\psi)}_{\text{smooth amplitude}}\ \underbrace{(\nu_\phi\cdot p^m)^{-1}\,|\nabla_y\phi|\,\delta(\phi(y))}_{\text{interface characterization}},$$

with $S^{(1)}$ strictly defined only for y on the interface defined as a zero level set of a function ϕ; the normal to the interface at point y is given by $\nu_\phi = \nabla_y\phi/|\nabla_y\phi|$. We note that if the singularity represents a jump (that is, a first-order discontinuity) the scattering coefficient represents the reflection coefficient.

For given x^s and x^r, let x^ϕ be the specular reflection point with associated interface dip $\nu_\phi(x^\phi)$. For y in the vicinity of this specular reflection point, p^m aligns with ν_ϕ. Using the first-order Taylor expansion of $\phi(y)$ about x^ϕ while noting that $\phi(x^\phi) = 0$, that is, $\phi(y) \simeq \nabla_x\phi_{x^\phi} \cdot (y - x^\phi)$, it follows that the interface characterization can be written as

$$(\nu_\phi \cdot p^m)^{-1} \, |\nabla_y \phi| \, \delta(\phi(y)) \simeq (p^\phi)^{-1} \, |\nabla_x \phi|_{x^\phi} \, \delta(\nabla_x \phi|_{x^\phi} \cdot (y - x^\phi)) = (p^\phi)^{-1} \delta(\nu_\phi \cdot (y - x^\phi)),$$

where

$$p^\phi = |p^m(x^\phi)|, \quad \text{such that} \quad \nu^m(x^\phi) = \nu_\phi(x^\phi). \tag{2.11}$$

Using the homogeneity of the delta function (easily checked in its Fourier representation), the interface can then be characterized as

$$(p^\phi)^{-1} \delta(\nu_\phi \cdot (y - x^\phi)) = \delta(p^\phi \nu_\phi \cdot (y - x^\phi)),$$

and the assumption of a (common) source signature leads to the factorization

$$\hat{S}_w(y; \theta, \psi) = \hat{S}^{(1)}(x^\phi; \theta, \psi) \, w_{(\theta,\psi)}(p^\phi \nu_\phi(\theta, \psi) \cdot (y - x^\phi)), \text{ for } |y - x^\phi| \text{ small}, \tag{2.12}$$

where $w_{(\theta,\psi)}$ are smooth functions that reveal the imprint at x^ϕ of the source-receiver distribution [*De Hoop and Bleistein*, 1997, equation (94)]. We recognize in (2.12) a dilation (scaling with $1/p^\phi$) and a translation (by x^ϕ). Hence, for given direction ν_ϕ, $\hat{S}_w(y; \theta, \psi)$ behaves like a *wavelet transform* of the singularity being imaged at y. We will see later (Section 2.3) that the dilation plays an important role in characterizing the (radial) scaling properties of the singularities being imaged. In(2.12), $\hat{S}^{(1)}(x^\phi; \theta, \psi)$ denotes the estimate

of the scattering coefficient $S^{(1)}(x^{\phi};\theta,\psi)$ (see Appendix).

By adjusting the weighting functions inside the integral of (2.7) we change the GRT given by $\hat{S}_w(y;\theta,\psi)$ to a mapping $\mathscr{I}(y;\theta,\psi)$ of $\tilde{u}^{(1)}$ to (approximately) uniform image gathers (IGs) in scattering angle θ and azimuth ψ (cf. (2.18), Appendix):

$$\mathscr{I}(y;\theta,\psi) := \int_{E_{v^m}} \frac{\tilde{u}^{(1)}(x^s,x^r,y)}{W(x^s,x^r,y)2} \frac{|p^m(y)|^3}{|\mathbf{w}(x^s,x^r,y)|} dv^m.$$

With this result we derive the structural image through integration over θ and ψ:

$$\mathscr{I}(y) = \iint \mathscr{I}(y;\theta,\psi)\, d\theta d\psi. \tag{2.13}$$

In the presence of caustics, $\mathscr{I}(y;\theta,\psi)$ commonly generates false image events while the stack $\mathscr{I}(y)$ over (θ,ψ) does not (see *Stolk and de Hoop* [2004] for details). In Paper II this linear stack will be replaced by an integration over ψ followed by formal statistical inference of singularities in the gather.

2.2.5 Sampling

Finally, in the construction and subsequent statistical analysis of the image gathers we have to understand the effects of sparse and uneven sampling. Typically, the global wavefield is sampled irregularly in x^s and x^r but regularly in t. By itself, irregular spatial sampling is an advantage for our approach; in fact, regularly sampled data from regional arrays should be subjected to aliasing tests. Irregular source-receiver sampling obtained by quasi-Monte Carlo sampling of migration dip and scattering angle and azimuth even results in optimal spatial resolution kernels for inverse scattering for a given number of data points (*De Hoop*

and Spencer [1996]).

The sampling properties of the GRT can be described using a table generated by map migration Σ (cf. (2.5))

$$\{(x^s, x^r, t, \omega\pi^s, \omega\pi^r, \omega; x, \omega p^m) \mid t = T(x^s, x^r, x)\}.$$

The braces, here, indicate that this set of points is a relation; note that, in general, there is *not* a mapping between $(x, \omega p^m)$ and $(x^s, x^r, t, \omega\pi^s, \omega\pi^r, \omega)$.

For a given grid, the Fourier duals in this table, that is, $(x^s, \omega\pi^s)$, $(x^r, \omega\pi^r)$, (t, ω) and $(x, \omega p^m)$, should satisfy the Nyquist criterion, which essentially provides an upper bound for frequency, $f_{\max} = \frac{\omega_{\max}}{2\pi}$. For example, if Δr denotes the (average) station spacing on a gridded array, then

$$f_{\max}|\pi^r| \leq \frac{1}{2\Delta r}.$$

If needed, an anti-alias filter (for a design, see *Lumley et al.* [2001]) can be applied. For sampling the image, we note that the magnitude of the (migration) slowness vector is given by

$$|p^m| = [(V^s)^{-2} + 2(V^s)^{-1}(V^r)^{-1}\cos(\theta) + (V^r)^{-2}]^{1/2}, \tag{2.14}$$

which also shows how $|p^m|$ changes with scattering angle θ. In practice, spatial aliasing in the image is not an issue because its sampling is part of the (computational) algorithm.

2.3 Resolution tests with synthetic data

We evaluate the performance of the methodology developed above with broad-band WKBJ seismograms (*Chapman* [1978]) calculated from a radially stratified wavespeed model

(*ak135*, *Kennett et al.* [1995]), on which we superimpose jumps in elastic parameters at certain distances above the CMB. For example, Figure 2-2 (left) depicts seismograms calculated for a 1.5% S-wavespeed contrast at 150 km above the CMB; the records on the right are generated with contrasts at 150, 200, and 250 km above the CMB. Tests such as the ones presented here demonstrate that the GRT can detect small medium contrasts in noise-free data or if the noise in the data is random and white. Other types of noise can degrade the GRT stacks, however, and in Paper II (**Chapter III**, *Ma et al.* [2006]) we assess the performance of the GRT in the presence of non-random 'noise' (e.g., due to uneven data coverage, errors in the assumed background medium, or multiple scattering) and discuss how the GRT stacks can be enhanced through statistical inference (with mixed-effects models).

We consider different geographical source-receiver distributions. In one series of tests we use an idealized geographic distribution of source-receiver pairs (Figure 2-3A,B); in another we use the actual earthquake-station distribution (Figure 2-3C). We show results for (synthetic) data bandpass filtered between 1-10 s.

For the wavespeed models and associated ray geometries considered we can calculate the reflection coefficient R as a function of scatter angle. The theoretical curve (Figure 2-3B) and the synthetic data (Figure 2-2B) suggest that three angles are of particular interest. (1) The intramission angle i_i, at which no energy is reflected (that is, $R = 0$). In accord with theoretical predictions, Figure 2-3A shows that the amplitude of S_dS decreases with increasing opening angle (with $R < 0$) for $i < i_i$, becomes zero at $i = i_i$, and increases again for $i > i_i$, but with opposite polarity (that is, $R > 0$). (2) The cross-over angle i_x, beyond which ScS, the CMB reflection, arrives before S_dS, the reflection off the shallower interface (which happens because of the imposed increase in wavespeed in the lowermost layer). (3)

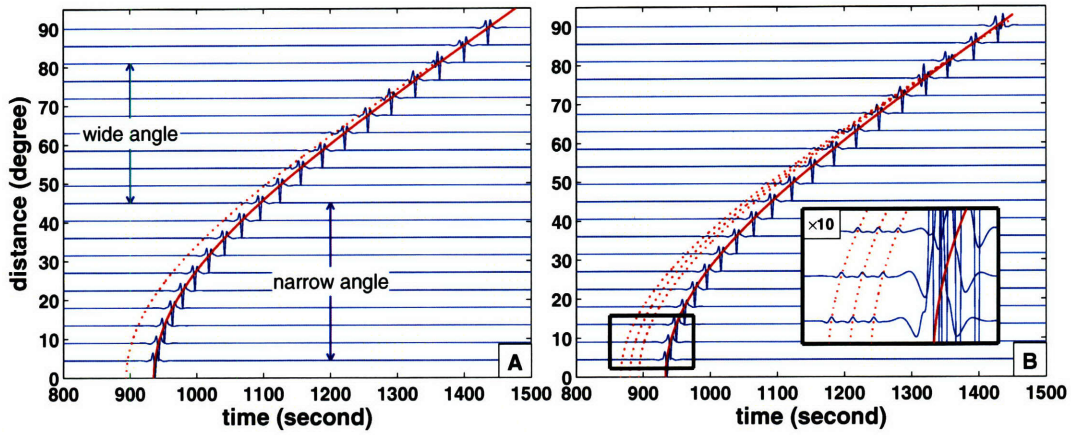

Figure 2-2: Record section of synthetic data for models with one (left) and three (right) contrasts above the CMB, calculated with WKBJ. The red solid lines are the travel time curves of ScS phase and the red dotted lines are the travel time curves of S_dS phases. At wide angles this reflection becomes stronger and crosses over with, and eventually arrives in the coda of direct ScS. The inset in B shows, for narrow angle reflections, the weak precursory energy (amplified). Narrow angle (0-i_i); Wide angle (i_i-i_c), with i_i and i_c the intramission and critical angle, respectively. For the parameters used here, i_i=44.6° and the critical angle for S_dS is at 80.6°; the cross-over between S_dS and ScS occurs at 83.5°.

The critical angle i_c, beyond which no energy is transmitted into the D'' (head wave). We define narrow angles as i less than i_i and wide angles for i between i_i and i_x (or i_c, if $i_c < i_x$). In Figure 2-3B we compare the magnitude of the reflection coefficient inferred from the GRT (see Appendix) and from Snell's law. The inferred reflection coefficient matches the theoretical curve remarkably well, except near and beyond the critical angle (which in this case is $2\times \sim 80.6°$).

In addition to illustrating how i_i, i_x, and i_c affect the appearance of the angle gathers, Figure 2-3A,C demonstrates that the width of the reflectors increases (dilates) with increasing scatter angle. This dilation – see (2.12) and the text below it – depends on θ as $1/\cos(\theta/2)$. The degradation of radial resolution with increasing distance can be understood from simple ray geometrical considerations: with increasing angle of incidence the vertical slowness decreases and the travel time becomes less sensitive to perturbations in

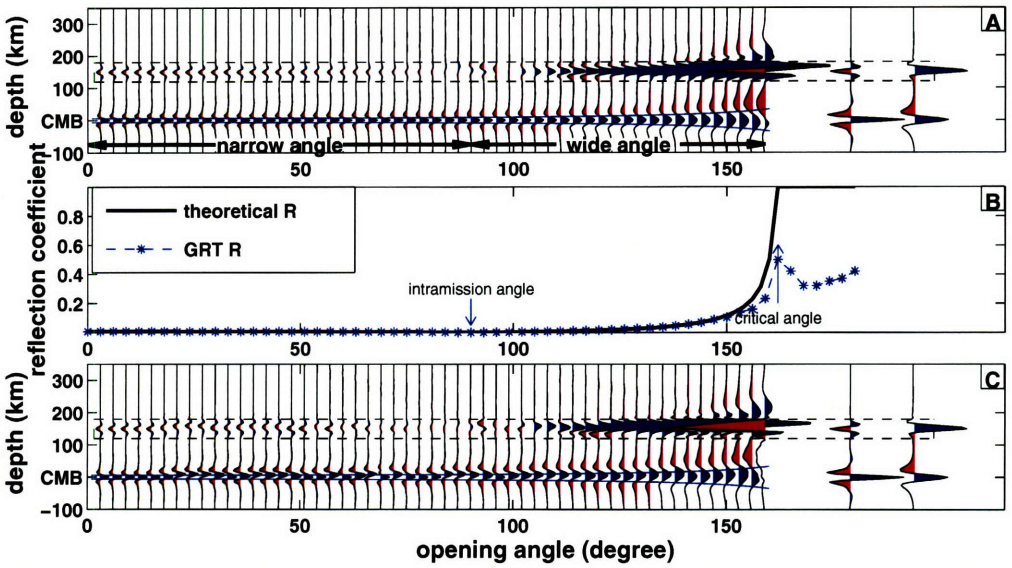

Figure 2-3: Illustration of the construction of GRT stacks (images) from image gathers at different scatter angles. The traces on the left of panels (A) and (C) are image gathers at 53 scattering angles produced from the synthetic data as in Figure 2-2. The traces on the right are stacks over narrow and wide angles (as defined in the text and in the caption to Figure 2-2). We integrate over narrow and wide angles seperately because of the change in polarity upon crossing the intramission angle i_i; stacking over all angles would involve signals with opposite polarities and cold thus mask interfaces. We use a 1-10s bandpass filter. The gathers and stacks in (A) are produced from an artificial (regular) source-receiver distribution; the results in (C) were computed using the data coverage depicted in Figure 2-6A. In (A) and (C) the dilation shows up as $1/\cos(\theta/2)$ – the theoretical values are depicted by the thin blue lines around the depth of the CMB. To aid visual inspection, the amplitude in the (dashed) box in (A) and (C) is amplified by a factor of 20. In (B), the solid line depicts the reflection coefficient calculated from the input model and the star is the reflection coefficient picked up by our GRT method (see Appendix). The intramission angle $i_i = 44.6\,^\circ$ and the critical angle $i_c = 80.6\,^\circ$.

discontinuity depth (in the limit of grazing rays the sensitivity vanishes because at the image point the depth perturbation is perpendicular to the ray). Most forward modeling studies consider (triplicated) waveforms at distances larger than $\sim 75^\circ$ (which is at the large distance end of what we call wide angle data) because the S_dS phase is weak at smaller distances. This leaves a fairly small epicentral distance range that can be used. Moreover, these data have relatively poor resolution to the depth of interfaces. For our purposes, however, the (predictable) variation in depth resolution provides valuable information; indeed,

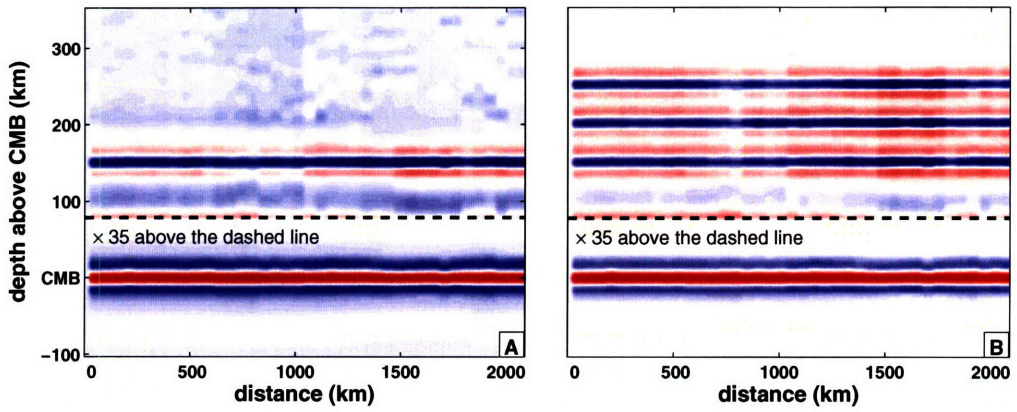

Figure 2-4: Images obtained from narrow angle stacks produced from synthetic data, illustrating the resolution of the GRT. (A) The recovery of the input model with a contrast at 150 km above the CMB. (B) The recovery of the input model with contrasts at 150 km, 200 km, and 250 km above the CMB. The amplitudes above the dashed line are multiplied by a factor of 35 to make them comparable to those of the CMB.

the redundancy contained in the narrow and wide angle data helps us constrain both the fine scale and coarser scale radial variations in elastic properties. This property can be exploited to quantify the space-scale properties of the singularities being imaged.

The ability to reproduce the dilation ($1/p^{\phi}$ in (2.12)) and the reflection coefficients confirms the correct behavior of our GRT. We can also demonstrate that our GRT is able to detect multiple interfaces. For this purpose, we chose 41 imaging points along the great circle transect from (-105W, 0) to (-75W, 30N). Figure 2-4 was generated by the lateral juxtaposition of image gathers stacked over narrow angles. Using the synthetic data in Figure 2-2A,B, the images depicted in Figure 2-4A,B demonstrate that the CMB and multiple interfaces within D'' are well recovered by the GRT.

Finally, we demonstrate that random noise in the data is suppressed effectively by the GRT. For this test we add noise to data generated from a model with, as before, a CMB and a wavespeed increase at 150 km above it (Figure 2-5). In Figure 2-5, top right, we show the image obtained by applying the GRT to the data without noise (shown in top left panel).

The CMB as well as the shallower reflector are resolved at depths that correspond to the contrast in wavespeed in the background model used. Next, we perform the GRT to the data contaminated with noise; the result is shown in Figure 2-5, bottom right. The result is practically the same as that of the noise-free data case. Tests like this demonstrate that the GRT is robust under the addition of white random noise. Even if the signal-to-noise is so low that it is impossible to see the signal from the 'top' reflector in the raw data (see inset in Figure 2-5, bottom left), the GRT still yields the contrasts at the right position because it makes use of the redundancy in the data.

2.4 Imaging the CMB beneath Central America

We apply the GRT to a broadband wavefield formed by *ScS* (and its precursors and coda) that sample a $50^\circ \times 50^\circ$ core mantle boundary (CMB) beneath Central America (Figure 2-6A). This region has been studied intensively and several investigators have found evidence for structural complexity within D'' (e.g., *Garnero* [2000]; *Buffett et al.* [2000]; *Thomas et al.* [2004]). Here we present a sample 2-D image of lowermost mantle structure; a more complete analysis and interpretation of such images is presented in **Chapter II** (*Ma et al.* [2006]) and **Chapter V** (*Van der Hilst et al.* [2007]).

2.4.1 *ScS* data selection and pre-processing

For all events considered here, origin times and source locations (hypocenters) were obtained from *Engdahl et al.* [1998] and moment tensors and magnitudes from the Harvard CMT catalog. For all events in our data set, three-component broad-band waveforms were retrieved from the Data Management Center of the Incorporated Research Institutions for

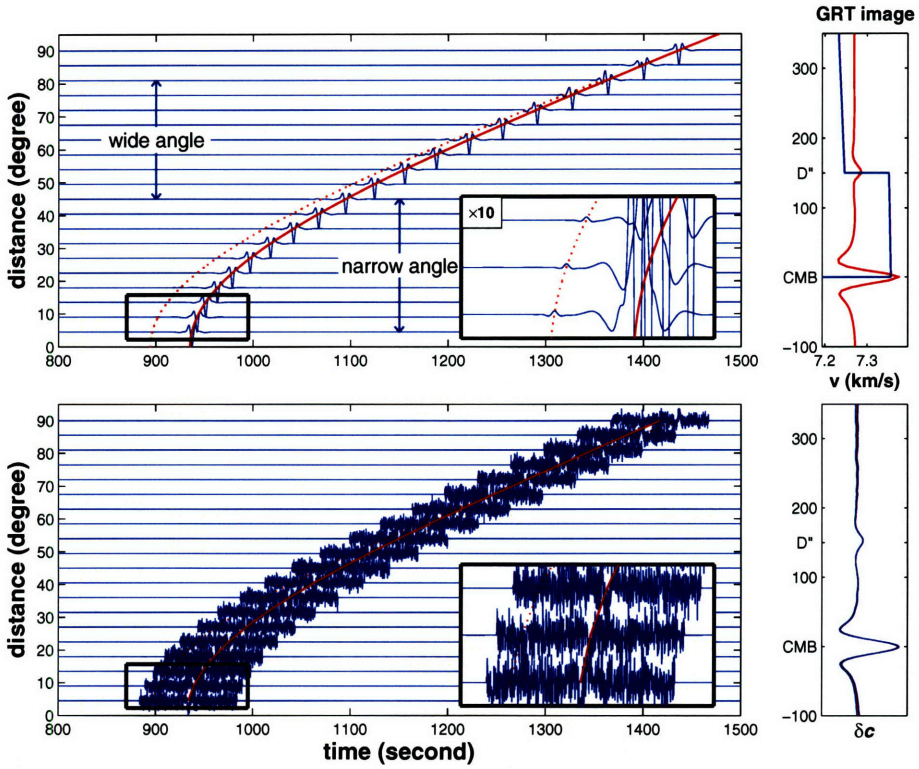

Figure 2-5: Illustration of the robustness of the GRT in the presence of random noise in the data. Top left: same as Figure 2-2A. Red lines are theoretical travel time curves for *ScS* (solid) and S_dS (dashed). Top right: GRT image trace (solid red line) constructed from the synthetic data shown on the left and the wavespeed profile used to generate the synthetic waveforms (blue curve). Bottom left: Data as in top panel after addition of (random) noise. The arrival of *ScS* can still be discerned in the noisy data, but signal from the top reflector has disappeared in the noise. Bottom right: GRT image trace (solid red line) constructed from the noisy data shown on the left. The image is practically identical to the noise-free image.

Seismology (IRIS).

The range of epicentral distances that show the most prominent specular *ScS* reflections is $20-70°$, but as input to the GRT we used data in the distance range from $0-80°$. We further require that the image points y are within the CMB patch considered here. The ~1,300 earthquakes (with $m_b > 5.2$, origin times 1988-2002) recorded at one or more of a total of ~1,200 stations (Figure 2-6A) yield a total of ~65,000 broad-band data (Figure 2-

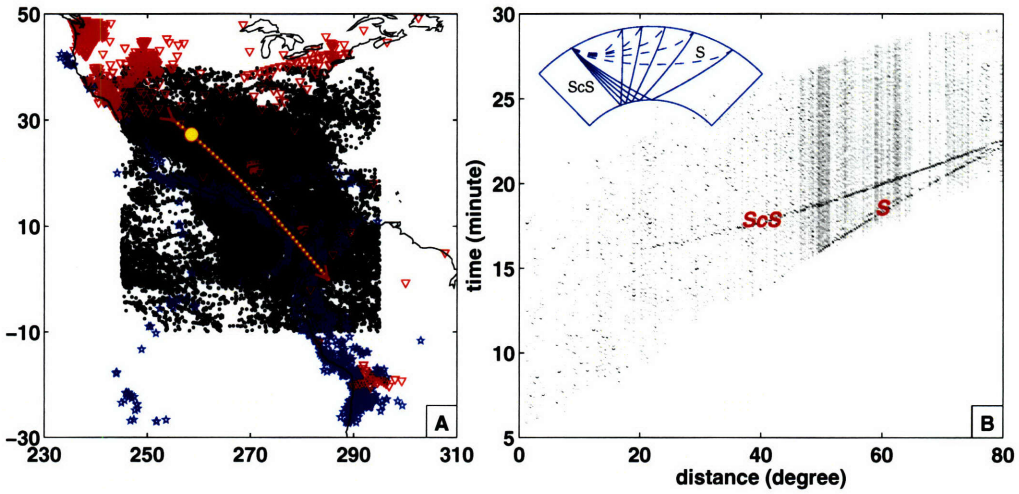

Figure 2-6: (A) Geographic map of the region under study, depicting the epicenters of the ~1,300 earthquakes (blue stars) and the locations of the ~1,200 stations (inverted red triangles) that yielded the data used in the construction of the common image-point gathers. The $50^\circ \times 50^\circ$ CMB bin is indicated by the densely sampled rectangle: small black dots mark specular CMB reflection points of the ~65,000 *ScS* data displayed in the panel on the right. The small yellow dots that delineate the NW-SE trending section line mark the locations of the image gathers constituting the 2-D profile presented in Figure 2-9; the large yellow dot represents the location of the IGs and angle stacks shown in Figure 2-8. (B) Stack of the ~65,000 $ScS(SH)$ data with reflection points in the CMB bin shown in (A). Processing details: data source IRIS-DMC; bandpass filter: $1-10$s; earthquakes: $m_b > 5.2$, origin time between 1988-2002. Inset, top left: generic ray geometry of *ScS*. NB. of these data, $\sim 35,000$ were used for the construction of the 2-D profile shown in Figure 2-8. ($\sim 30,000$ data were rejected either because the specular reflection point was too far from the image points or because the number of seismograms for particular earthquake was inadequate for PCA.)

6B). Subsets of this data set are use to construct GRT images at specific CMB locations.

Before we can perform the GRT we subject the data to several pre-processing steps. First, we account (by deconvolution) for the different instrument responses of the seismograph stations from which data are used. Second, we band-passed all data between 1-10 s. Third, we remove effects of source and receiver differences on the displacement field u. To obtain a (common) band-limited source signature, which allows the factorization in (2.12), we account as well as we can for the differences in rupture mechanisms of all

earthquakes involved. For this purpose we check the first-onset polarity and deconvolve the time derivative of the source time-rise functions, which can be estimated from CMTs or from a principal component analysis (see Section 2.4.2). Furthermore, all travel times are corrected for Earth's ellipticity (*Kennett and Gudmundsson* [1996]).

For the CMB bin and the source-receiver pairs considered here, the cone of associated dip directions has an opening of about 24 deg, which restricts the detectable structural dip angles. (We note that we only consider a cone perpendicular to the CMB, which restricts the structural dip angles; this limitation can be removed by considering multiple cones.) Furthermore, we invoke a bound on the difference between two-way travel time $T(x^s, x^r, y)$ and the travel time of the specular reflection at image point y. Finally, the broad-band seismograms are subjected to windowing in order to obtain time series that comprising main arrivals, their coda, and their precursors.

2.4.2 Principal Component Analysis (PCA)

As with other techniques involving stacks of earthquake data, we need to account for the differences in source signature of the many different earthquakes involved. Of primary interest here is the source time (rise) function. This can be inferred from the Harvard CMTs, but the frequency content and the type of data used to obtain the CMT solutions is quite different from those used in our study. Instead of CMTs we use a principal component analysis (PCA) – see, for instance *Rondenay and Fischer* [2003] – to estimate the relevant parameters directly from the data. As the direct wavefield we can use either *ScS* or (teleseismic) *S*; the latter has the advantage of not being involved in scattering in the CMB region but a disadvantage is that we cannot use some very narrow angles (associated with epicentral distances less than 30°). The following steps are used to estimate the

time-rise functions and to separate the direct and scattered wave fields: (1) The transverse components excited by the same earthquake are divided into different groups according to epicentral distances – if there are fewer than three records in one or more groups, the event (and associated data) is not used; (2) For each group, a Hilbert-transform is applied to the seismograms; (3) The transformed seismograms are time-normalized using delay times obtained from multi-channel cross-correlation (e.g., *VanDecar and Crosson* [1990]); (4) The seismograms are projected onto the first principal components (see *Ulrych et al.* [1998]), which are determinined for each group. (5) A 100 s window around the calculated travel time of the direct wave is then used to obtain an emperical 'time-rise' function for each record, which is deconvolved from the direct and scattered wave fields to obtain the data used for imaging of the CMB and structure above it, respectively. For PCA with *ScS* as the reference phase this process is illustrated in Figure 2-7.

2.4.3 Preliminary 2-D image

We draw from the ~65,000 *ScS* displacement records in Figure 2-6B to construct GRT images of the lowermost mantle beneath Central America (Figure 2-6A). We first consider an image gather and angle stacks at a particular image point y and then construct a 2-D profile by lateral juxtaposition of 41 angle stacks. In the current study we restrict the analysis to the bottom 400 km of the mantle in order to avoid contamination with the *S* wavefield.

For y at (−90°W, 15°N) we integrate over scattering azimuth ψ and form image gathers for different scattering angles θ and, hence, radial resolution bands (Figure 2-8). The preliminary results shown here are obtained by integration (according to (2.13)) either over narrow and wide opening angles, for *ScS* as the primary phase (Figure 2-8A), or over wide

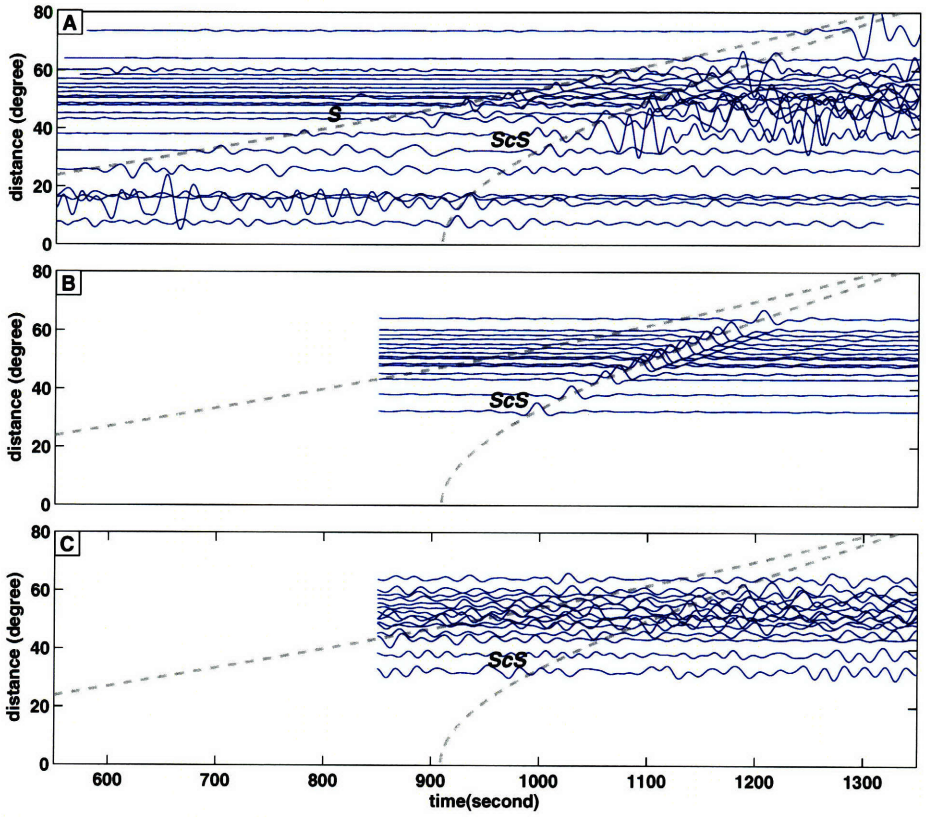

Figure 2-7: Illustration of principal component analysis. (A) (Preprocessed) raw data (see Section 2.4.1). (B) Deconvolution of the raw data using PCA estimation of the derivative of the *ScS* time rise function (using a time window of 100 s around *ScS*). (C) Raw data deconvolved with the *ScS* estimate minus the field shown in (B). Traces as in (B) and (C) are used to image the CMB and shallower structure, respectively, as shown in Figure 2-8A. Similar such estimates based on teleseismic *S* as the direct wavefield were used for Figures 2-8B and 2-9.

angles only, for *S* (Figure 2-8B). We integrate over narrow and wide angles seperately because of the change in polarity upon crossing the intramission angle; stacking over all angles would involve signals with opposite polarities and cold thus mask interfaces. The resulting angle stacks are depicted on the right of Figures 2-8A,B. The stacks for either *ScS* or *S* as the primary phase used in PCA both reveal contrasts in elastic parameters at $\sim$0 km (that is, the reference depth of the CMB) and $\sim$280-340 km above it. There is also strong evidence for structure in between.

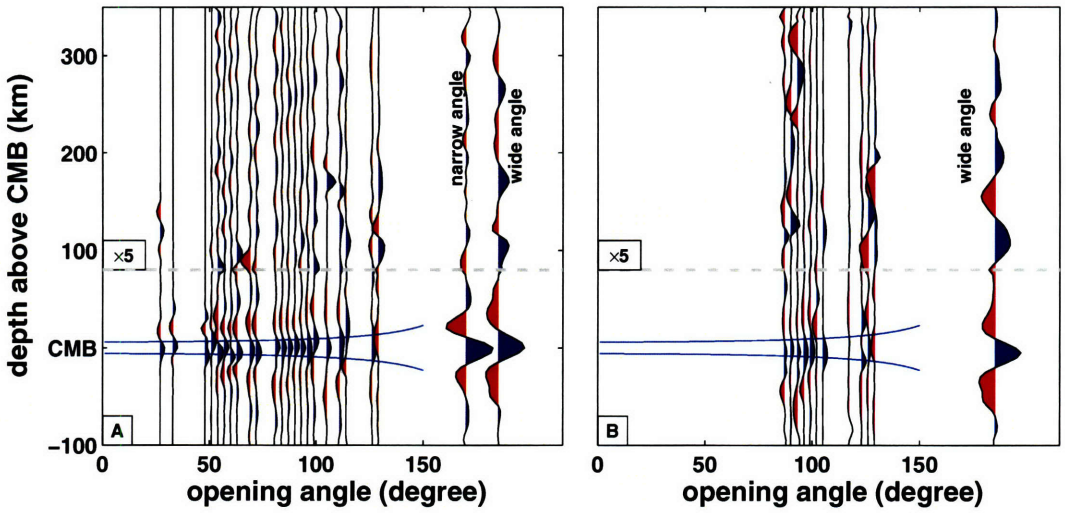

Figure 2-8: Construction of image gathers and angle stacks with real data for an (arbitrary) image point marked by the yellow dot in Figure 2-6: (A) PCA with *ScS* as the direct wavefield, (B) PCA with (teleseismic) *S*. Similar to Figure 2-3, in each panel we show to the right of the image gathers the stacks over the scattering angles. (As before the theoretical prediction of the dilation is given by the thin blue lines around the depth of the CMB.) Note that for PCA with *S* we only considered wide angle data.

We repeat this procedure to create angle stacks at other image points along a ~2,500 km long great-circle transect from (−105°W, 30°N) to (−75°W, 0). Lateral juxtaposition of these stacks creates a (scatter) density plot for the deepest 400 km of Earth's mantle. Of the $\sim 65,000$ records depicted in Figure 2-6B, $\sim 35,000$ were involved in this calculation. The other data were not used either because their specular reflections were too far from the line of section considered or because the number of seismograms for particular earthquake was inadequate for the principal component analysis (see above). The image in Figure 2-9 shows high scatter from the depth corresponding to the CMB, and it indicates substantial structural complexity in the lowermost mantle above it. In addition to a weakly undulating feature near 280-340 km above CMB, which seems laterally continuous over many hundreds of km, the image reveals pronounced structures at smaller distances above the CMB. We note that these structures are not present in the tests with synthetic data calculated from

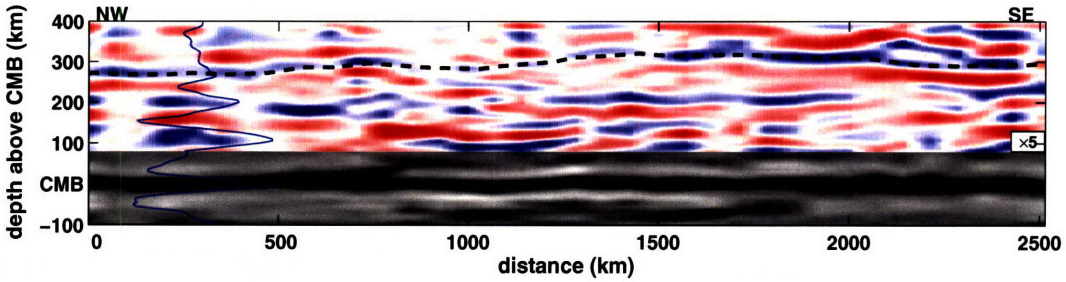

Figure 2-9: 2-D image of the CMB and lowermost mantle beneath Central America. Using a total of $\sim 35,000$ broad band records, this 2,500 km long profile is produced by juxtaposition of and linear interpolation between angle stacks of the image gathers for 41 image points, evenly spaced along the line of section depicted in Figure 2-6A. As an example, the stack depicted in Figure 2-8B is plotted at the NW (that is, left) end of the profile (large yellow dot in Figure 2-6A). The gray-scale part of the image depicts the CMB contrast whereas the part in color (amplified by a factor of five relative to the CMB part) reveals structure (stratification?) in the lowermost mantle. The dashed line marks the blue contrast (with side lobes in red) at ~280-340 km above the CMB, which may represent the top of the so called D'' region. The image is rich in structure at depths between the CMB and the top of D'' but we refrain from further interpretation until we have performed a rigorous statistical analysis (*Ma et al.* [2006]).

a signle contrast above CMB (e.g., Figure 2-4).

2.5 Discussion and concluding remarks

To enable the efficient exploration of interfaces in Earth's lower mantle over large geographical areas we will combine inverse scattering (through a generalized Radon transform) with (mixed-effect models) statistical inference and model validation. The generalized Radon transform (GRT) of broad-band *ScS* data is developed here; the statistical analysis is presented in Paper II (*Ma et al.* [2006]). The GRT method uses three-component, broad-band waveforms and exploits the redundancy in large modern data sets. In this context, with redundancy we mean that the combination of specular and non-specular reflections – at different scattering angles – produce multiple images of the same points in the image.

The radial resolution of interface depth depends on (i) the scattering angle (through the dilation, as discussed above and as shown, for instance, in Figure 2-3) and (ii) the frequency of the data used. For periods of ~ 5 s, scatter angles of $\sim 100°$, and shear wavespeed of ~ 8 km/s, the (quarter wavelength) radial resolution is $(0.25 \times 5 \times 8)/cos(100°/2) \sim$15 km.

In contrast to labor intensive forward modeling of individual or stacked waveforms, our imaging method is highly automated and imposes few *a priori* constraints on the geometry and nature of the structures that we attempt to image. Indeed, the only prior knowledge concerns the type of seismic phase considered (here, *ScS* and its coda and precursors), so that appropriate time-windows can be extracted from the recorded wavefield, and the requirement that (at least locally) the singularities form a contiguous interface. With this information, the data themselves will yield structure in the neighborhood of a predefined imaging point. Since submission of this manuscript, several studies have published results from Kirchoff migration stacking methods (e.g., *Avants et al.* [2006a]). Our method has several aspects in common with this technique but differs in that it is explicitly 3-D and that it accounts for wave amplitudes (geometrical spreading and source radiation) and caustics due to wave propagation in a heterogeneous background model. Furthermore, in our automated approach, statistical inference (Paper II) replaces modeling of the (Kirchoff) stacks as the primary means for model validation. Of course, forward modelling can be used to explore in more detail structures of particular interest revealed by our method, but that is beyond the scope of the analysis presented here.

We have performed a series of tests with synthetic data to confirm theoretical predictions, to establish the accuracy of the GRT, and to test the performance of the GRT in the presence of (random) noise. Adding noise to the synthetic data, giving signal-to-noise ratios well below 1, demonstrates that the GRT can detect and locate medium contrasts

correctly even if the pertinent signal is not apparent from visual inspection of the 'raw' data. With sufficient data coverage, multiple interfaces as well as interfaces marking small medium contrasts (of the order of a few per cent) will then be imaged correctly. Sources of error that are not 'random' will, however, degrade the GRT images and may increase the minimum medium contrast that can be reliably detected and imaged. Examples of more realistic, non-random noise include the distortion of the image gathers due to the uneven source-receiver distribution, the presence of signal due to multiple scattering, and the effects of using an incorrect background wavespeed model. In Paper II we use statistical inference and validation methods to deal with such complications and to quantify the uncertainty of the resulting GRT images. Because this is not done here we will refrain from detailed interpretations of the sample result presented above.

An important aspect – and source of uncertainty – of the GRT presented here as well as similar such methods based (somehow) on the stacking of earthquake data is the estimation and removal (by deconvolution) of the different source pulses. Imperfect removal of the pulse can produce artificial structure in the stacks. *Chambers et al.* [2005] and *Avants et al.* [2006a] visually inspect the deconvolved data and remove bad traces. This labour intensive approach is feasible if one uses 'only' several hundred waveforms. For the applications that we are interested in – that is, automated processing of tens or hundreds of thousands of records – this is not feasible and other approaches toward source pulse estimation and deconvolution must be sought. We have considered here a principal component analysis (PCA) but we are exploring the use of more robust (statistical) methods.

We have demonstrated the feasibility of the GRT method with an application to $\sim$ 35,000 broad-band records of *ScS* waves that reflect off the CMB beneath Central America. The (preliminary) image profile inferred from these data (Figure 2-9) reveals strong

contrasts in elastic parameters at about 0, that is, at the depth of the CMB. It is encouraging that with neither visual data inspection nor prior assumptions about the geometry of target structures the CMB appears so clearly in the images (e.g., Figures 2-8 and 2-9). The 2-D image also reveals a quasi-continuous structure between 280 and 340 km above the CMB. It is tempting to interpret this as the 'top' of the so called D'' layer. Changes in elastic parameters near this depth have been the subject of many seismological studies (e.g., *Lay and Helmberger* [1983b]; *Tromp and Dziewonski* [1998]; *Sidorin et al.* [1999]; see *Garnero* [2000] for a comprehensive review), but there is as yet no consensus on this transition and its radial and lateral extent. Our preliminary results are consistent with an undulating surfaces (*Thomas et al.* [2004]), but they also reveal structures that could be interpreted as jump-like steps in the discontinuity (*Hutko et al.* [2006]). Further study is required in order to establish whether such steps are real or whether they can represent a continuous phase boundary (as suggested by *Sun et al.* [2006]) and whether they are a unique, local phenomenon (as implied by *Hutko et al.* [2006]) or a more general feature of D''.

The image presented here also provides tantalizing evidence for interfaces closer to the CMB. This may suggest that the lowermost mantle is stratified and more complicated than hitherto thought. Further study is needed, however, to establish the (statistical) significance and lateral extent of these multiple transitions. We expect that incorporation of constraints from experimental and theoretical mineral physics with the seismological estimates of interface regularity will help distinguish between compositional, mineralogical, and petrological boundaries. Mineral physicists have recently discovered a phase transformation of $MgSiO_3$ perovskite at a pressure that could coincide with the contrast depicted between 280-340 km above CMB (*Murakami et al.* [2004]; *Shim et al.* [2004]). It should be noted, however, that important aspects of this purported phase transition remain either

unknown or uncertain (Dan Shim, MIT, personal communication, 2005).

In addition to statistical analysis (Paper II) and cross-cutting seismology-mineralogy analyses, a logical follow-up of the study presented here would be the exploration of much larger CMB regions. This extension of our current study is possible because of the availability of large volumes of data through international data centers. We note, however, that elsewhere in the world the data coverage may not be as good as considered in the area of interest here, which would make the statistical analysis for image enhancement and validation (Paper II) all the more important.

2.6 Acknowledgments

We thank Gary Pavlis, Jesse Lawrence, and Justin Revenaugh (Associate Editor, JGR) for their rigorous and constructive reviews, which helped us improve the manuscript. We acknowledge Stéphane Rondenay (MIT) for suggesting to use the Principal Component Analysis to estimate source parameters from the direct wavefield. This study would not have been possible without the efficient data storage at and retrieval from the Data Management Center of the Incorporated Research Institutions for Seismology (IRIS). The research presented here was funded by the Collaborative Mathematics and Geosciences program of the US National Science Foundation (grant EAR-0417891).

2.7 Appendix: Amplitude analysis of the GRT image gathers

In this appendix we discuss an expression for $S^{(1)}$ by accounting properly for the source and contrast-source radiation patterns in modeled seismic data. This also leads to an introduction of residual amplitude compensation within the GRT – here, derived from the Born-Kirchhoff approximation for scattered body waves.

The relative contrast in the medium parameters is formally defined by the 'vector'

$$\mathbf{c}^{(1)}(y) = \left\{ \frac{\rho^{(1)}(y)}{\rho^{(0)}(y)}, \frac{c^{(1)}_{ijkl}(y)}{\rho^{(0)}(y) V^s_\circ(y) V^r_\circ(y)} \right\}^T . \tag{2.15}$$

Its dimension (number of independent parameters) depends on the local symmetry of the elastic medium. Here, $V^s_\circ$ and $V^r_\circ$ are the phase velocities at y averaged over phase angles. We have assumed that $\mathbf{c}^{(1)}(y) = \tilde{\mathbf{c}}^{(1)}(y, \phi(y))$ with $(\tilde{\mathbf{c}}^{(1)})'(y, \phi(y)) = \mathbf{C}(y)\, \delta(\phi(y))$, where $'$ denotes the derivative with respect to the second argument, and $\mathbf{C}$ denotes the local magnitude of the jump across (a specific vaulue) of the function ϕ that defines the interface. Then [*De Hoop and Bleistein*, 1997, (38),(62)]

$$S^{(1)}(x^\phi; \theta, \psi) = W(x^s, x^2, x^\phi)^2 \mathbf{w}^T(x^s, x^r, x^\phi) \mathbf{C}(x^\phi),\ x^{s,r} = x^{s,r}(x^\phi, \nu_\phi, \theta, \psi), \tag{2.16}$$

where $\mathbf{w}$ denotes the 'vector' of radiation patterns

$$\mathbf{w}(x^s, x^r, y) = \left\{ h^s_m(y) h^r_m(y), [h^s_i(y) p^s_j(y) h^r_k(y) p^r_l(y)] V^s_\circ(y) V^r_\circ(y) \right\}^T . \tag{2.17}$$

We refer to $S^{(1)}$ as linearized scattering coefficients; $\hat{S}^{(1)}$ is a filtered realization of $S^{(1)}$,

where the filter is determined by the actual illumination. From the expression for $S^{(1)}$ we may extract the linearized reflection coefficient

$$\frac{S^{(1)}(x^\phi;\theta,\psi)}{W^2(x^s,x^r,x^\phi)^2}\,\frac{1}{[V^s(x^\phi)V^r(x^\phi)^3]^{1/2}(\nu_\phi\cdot p^m(x^\phi))^2}.$$

To estimate directly this reflection coefficient we thus replace in the GRT the obliquity factor $|p^m(y)|^3$ by $|p^m(y)|$. If, in contrast, we want to relate the image directly to the stiffness perturbation, we have to use another modification of transform (2.7): instead of $\hat{S}_w(y;\theta,\psi)$ we then define the image gather $\mathscr{I}(y;\theta,\psi)$ as

$$\mathscr{I}(y;\theta,\psi) := \int_{E_{\nu^m}} \frac{\tilde{u}(x^s,x^r,y)}{W(x^s,x^r,y)^2}\,\frac{|p^m(y)|^3}{|\mathbf{w}(x^s,x^r,y)|}\,d\nu^m. \tag{2.18}$$

Here, $|\mathbf{w}(x^s,x^r,y)|$ is the Euclidean norm of the 'vector' of radiation patterns. Thus defined, $\mathscr{I}(y;\theta,\psi)$ represents a dimensionality preserving transformation of data to a set of images: the common image-point gathers (IGs).

At specular reflection points, $S^{(1)}$ in (2.16) gets replaced by

$$\mathbf{e}(x^\phi;\theta,\psi)^T\mathbf{C}(x^\phi) \quad \text{with} \quad \mathbf{e}(x^\phi;\theta,\psi) = \frac{\mathbf{w}(x^r(x^\phi,\nu_\phi,\theta,\psi),x^\phi,x^s(x^\phi,\nu_\phi,\theta,\psi))}{|\mathbf{w}(x^r(x^\phi,\nu_\phi,\theta,\psi),x^\phi,x^s(x^\phi,\nu_\phi,\theta,\psi))|}.$$

We anticipate that $\mathbf{e}$ is only weakly dependent on (θ,ψ) so that the image gathers $\mathscr{I}(y;\theta,\psi)$ are approximately uniform in (θ,ψ).

Chapter 3

Imaging of Structure at and Near the Core Mantle Boundary using a Generalized Radon Transform: II - Statistical Inference of Singularities[†]

Abstract

We present Part II of our approach to high resolution imaging of deep Earth's interfaces with large volumes of broad-band, three-component seismograms. We focus on the lowermost mantle – also referred to as D'' region – but the methodology can be applied more generally. Part I (**Chapter II**, *Wang et al.* [2006]) describes the generalized Radon transform (GRT) of broad-band *ScS* data (comprising main arrival, precursors, and coda). The GRT produces "image gathers", which represent multiple images of medium constrasts at

[†]Published as: Imaging of Structure at and Near the Core Mantle Boundary using a Generalized Radon Transform: II - Statistical Inference of Singularities, *J. Geophys. Res., inpress*. As the second author of this paper, I helped build the statistical models. Furthermore, I performed all synthetic tests and applied the statistical models to the seisimc images.

the same image point near the base of the mantle. With a method for statisticial inference we use this redundancy to (i) enhance the GRT images through improved recovery of weak contrasts and through suppression of spurious oscillations in the GRT image gathers and (ii) provide uncertainty estimates that can be used to identify the robust features in the images. Using the image gathers from Paper I (**Chapter II**, *Wang et al.* [2006])as input, we use mixed-effects statistical modeling to produce the best estimates of reflectivity along with their uncertainty. In this framework, random noise in the signal is separated into white and coherent components using the geometry of the (GRT) imaging operators and a generalized cross-validation method. With synthetic data we show that conventional GRT images deteriorate substantially, in some cases to the point at which weak reflectors can no longer be detected, due to effects of uneven sampling, wave phenomena that are not accounted for in the underlying single scattering approximation, or errors in the assumed background wavespeed model. We demonstrate that even in these circumstances statistical analysis can yield adequate estimates of the true model. GRT imaging produces robust images of the core mantle boundary (CMB) beneath Central American and suggests the presence of several structures in the D'' region, in particular between 100-200 and between 270-320 km above the CMB proper. Most of these structures are significant at the 1σ (that is, 68%) level, but at 2σ (95%) confidence the images show, at various depths above the CMB, intermittent instead of laterally contiguous features.

3.1 Introduction

The remote sensing of deep mantle discontinuities, for instance the core mantle boundary (CMB), is a challenge in part because the seismic waves used to probe it propagate (at least twice) through Earth's heterogeneous mantle before they are observed and in part because the diagnostic seismological signals are often too small for direct observation and (forward) waveform modeling. Since the pioneering work by *Lay and Helmberger* [1983a], many seismologists have tried to image the so called D'' region of enhanced heterogeneity in the bottom 300 km or so of the mantle (see *Garnero* [2000] and *Helmberger and Ni* [2005] for extensive reviews). Recently, mineral physicists presented compelling evidence for a phase transition in the mantle silicate $(Mg,Fe)SiO_3$ – from perovskite (pv) to the so-called post-perovskite (ppv) phase – at pressures that roughly coincide with the changes

in elastic parameters inferred from seismic imaging (*Murakami et al.* [2004], *Oganov and Ono* [2004], *Shim et al.* [2004]). The presence of a phase transition is consistent with earlier seismological observations and geodynamical arguments (*Sidorin et al.* [1999]; *Van der Hilst et al.* [2007]).

These are exciting developments for studies of Earth's deep interior. But much uncertainty remains. On the one hand, estimates from theoretical and experimental mineral physics of the pressure at which the transition occurs show a large uncertainty, and the temperature and composition dependencies are not yet precisely known (*Shim* [2005]). Consequently, the depth at which the pv-ppv transition occurs has an uncertainty of several 100 kilometers. On the other hand, the seismological detection and characterization of such subtle and remote changes in elasticity faces formidable observational and theoretical (and computational) challenges. Various types of noise and scatter from 3-D heterogeneity can mask weak signals in seismic data. Along with the massive size of modern data sets, this poses severe limitations on forward (waveform) modeling. The increasing availability of large volumes of densily sampled broad-band data has begun to allow application of subsurface imaging methods based on inverse scattering, which exploits more efficiently the rich information contained in seismic waveforms.

To meet the challenge of imaging and characterizing structure at and near remote interfaces and boundary layers, we are developing techniques for the automated identification, extraction, and interpretation of structural signal pertinent to subtle medium contrasts. Our approach differs in several important ways from forward modeling: First, we exploit the redundancy in large data volumes. Second, we make only a few restrictive *a priori* assumptions about the structures of interest. Third, we use data from a wide range of incidence angles. A simple analysis of reflections at an interface that marks a wavespeed increase shows

that the reflection coefficient is very small for near vertical incidence and increases dramatically toward critical incidence. For this reason, waveform modeling is usually restricted to observations made near (and beyond) the critical incidence, that is, large epicentral distances. In addition to the obvious reduction in target regions that can be studied, and the need to deal with triplicated waveforms, one should also realize that the radial resolution to interface depth degrades markedly with decreasing vertical slowness. Indeed, the wide angle reflections considered in forward modeling have, in general, rather poor sensitivity to the depth of the contrasts being studied. Narrow angle data provide better resolution of discontinuity depth, but because of their small amplitudes they are rarely used in forward modeling. A further difference is that we can estimate formal uncertainties on the estimates of interface properties.

We combine concepts from inverse scattering and modern statistics into a two-step strategy. As the first step, *Wang et al.* [2006], hereinafter referred to as Paper I (**Chapter II**, *Wang et al.* [2006]), developed a generalized Radon transform (GRT) of global seismic network data in heterogeneous, anisotropic elastic media to map tens of thousands of seismograms to a set of multiple images of the same target structure. These 'common image-point gathers' reveal multi-scale variations in elastic properties. For a detailed discussion and a historical perspective we refer readers to Paper I (**Chapter II**, *Wang et al.* [2006]).

In the second step, which is the focus of this paper, we obtain estimates of variations in Earth's deep interior from the image gathers using a statistical approach, in which the image gathers are modeled nonparametrically using mixed-effects statistical models. In this framework, the random noise in the signal is allowed to have white and coherent components, and the latter are estimated from the data through prediction error minimiza-

tion (also known as generalized cross-validation). This methodology, a flexible type of Tikhonov regularization, can be used with different types of correlated noise and with the typically sparsely and unevenly sampled image gathers owing to the geographic distribution of sources and receivers.

The mixed-effects estimate of the reflectivity profile replaces the linear stack in the conventional GRT. This achieves three specific goals: first, it enhances signal-to-noise in the image gathers; second, it adapts to and mitigates effects of error in the background wavespeed model, and third, it provides quantitative uncertainty estimates, which are more satisfactory than the ones from ubiquitous bootstrapping of slowness stacks. Of key importance is the 'pre-stack' aspect of the analysis: the 'common image point gathers' (and not the 2-D image profiles) are subjected to statistical analysis, and the optimal gathers are combined into the stack for that image point. This allows us to exploit the additional information contained in the dependence on scatter angle and to identify and remove poorly constrained gathers as well as artifacts due to, for instance, incorrect back ground velocities and presence of signal that cannot be modeled with single scattering.

In Section 3.2 we briefly summarize the (geometrical) aspects of the GRT that we need for the statistical analysis. In particular, we mention how three-component broad-band global network data can be transformed to so called common image-point gathers. In Section 3.3, we explain the concept of mixed-effects statistical models and describe how we transform the image gathers to statistical estimates of discontinuities. Technical aspects of parameter estimation in mixed-effects models are presented in Appendix 3.8. In Section 3.4 we test the effectiveness of the methodology on synthetic data against the presence of various types of noise, the imprint of source-receiver geometry, and imperfections of the mantle (wavespeed) model. Finally, in Section 3.5, we apply the method to the *ScS* wave-

field (containing ~80,000 broad-band records) for the purpose of imaging the CMB and overlying D'' region beneath Central America. In addition to a laterally continuous image of the core mantle boundary, the resulting 2-D image profile reveals intriguing structure, including multiple interfaces, in the few hundred km above the CMB.

3.2 Common image-point gathers

For a detailed discussion of the construction of the GRT we refer the readers to Paper I (**Chapter II**, *Wang et al.* [2006]), but for completeness sake we mention here the aspects that are relevant for the development of the statistical methods. In essence, The GRT enables the automated transformation of a large number of seismic waveform data into a set of multiple images of the same structure in the vicinity of a target region. In order to do so, one has to account for variations in volumetric wave speed or, in general, density normalized stiffness (in case of an anisotropic model) and such geometrical aspects as the focal depth and radiation patterns of the earthquakes considered, the various move-outs due to the large range of slownesses and epicentral distances of the data used, and the effects of uneven sampling. The geometry involved in the reconstruction is illustrated in Figure 3-1A. Figure 3-1B depicts the study region considered here as well as the distribution of *ScS* reflections at the CMB associated with the broad-band wavefield used used to construct the image profiles presented later in this paper.

With $y = (y_1, y_2, y_3)$ the image point and superscripts s and r the association with a ray from a source and a receiver, respectively, the 'two-way' travel time for a particular diffraction branch associated with a ray path connecting x^r with x^s via y is denoted by $T = T(x^s, x^r, y)$. The slowness vector of the ray connecting a source point x^s with image

point y, evaluated at y, is given by $p^s(y)$, and $p^s(x^s)$ indicates the slowness along this ray evaluated at the source. Furthermore, we introduce the phase direction $\alpha^s = p^s/|p^s|$ and the phase velocity V^s according to $|p^s| = 1/V^s$. A similar notation is employed for the slowness vector related quantities along the ray connecting the receiver with the image point, namely $p^r(y)$, $p^r(x^r)$, and α^r and V^r.

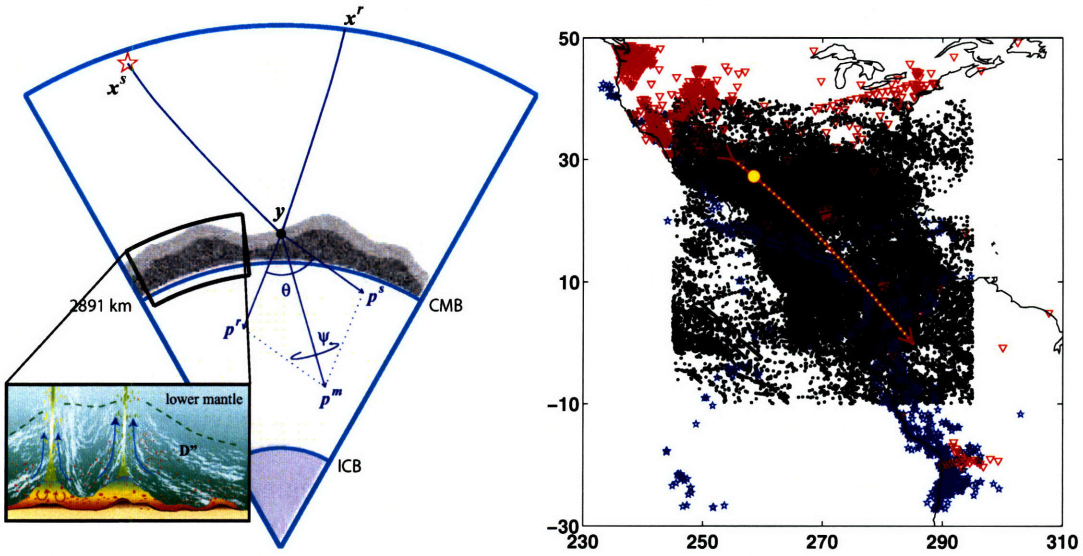

Figure 3-1: Left: The geometry of the GRT with ScS precursors and coda waves. For illustration purposes, the ray geometry and associated imaging parameters are shown for scattering at image point y. The objective of the work presented here – and in *Wang et al.* [2006] – is the high resolution imaging of the structures in the lowermost mantle, also referred to as D'' region, that may arise from boundary layer processes (e.g., flow), lateral variations in composition, and pressure induced phase changes (as depicted in the inset, after *Garnero* [2000]). Right: geographical map of study region. Small black dots depict ~80,000 (specular) CMB reflection points associated with the broad-band *ScS* wavefield used to construct the image profiles presented in this paper.

A key element in the GRT, the migration dip, $\nu^m(y)$, is the direction $\nu^m(y) = p^m(y)/|p^m(y)|$ of the migration slowness vector, $p^m(y) = p^s(y) + p^r(y)$. Together, the migration dip and the phase directions of incoming and scattered rays define the scattering vector,

$$\psi = (\alpha^s \times \alpha^r) \times \nu^m \quad \text{at } y. \tag{3.1}$$

For a particular travel time diffraction branch away from caustics at x^r or x^s, the opening (or scattering) angle, θ, between incoming and scattered rays is related to the scattering vector according to

$$\sin\theta = |\psi| \quad \text{at } y; \quad \theta = \theta(x^s, x^r, y). \tag{3.2}$$

The scattering azimuth, ψ, is the angular displacement of the scattering vector, normalized to one: $\psi/|\psi|$. An image gather at y can now be obtained by integrating, for each (θ, ψ), the pre-processed global network data over migration dip v^m (Figure 3-1A).

With a generalized Radon transform the data are transformed to an extended image volume, $\mathscr{I}(y;\theta,\psi)$. Typically, one distinguishes the depth coordinate from the other coordinates representing the image point y; an image gather is formed by plotting the image (or reflectivity) as a function of depth, which forms a radial reflectivity profile, against scattering angle and azimuth (θ, ψ). Integration over scattering azimuth, ψ, then yields azimuth-integrated (normalized) reflectivity profiles as a function only of scattering angle. On the right of Figure 3-2 we plot for each opening angle the best estimate of reflectivity as a function of depth (using statistical models as discussed in the next section) along with the best estimate of radial variations in reflectivity inferred from the angle dependent traces. Lateral juxtaposition of this estimate then produces 2-D image profiles.

The fact that image gathers form multiple images (namely, one for each pair of scatter angle and azimuth (θ,ψ)) of the same image point represents redundancy in the data. This is exploited in the statistical development of the GRT but should be accounted for with care. First, in Paper I (**Chapter II**, *Wang et al.* [2006]) we noticed and analyzed the dilation with scattering angle θ as well as the possible presence of a polarity flip at a particular scattering angle. Second, artifacts (with residual moveout in (θ, ψ)) may appear due to the presence

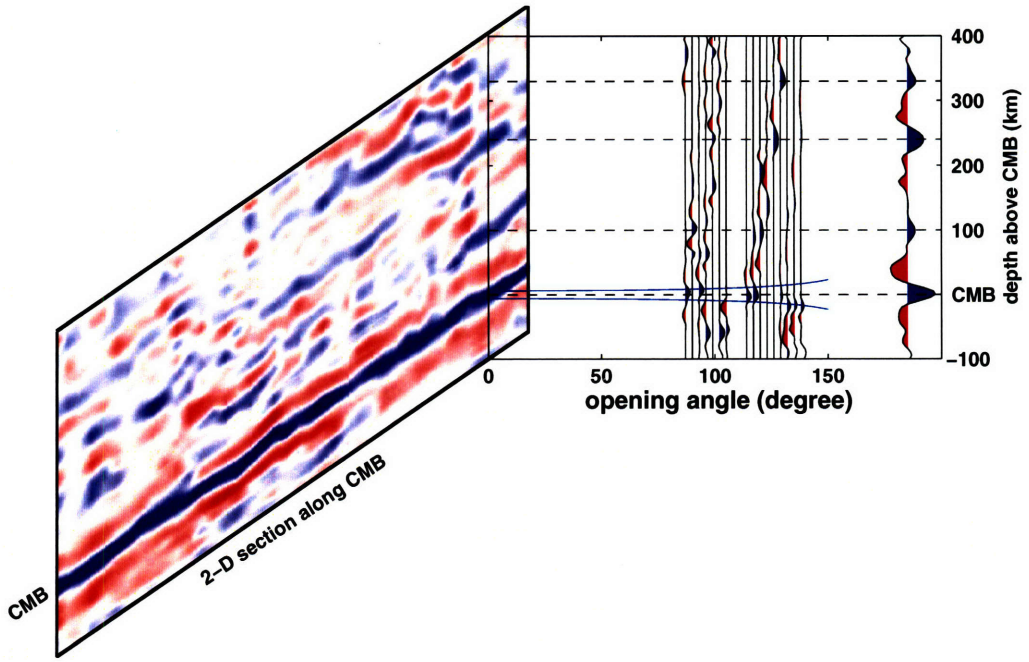

Figure 3-2: Image volume generated in Paper I (**Chapter II**, *Wang et al.* [2006]). Right: common image point gathers in (scatter, or opening) angle for a selected geographic location at the CMB; for each scatter angle the best estimate of the (azimuth-integrated) reflectivity as a function of depth is shown. To the right of these gathers we show the GRT trace, which represents the optimal estimate of the radial contrasts in reflectivity at that particular CMB location. We note that the statistical inference described in this paper is used to produce the best possible estimate of such an GRT trace for a specific location; as such, it replaces direct (non-) stacking linear stacking over azimuth and scatter angle. Left: 2-D image profile that results from lateral juxtaposition of 40 of such GRT images. Peaks in contrast as revealed by the GRT produce blue 'events' in the seismic section to the left. We note that except for interpolation between the GRT traces, no lateral smoothing or other image processing (or statistical inference) is used to produce such 2-D image profiles.

of caustics. (These can be removed by extending the GRT as in *Stolk and de Hoop* [2002].) Third, imperfections in the wavespeed model will lead to residual moveout with (θ, ψ). Fourth, the limited acquisition footprint leads to small shifts in (radial) depth in reflector images, different for each (θ, ψ). Fifth, scattered phases different from the phases scattered off the neighborhood of the CMB can lead to artifacts (locally) in the image gathers.

3.3 Statistical inference of singularities

3.3.1 From image gathers to medium contrasts

How can we best extract (robust) information about contrasts in elastic (medium) parameters, that is, radial reflectivity profiles $\mathscr{I}(y)$, from the above-mentioned image gathers (IGs)? Traditional methods involve stacking over (θ, ψ),

$$\mathscr{I}(y) = \int\int \mathscr{I}(y;\theta,\psi)\, d\theta d\psi \,. \tag{3.3}$$

Indeed, a structural image could be obtained by performing a GRT, followed by an integration over scattering angle and azimuth. With array observations, the signal-to-noise ratio could be improved, for instance, by means of phase-weighted stacking, which is non-linear in the data (*Schimmel and Paulssen* [1997]). However, this approach does not fully benefit from structural information that might be contained in the 'noise'. Moreover, it would not be clear how to assess the uncertainty of the final estimate.

Singularities in one-dimensional signals could be detected by means of wavelets, but the resolution and uneven coverage of the (multi-dimensional) image gathers make an approach based on the wavelet transform unfeasible.

The analysis that we develop here can be viewed as a focusing procedure, in which the geometry and statistics of 'noise' in the data is used to enhance the scanning for singularities or discontinuities through common image-point gathers. In the data we distinguish and model separately the systematic (non-random) and random components, hence the name 'mixed effects' statistical modeling. (We note that this approach is similar to Tikhonov regularization with random effects.)

Assuming that we have IGs in the angle domain (that is, $\mathscr{I}(y;\theta,\psi)$) we estimate an optimal reflectivity profile $\mathscr{I}(y)$ through 'mixed-effects' statistical inference instead of (linear or non-linear) stacking, as in (3.3). This allows us to mitigate more effectively the systematic errors due to uncertainty in background velocity, source location and origin time, and for imaging artifacts, artifacts owing to phase misinterpretation and, e.g., near source scattering, and errors introduced during preprocessing (e.g., filtering).

3.3.2 Mixed-effects models

Motivation and Strategy

To motivate the methodology, we recall how a line is fitted to observed data using the linear model $Y_i = \alpha + \beta x_i + \varepsilon_i$ $i = 1,\ldots,n$, where x_i are fixed design points and ε_i are independently and identically distributed (usually Gaussian) with mean zero and variance σ^2. Typically we estimate the slope β and intercept α using a least squares approach based on minimizing the residual sum of squares $\frac{1}{n}\sum_{i=1}^{n}(Y_i - \alpha - \beta x_i)^2$.

Suppose that instead of fitting a straight line we wish to fit a curve to the data; thus we write our model as $Y_i = \eta(x_i) + \varepsilon_i$. The residual sum of squares that is to be minimized can be written as

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - \eta(x_i))^2. \tag{3.4}$$

By fitting a curve we mean that not only do we want to estimate the value of $\eta(x)$ at $x = x_i$ but, in fact, at any x in the domain. Clearly this problem is ill-posed as there are many functions that pass through all the observed data points with a zero residual sum squares. We need to impose some restrictions on $\eta(x)$ in order to regularize the problem and obtain a reasonable function estimate (see Figure 3-3).

One approach is to assume that the function $\eta(x)$ is a linear combination of polynomials. However, polynomials over large intervals often display undesirable oscillations and other artifacts, especially when such polynomials are of order greater than three. Another approach is to find a balance between the residual sum of squares and a measure of the smoothness of the unknown function, measured, for example, by a functional $J(\eta)$. To obtain such a functional, we use the norm of the first derivative: $J(\eta) = \int \eta'(x)^2 dx$. In our case, smoothness is justified by the smoothing effect of a convolution of the reflector with a deterministic resolution filter, which reflects the acquisition imprint.

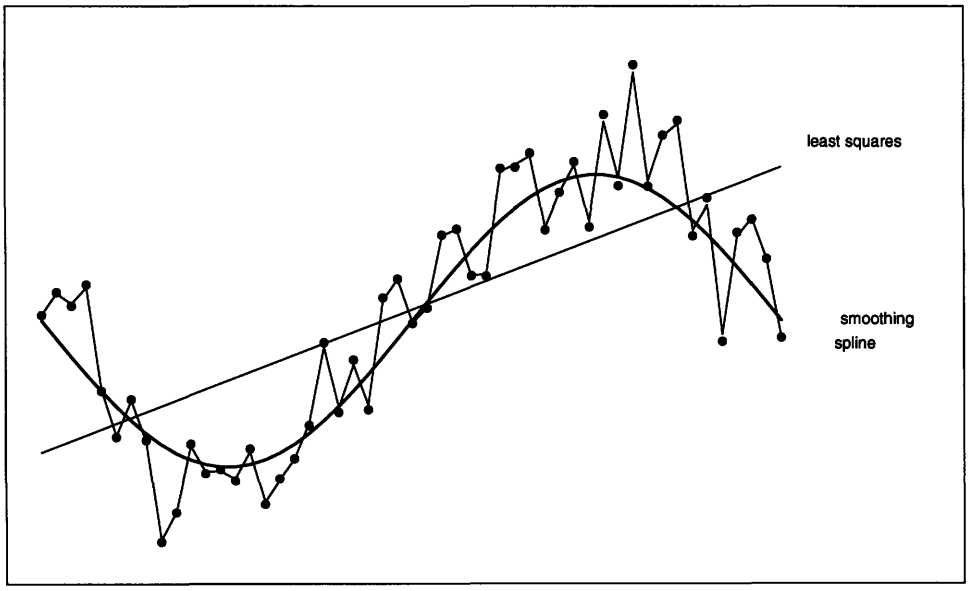

Figure 3-3: Curve fitting with an arbitrary function showing a simple linear interpolation (over-fitting), a least squares fit, and a smoothing spline estimate.

To fit the data with smooth function, instead of (3.4) we minimize

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i-\eta(x_i))^2+\lambda J(\eta), \tag{3.5}$$

where λ is a smoothing parameter that controls the trade-off between the goodness-of-fit and smoothness of η. This method is commonly called penalized least squares or Tikhonov regularization. The minimization of (3.5) is performed in a space of functions where the evaluation of a function at a point is a continuous linear functional (i.e., a reproducing kernel Hilbert space). Remarkably, the function $\eta(x)$ that minimizes (3.5), η_λ, turns out to be a finite linear combination of particular basis functions. The minimization problem is, thus, reduced to solving a linear system for the coefficients of this linear combination. The parameter λ is estimated through a refined leave-one-out cross-validation. Penalized least squares has been studied extensively in the literature; see, e.g., *Wahba* [1990] and *Gu* [2002] for comprehensive treatments of the subject.

A disadvantage of the described method is that it does not perform well with correlated noise, which limits its applicability in a variety of settings. To overcome this difficulty we use 'mixed-effects' statistical models, which explicitly distinguish systematic (non-random) from random components. They provide a unified framework for modeling a variety of correlated data (*Vonesh and Chinchilli* [1997], *Wang* [1998] and *Pinheiro and Bates* [2000]). For our purposes, an important characteristic of these models is that it can flexibly accommodate angular dependence and account for coherent noise and artifacts.

Azimuth-integrated angle gathers are functions of depth and angle that can be modeled as

$$G_{ij}=\int \mathscr{I}(y_j;\theta_i,\psi)\,d\psi = g(\theta_i,y_j)+\varepsilon_{ij},$$

where $g(\theta_i, y_j)$ is the ideal noiseless angle gather at angle θ_i and (radial) depth y_j, and ε_{ij} are random noise terms that are usually coherent. Our goal is to estimate $g(\theta_i, y_j)$, which should be a singular function of depth (smoothed by the convolution with some pulse) that varies slowly with angle (including a dilation). To achieve this goal, we make additional assumptions on g and the noise. For g, we use a flexible functional representation in terms of some basis functions φ_ℓ in depth. The noise is modeled as a sum of a discrete harmonic process and white noise. The angle gather is thus modeled as

$$G_{ij} = g(\theta_i, y_j) + h(\theta_i, y_j) + \varepsilon_{ij}, \qquad i = 1, \ldots, a,\ j = 1, \ldots, b. \tag{3.6}$$

Here, $h(\theta_i, y_j) = \sum_{k=1}^{m} h_k(\theta_i)\, \phi_k(y_j)$ represent a harmonic process that models coherent noise through sinusoid functions $\phi_k(y)$ in the depth coordinate. For a fixed angle θ_i, the coefficients $h_k(\theta_i)$ are assumed to be uncorrelated, but $h_k(\theta_i)$ and $h_k(\theta_j)$ may be correlated for $\theta_i \neq \theta_j$. The errors ε_{ij} are independent Gaussian random variables with mean zero and variance σ^2. A refined model for the noiseless IG that include angle dilation is given by $g(\theta_i, \alpha_i y_j)$, which replaces $g(\theta_i, y_j)$.

To estimate g, we need estimates of the covariance parameters of the random effects (i.e., h_k), including sinusoid frequencies determining the ϕ_k. We will consider three particular cases of (3.6) that can be analyzed efficiently. We use the same notation for random errors, ε_{ij}, for the different models below.

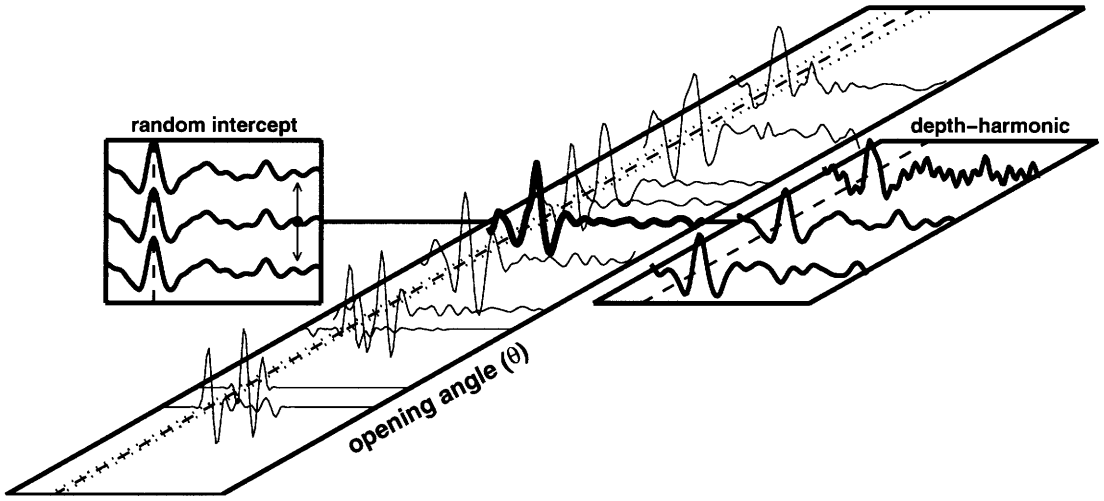

Figure 3-4: An illustration of how the noise adds to an image trace in the random intercept (left), depth-harmonic (right), and angle-dilation depth-harmonic (center, indicated by dotted dilation curves) models. In the center we plot actual image gathers for different opening (or scatter) angles θ. The random intercept model estimates the best 'common component' in theses traces (left panel surface, middle trace), which is allowed to move up-and-down in order to find the best fit to the gathers. The depth-harmonic models detect and correct for spurious oscillations in radial direction: at the right, the middle trace depicts the best image estimate, and the adjacent traces show this trace with a (randomly chosen) harmonic either added to or subtracted from it. The third model, the angle-dilation depth-harmonic model, is similar to the depth-harmonic model proper, but it also accounts for the angle-dilation (depicted with dotted lines in the center panel) that results from the geometry of the GRT image problem at hand (see Paper I).

Random-Intercept Model

For weak angle dependence we can approximate $g(\theta_i, y_j)$ in (3.6) by

$$g(\theta_i, \alpha_i y_j) \approx b_i + g(\alpha_i y_j). \tag{3.7}$$

We start with the simple model in (3.7) that uses a single parameter to aggregate higher order approximation terms and which also ignores the angle-dilation effect:

$$G_{ij} = b_i + g(y_j) + \varepsilon_{ij}, \qquad i = 1,\ldots,a,\ j = 1,\ldots,b, \tag{3.8}$$

where the b_i are Gaussian $N(0,\sigma_s^2)$. Clearly (3.8) is a particular case of (3.6) obtained by setting $g(\alpha_i y_j) = g(y_j)$ and $h(\theta_i, y_j) = b_i$.

We estimate $g(y)$ by minimizing a penalized sum of squares similar to (3.5) *Robinson* [1991]

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(G_{ij} - g(y_j) - b_i)^2 + \frac{\sigma^2}{\sigma_s^2}\sum_{i=1}^{a} b_i^2 + n\lambda J(g), \tag{3.9}$$

where $J(g) = \int g'(y)^2 \mathrm{d}y$ quantifies the smoothness of g and the smoothing parameter λ controls the trade-off between the goodness-of-fit and smoothness of g. Since the parameters λ and σ^2/σ_s^2 provide a large family of possible estimates, one has to choose values that lead to good estimates as measured by the goodness of fit and smoothness of the estimate. To choose the parameters we use a method known as generalized cross-validation (GCV), a method that is widely used in Tikhonov regularization (*Wahba* [1990]). The basic idea is as follows. We start by fixing the angle (optimal parameters are chosen for each fixed angle, i.e., each fixed i). For each j and choice of parameters we obtain estimates $\widehat{b}_{i,-j}$ and $\widehat{g}_{-j}(y_j)$ of b_i and $g(y_j)$, respectively, using all the data but G_{ij}. That is, we predict the value of G_{ij} using the data $G_{i\ell}$ for $\ell \neq j$ (the notation $-j$ means that the jth observation has been deleted from the data set). These estimates in turn provide the prediction $\widehat{G}_{ij,-j} = \widehat{b}_{i,-j} + \widehat{g}_{-j}(y_j)$ of G_{ij}. It has been shown (see *Wahba* [1990]) that by using the deleted estimates one obtains better estimates of the prediction error which is what we want to minimize. Ordinary cross validation chooses the parameters that lead to the smallest value of $\sum_j (G_{ij} - \widehat{G}_{ij,-j})^2$.

In the Appendix this is discussed in more detail.

We use Bayesian confidence intervals as a benchmark for assessing the uncertainties in the estimate of $g(y)$. We calculate posterior means and variances of $g(y)$ under the Bayes model and use them to construct 95% and 99% confidence interval estimates. An important feature of these intervals is that they approximately have the correct across-the-function coverage probability. That is,

$$\frac{1}{n}\sum_{i=1}^{n} P\left[g(y_i) \in C_\beta(y_i)\right] \approx \beta,$$

where $C_\beta(y_i)$ is the Bayesian confidence interval and β is the coverage (we use $\beta = 0.95$ and 0.99). Although these intervals were originally derived under the independence assumption, there are straightforward extensions that can be used with correlated data. See *Wahba* [1983], *Nychka* [1988] and *Ma* [2003] for details.

Depth-Harmonic Model

Now we take a different approach to model the noise in the data. Instead of combining all the effects in a simple random effect, we model the harmonic process explicitly. We extend the random-intercept model (3.8) to

$$G_{ij} = \sum_k a_{ik}\cos(\omega_{ik} y_j + \beta_{ik}) + b_i + g(y_j) + \varepsilon_{ij}, \tag{3.10}$$

where the coefficients a_{ik} and phases β_{ik} are random and independent, and the frequencies ω_{ik} are fixed but unknown.

Since the harmonic component is fixed for each angle, we may think of it as a sinusoid signal contaminating $g(y)$. Thus, although it could be included as a random effect, it is

easier to fit and remove it from each trace. This fitting is done using ARMA time series techniques described in the Appendix. The corrected G_{ij} is then analyzed using the simpler random-intercept model.

Angle-Dilation Depth-Harmonic Model

To account for the angle-dilation (see Paper I) as well as a harmonic process of coherent noise, we use the following modification of (3.10) to fit the data:

$$G_{ij} = \sum_k a_{ik} \cos(\omega_{ik} y_j + \beta_{ik}) + b_i + g(\alpha_i y_j) + \varepsilon_{ij}, \qquad i = 1,\ldots,a,\ j = 1,\ldots,b, \quad (3.11)$$

where α_i is the angle-dilation effect. As the remaining noise ε_{ij} is assumed white, the term b_i has to account for higher order angle-dilation effects not modeled by $g(\alpha_i y_j)$. The lack of structure in the boxplots of the residuals shows that this assumption is reasonable.

The parameters are estimated in two stages: The first stage is the same harmonic correction as before. In the second stage the parameters α_i and b_i are simultaneously estimated with other parameters through penalized least squares (see Appendix for more details).

In Figure 3-4, we illustrate the properties of the different mixed effects models. In the center, we plot an actual image gather (as in Figure 3-10). On the left we illustrate the random intercept effects: the best image estimate (middle curve) is randomly shifted up or down. On the right we illustrate the depth harmonic effects: a randomly chosen harmonic is added (back) or subtracted (front) to the best image estimate (middle curve). The angle-dilation of the third model is indicated by two parting (dotted) curves in the image gather in the center plot.

3.3.3 Model validation

The adequacy of a statistical model is checked by comparing, at each scattering angle, the data G_{ij} to the estimates $\widehat{G}_{ij}$ under the corresponding model. We study plots of the residuals defined as $e_{ij} = G_{ij} - \widehat{G}_{ij}$ normalized by the model estimate $\widehat{\sigma}$ of the noise standard deviation σ (i.e., $e^*_{ij} = e_{ij}/\widehat{\sigma}$). Boxplots of the angle-dependent residuals provide information about the shape of the residual distribution (that is, its median, interquartile range and presence of outliers). For example, inadequacies in the background elastic properties would lead to biased residuals not centered around zero, with a magnitude that depends on how the structure is sampled (i.e., the scatter angle). Residuals centered at zero without discernible patterns indicate that the estimate model is reasonable. Examples of boxplot analysis are given below.

3.4 Study: Synthetic *ScS* data

In **Chapter II** we used WKBJ synthetics to explore certain aspects of image recovery with GRT, including its performance on random, additive noise in the data. Here we use the same synthetic data (Figure 3-5, left) to test the performance of our statistical inference method on non-random noise in the image gathers. The data are generated from a 1-D, spherically symmetric background model with a wavespeed increase at 150 km above the CMB (Figure 3-5, right). Figure 3-5 also illustrates the distance ranges associated with 'narrow' and 'wide' scattering angle data. At narrow angles, the signal from the top of D'' is a precursor to *ScS*, see inset in Figure 3-5 (left), but owing to a small reflection coefficient it has a small amplitude and will be difficult to detect in raw data. At larger scatter angles the reflection becomes stronger, and it arrives closer to and, eventually, after the direct *ScS*

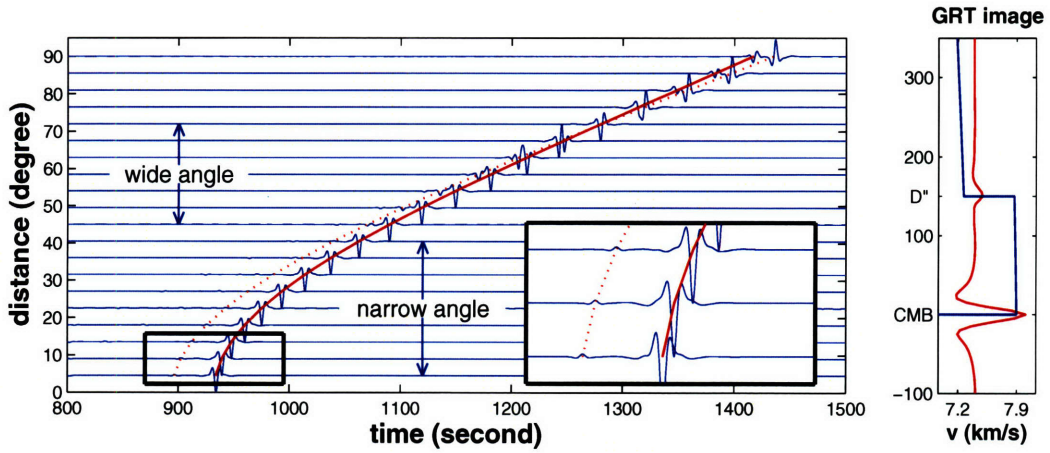

Figure 3-5: Synthetic data and illustration of the robustness of the GRT in the presence of random noise in the data. Left: synthetic (WKBJ) records of ScS and S_dS, that is the signal from a weak reflector at 150 km above the CMB. The inset shows the weak precursory energy, for narrow angle reflections. At wide angles this reflection becomes stronger and crosses over with, and eventually arrives in the coda of direct ScS. Red lines are theoretical travel time curves for ScS (solid) and S_dS (dashed). Right: GRT image trace (solid red line) constructed from the synthetic data shown on the left and the wavespeed profile used to generate the synthetic waveforms (blue curve).

arrival.

We explore how each of the statistical models described in the previous section handles the following constituent effects in the estimation: highly irregular sampling (due to the actual, uneven distribution of stations and events), random additive noise in the image gathers (for example, due to scattering not explained by the single scattering approximation), and random harmonic noise in the image gathers (for example, due to isolated spectral components generated by the imaging). For each test, we show the GRT image (that is, the substack over the image gathers for different scatter angle) both for the narrow and the wide angle data. However, the statistical results discussed and shown here are based on data associated with the large scattering angles (in Figures 3-6-3-9 the GRT image restricted to large scattering angles is indicated by a vertical arrow). In Figures 3-6-3-9, the black curves represent the statistical estimate of $g(y)$ and the light green bands depict the

95% Bayesian confidence intervals. For comparison, we also plot the true model (red line).

3.4.1 Effects of 'station-event' sampling

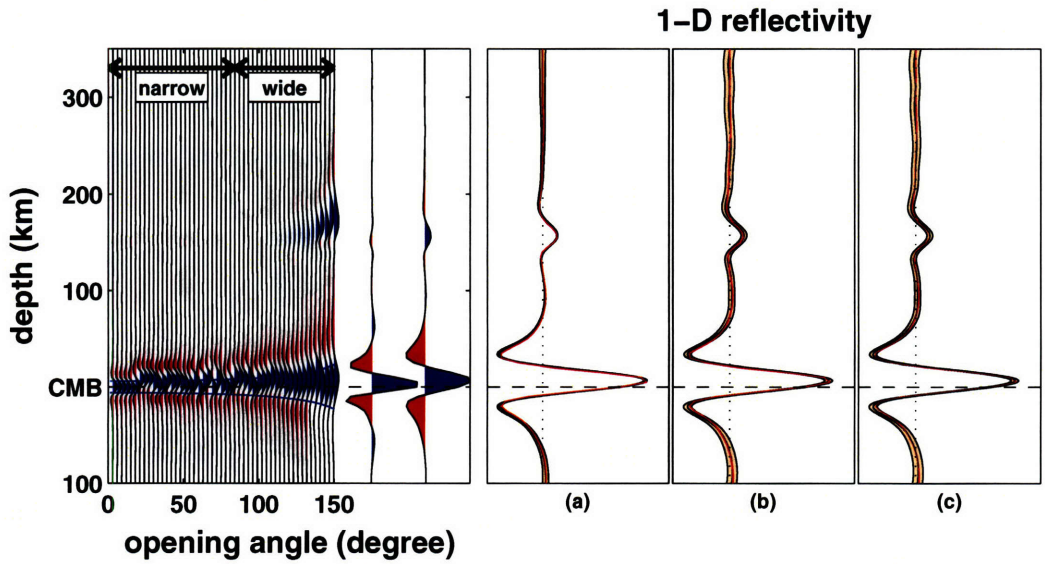

Figure 3-6: Effect of the acquisition imprint, that is, source-receiver distribution. The left panel shows image gathers for a range of opening (scatter angles) and, right next to them, the result of the GRT restricted to narrow (left) and wide (right, with arrow) scattering angles. The narrow and wide angle data are illustrated in Figure 3-5. The three panels on the right show estimates of the reflectivity profiles using the random intercept (a), depth-harmonic (b), and angle-dilation depth harmonic (c) statistical models. The bands around the mean estimate (black line) correspond to 95% Bayesian confidence intervals. For reference, the red line depicts the true signal.

First, we analyze the effect of acquisition imprint. Acquisition is here viewed as the spatial distribution of stations and events. If the stations and events were regularly spaced, aliasing would be an immediate concern. This is not quite the situation, but the effects due to the actual acquisition, locally, are related. In Figure 3-6 we illustrate how our approach treats these effects, using a realistic acquisition geometry. The image gathers on the left reveal significant scatter associated with non-random sampling. The two traces directly

to the right of the image gathers illustrate the action of the GRT restricted to small scattering angles (left trace) and large scattering angles (right trace, with arrow). The image is distorted in that the phase of the two events associated with the reflectors has changed. The panels on the right show the mixed-effects model estimates ((a): random intercept, (b): depth harmonic, (c): angle-dilation depth harmonic). The black line in the middle of the 95% confidence band shows the mean estimate and the red line shows the signal to be recovered. We see that the three different methods are all capable of undoing the effects of irregular event-receiver sampling, and restore the phase.

3.4.2 Effects of noise in the image gathers

In Figure 3-7 we have added random noise (energy not explained by the single scattering approximation) to the image gather obtained from the noise-free data in Figure 3-5 (left), subjected to a realistic acquisition geometry. The GRT image traces (small and large scattering angles) directly to the right of the image gather are strongly affected by the noise; in fact, the top reflector is no longer visible, whereas multiple weak, false reflectors appear. The random-intercept model is capable of providing a clean estimate of the image (Figure 3-7A), including the top reflector. The estimate is consistent with the true model (red line). Because the added noise did not have a harmonic component, there is no perceptible difference in performance between the random intercept model and either of the harmonic models (Figure 3-7B-C).

In Figure 3-8 we have added random noise with a harmonic component to the image gather obtained from the noise-free data in Figure 3-5 (left) again subjected to a realistic acquisition geometry. The GRT image traces (small and large scattering angles) directly to the right of the image gather are both affected by the noise; the top reflector has disap-

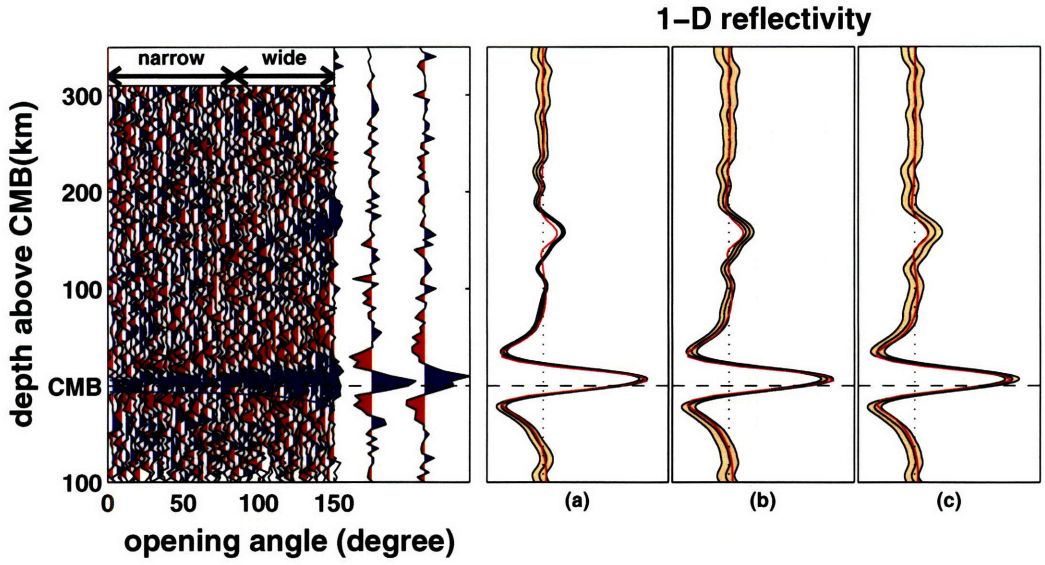

Figure 3-7: Estimation in the presence of random noise in the image gathers. For this purpose, we added random noise to the image gathers shown in Figure 3-6, that is, for a realistic acquisition imprint. Both the narrow and wide angle GRT stacks reveal significant jitter and neither suggests the presence of a reflector at 150 km above CMB. In contrast, the top reflector is detected in the statistical estimates, even though the images of it are slightly distorted compared to the true model (red line). In the absence of a harmonic noise component, all three models detect the contrast at the CMB.

peared, and more spurious signals are visible. If untreated, the latter can produce spurious events in the image profiles. The random-intercept model estimate (Figure 3-8A) is beginning to have problems recovering the top reflector as seen by the true signal (red line) being mostly outside the error band. We also see remnant oscillations of the harmonic components that the random-intercept model could not remove. The depth-harmonic models (Figure 3-8B-C) clearly perform better; many of the harmonic oscillations have been removed and the the estimate (black line), with the error bars, is consistent with the true signal.

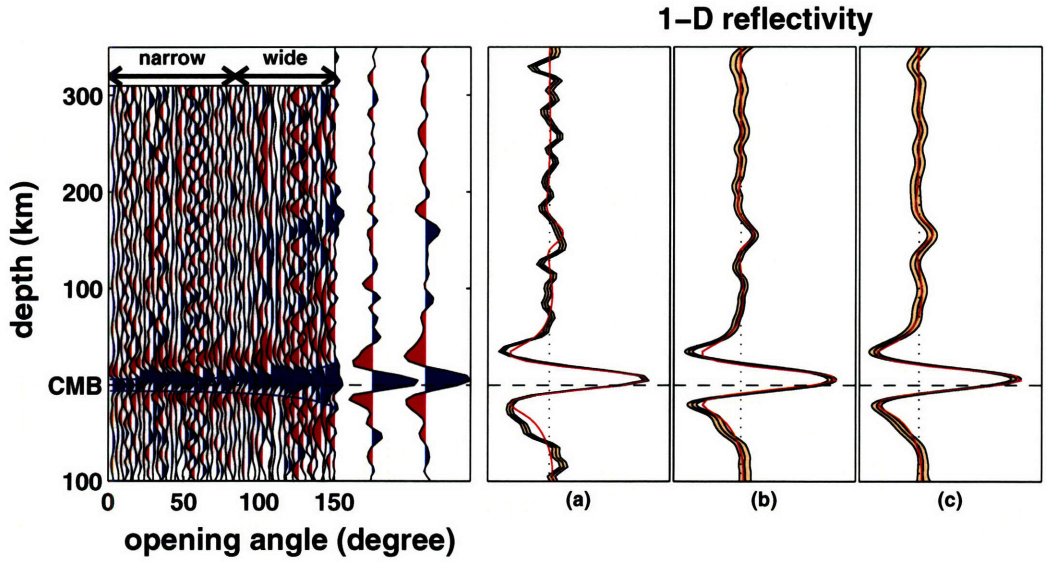

Figure 3-8: Same as Figure 3-7 but now for harmonic instead of random noise. While performing better than the GRT without statistical estimation, the random-intercept model begins to break down in the presence of a strong harmonic component in the noise structure; indeed, trace (a) reveals much spurious structure and the estimates of the contrasts at CMB and 150 km above it deviate significantly from the actual model (red line). Both harmonic models – traces (b) and (c) – retrieve the model well (that is, within 2 σ they are the same as the true model), but the angle-dilation depth harmonic model (c) performs slightly better than the standard depth harmonic (b).

3.4.3 Effects of an inaccurate wavespeed model

An important assumption in the application of the GRT as developed in Paper I (**Chapter II**, *Wang et al.* [2006]) is that we have a reasonable estimate of the elastic properties, say the wavespeed, of the background. Incorrect properties of the back ground model would produce (scatter) angle-dependent artifacts in the image gathers. We recall that the model validation operates on the pre-stack image gathers, which helps us recognize and correct such artifacts. In order to demonstrate this premise, we again form GRT images using the noise free, WKBJ modeled data (Figure 3-5, left), but now we assume background properties that are different than those used to produce the synthetic data. The perturbed

wavespeed model assumed in the GRT and the one used to produce the data are shown in the inset to Figure 3-9.

The GRT image traces are affected by the use of the incorrect model. In fact, for small scattering angles, the top reflector has become almost invisible. Figure 3-9A-C demonstrate that the different mixed-effects models are all capable of recovering the image of both reflectors. However, we observe a clear deterioration in spatial resolution, and the depth estimate of the reflectors have decreased. Tests like these show that not knowing the background wavespeeds well does not prevent us from detecting interfaces, but it may produce artificial boundary topography.

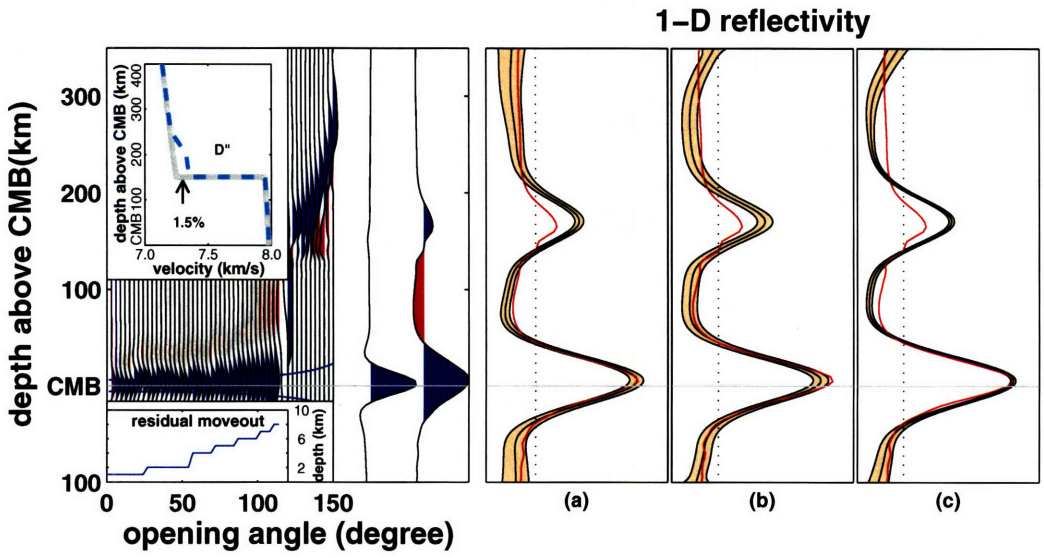

Figure 3-9: Same as Figure 3-6 but for image estimation with an inaccurate wavespeed model. As before, we use synthetic data generated from a model with a simple step-wise increase in wavespeed at 150 km above the CMB (solid gray line in inset, top left), but as back ground model for the GRT we assumed a model with a slightly different D'' structure (blue dashed line in inset, top left).

3.5 Imaging the lowermost mantle beneath Central America

We illustrate the proposed methodology with a study of inference of singularities over a $50^\circ \times 50^\circ$ patch of the core mantle boundary (CMB) beneath Central America, using the broad-band data from Paper I (**Chapter II**, *Wang et al.* [2006]). The data selection and preprocessing is explained in Section 2.4.1.

3.5.1 Statistical analysis of an *ScS* common image-point gather

We select a particular location and image gather, and apply the analysis of the previous section. In Figure 3-10 we show the gather for large scattering angles (left), the associated GRT stack (first trace to the right), and an image estimate with the depth-harmonic model (second trace to the right). The blue lines indicate the dilation derived from the GRT (see **Chapter II**). The (linear) GRT stack contains multiple reflectors, but the (nonlinear) image estimate suggests that not all of them are real. Indeed, we note the general difference in appearance of the statistical estimate compared to the GRT stack. The estimate shows the CMB – symmetric, zero phase, unlike in the GRT stack – as well as a clear indication of a reflector about 240 km above it – within 95% confidence level. This suggests that the location and width of the two main peaks are consistent with the model of the lowermost mantle and the data resolution bounds used. The image estimate of the CMB appears on the coarsest resolution viewed with respect to the expected dilation in the image gather (blue lines). One may argue for the presence of a weak reflector about 115 km above the CMB, marked as X in Figure 3-10, but it remains within the error bars associated with 95% confidence.

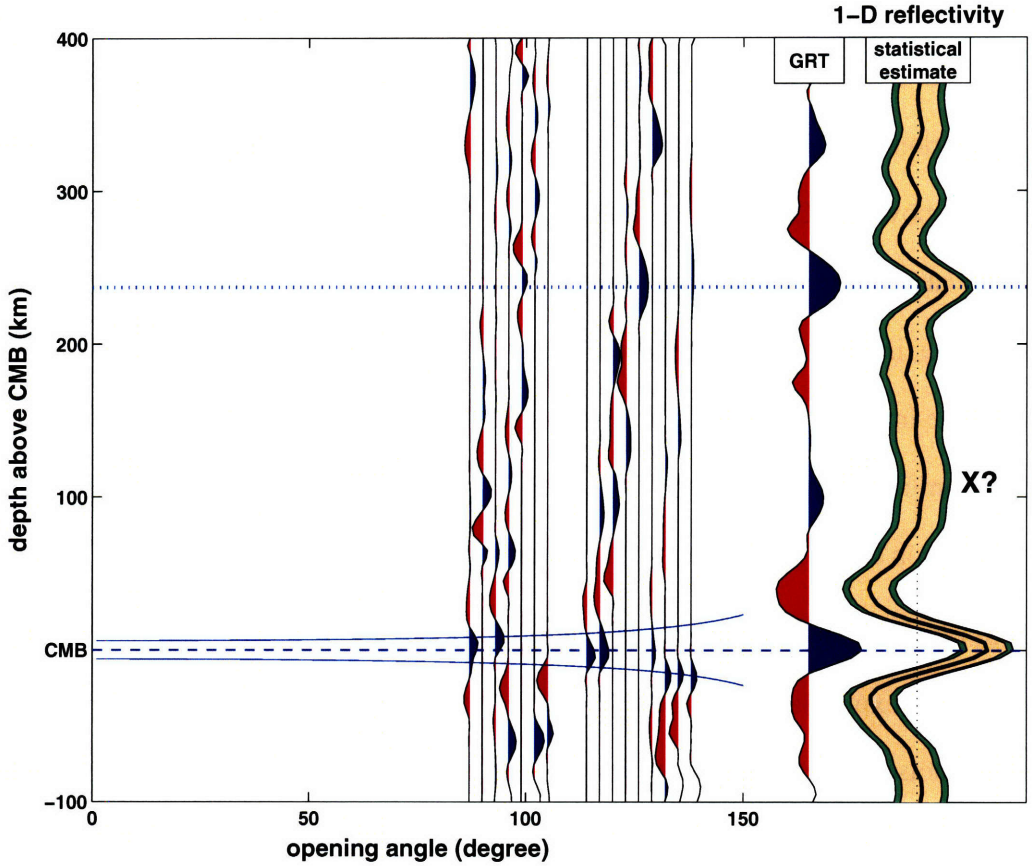

Figure 3-10: A typical image gather associated with large scattering angles, and a GRT stack versus an image estimation based on the depth-harmonic model. The bands of two colors in the depth-harmonic estimate correspond to 95% and 99% confidence error bars. The GRT stack suggests a scatterer ~100 km above the CMB , but the statistical estimate only shows a weak, broad structure (marked by X), which may suggest that it is not (statistically) significant.

3.5.2 2-D image profile

We form an 2-D image profile (which in the context of this paper represents a synthesis of the mean statistical estimates) by repeating the analysis in the previous subsection for a large number of CMB points along a 2500 km long great circle path. We also address the validity of the different mixed-effects models – for this purpose, we have selected 6 locations in the profile, indicated by (a) through (f).

In Figure 3-11 we present the results collectively. The top image profile is obtained with the random-intercept model, whereas the bottom two image profiles are obtained with the angle-dilation depth-harmonic model. The bottom profile uses a dual color scale to display structure above the CMB (color) at an amplification by a factor of five compared to the structure around the CMB proper (grey tone). Superimposed on the bottom panel are the statistical estimates at the mentioned 6 locations. Below these profiles, for these locations, we plot the image estimates accompanied with the 95% Bayesian confidence intervals, for the random-intercept model (top row), the depth-harmonic model (middle row), and the angle-dilation depth-harmonic model (bottom row).

For location (a) we carry out a detailed validation of the entire image gather. The corresponding boxplots are shown on the lower left. Each mixed-effects model accounts for coherent noise in the traces in its own way. For the three mixed-effects models introduced in this paper, we provide illustrations of the within-angle residuals e_{ij} as defined in Section 3.3.3. Each boxplot shows a box bounded by the 25% and 75% percentiles of the residual distribution (for e_{ij} with j fixed). This interval provides information about the spread of the distribution. The line within the box indicates the median, and the whiskers provide information about the symmetry and tails of the distribution. For all three models we observe that the residuals are close to centered at zero, which is an indication that the mean estimates seem reasonable. Their variability changes with angle, however, which is a measure of the degree of adequateness of the models. Such unmodelled variability may have an effect in uncertainty estimates.

As was the case in the examples presented in Figures 3-6-3-9, at first glance the three statistical models may seem to perform similarly. Upon closer inspection, however, we can see differences. We compare the models by viewing how the boxplots of residuals are

scattered around zero. The depth-harmonic models show less structure and fewer outliers than the random-intercept model, which may indicate that the they provide a better fit for our data. Furthermore, the spatial (depth) resolution has improved slightly (the peaks are sharper) by including angle-dilation. With the depth-harmonic models, however, we still see different dispersion across angles that may affect uncertainties. This may be caused by angle dependence in the depth of the singularities, an indication of errors in velocity model, perhaps associated with anisotropy in the lowermost mantle.

In Figure 3-12 we compare the best image estimate with the profile of Paper I (**Chapter II**, *Wang et al.* [2006])(also shown in Figure 3-2). Visually, the differences between results of GRT imaging with or without statistical analysis are fairly small. This is reassuring because it demonstrates that the structures are constrained by the data and not introduced by the statistical analysis. In detail, however, the effects of the statistical inference are visible (as marked). Statistical inference and validation lead to the suppression of rapid oscillations that are not required by the data, resulting in a smoother image. Furthermore, most of the seismic events in the image reveal more lateral continuity than in the original result presented in Paper I (**Chapter II**, *Wang et al.* [2006]). As important as the visual effects, however, the statistical analysis presented here provides a means for estimating uncertainty, which will be key for subsequent interpretation of the structures that are visible in the image profiles. As an example, in Figure 3-13 we show the structures in Figure 3-12B at various levels of probability by muting structure that does not exceed the width of a particular confidence interval at that location. Figure 3-12B appears to be significant at 68% (1σ) confidence level, but – as expected – only a few structures appear significant at 95% confidence. This example demonstrates how our analysis can be used not only to detect structure but also to identify and isolate the most robust features.

It should be noted that in the application introduced here 'mixed-effects' modeling is only used to estimate an optimal radial (reflectivity) profile at a particular image point at the CMB. In a similar vein, the lateral coherence in Figures 3-12 and 3-13 can be enhanced by applying the statistical models to the horizontal (space) distance parameter, but such image processing is not done here.

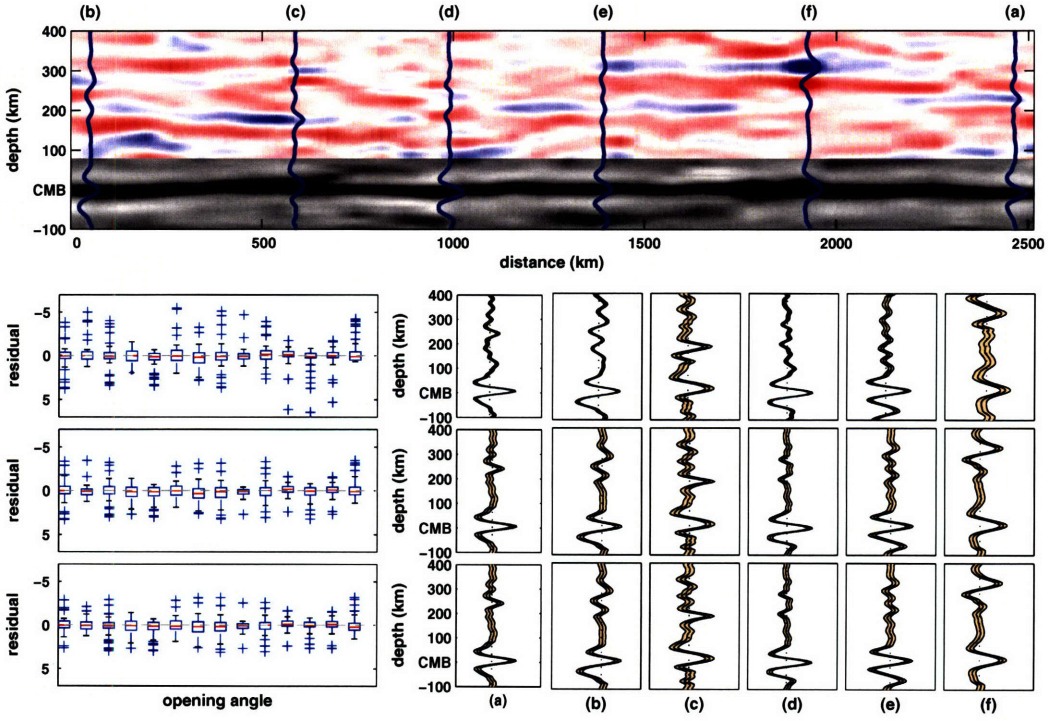

Figure 3-11: Top: 2-D image profile resulting from processing with angle-dilation) depth harmonic model. A dual color scale is used to display structure above the CMB (color) at an amplification by a factor of five compared to the structure around the CMB proper (grey tone). Superimposed are the statistical estimates at six arbitrary positions. Below the image profiles, from left to right we show the boxplots used in the validation (see Section 3.3.3) and the statistical estimates for six locations along the 2D image profiles. Top row: random-intercept model; Middle row: depth harmonic model; Bottom row: angle-dilation depth harmonic model.

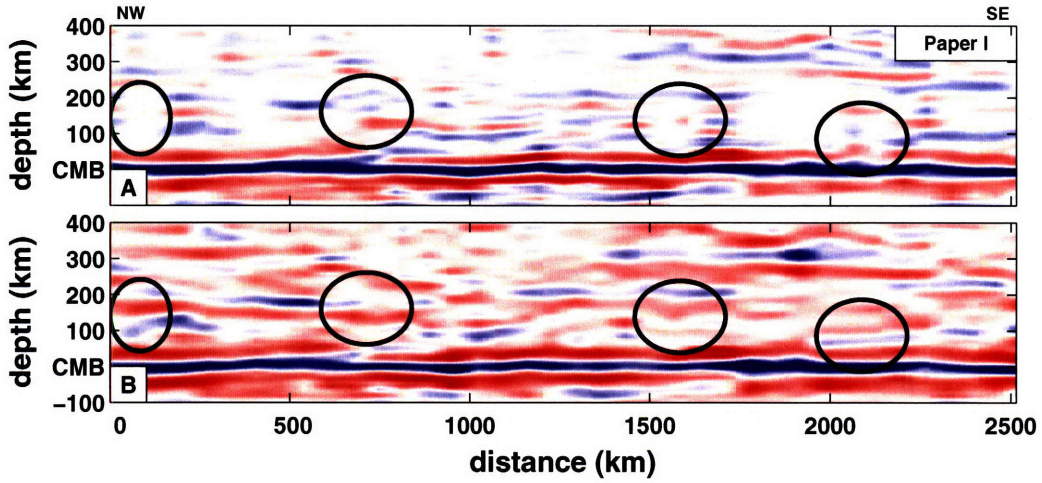

Figure 3-12: Comparison of original 2-D image profile (presented in **Chapter II**), at the top, through application of the (angle-dilation) depth harmonic model, at the bottom. There is no exaggeration in the vertical scale.

3.6 Discussion and concluding remarks

The aim of the research described here and in Paper I (**Chapter II**, *Wang et al.* [2006]) is to develop a novel approach to 3-D seismic imaging of the lowermost mantle using concepts from inverse scattering and modern statistics. In particular, we aim to exploit both the wide angle reflections, which are used in most modeling studies because the large reflection coefficient produces arrivals that can be recognized in raw data, but which produce very poor depth resolution, as well as the narrow angle reflections, which are associated with weak reflections but which produce superior depth resolution. In fact, the recognition – and use – of this angle dependence of radial resolution, which is reflected in the angle-dilation discussed in Paper I (**Chapter II**, *Wang et al.* [2006]), allows high resolution imaging and multi-scale analysis of weak interfaces in Earth's deep interior.

The combined use of the generalized Radon transform (GRT) and the mixed-effect statistical inference presented here exploits the redundancy in the broad band data and al-

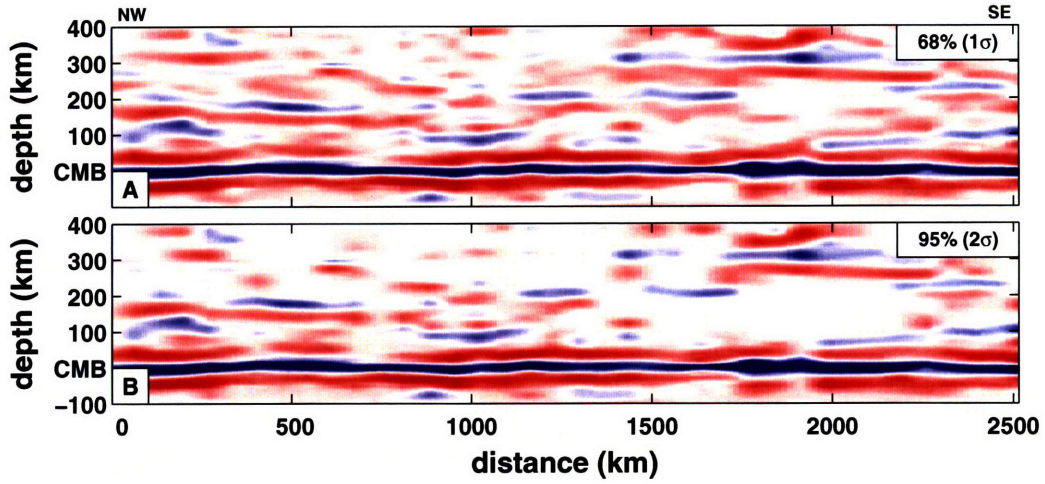

Figure 3-13: Panels (A) and (B) show, respectively, the structure in Figure 3-12B that is significant at the 68% and 95% confidence levels. This display is obtained by keeping only the features at each depth that are significantly different from zero at the chosen confidence level and muting the rest.

lows the transformation of large volumes of global network data to statistical estimates and quantitative analysis of elastic singularities (such as discontinuities). Through the use of mixed-effects models we can distinguish between and deal with true random noise in the data, random noise in the GRT images due to scatter that is not considered in the GRT theory, and artifacts between the assumed and real wavespeeds of the medium. The examples discussed in Section 3.4, and illustrated in Figures 3-6-3-9, demonstrate that this does, indeed, enhance the ability to extract weak signal from noisy data. In some cases the changes may seem subtle, but given the overall difficulty of – and interest in – imaging D'' structure even small improvements are very welcome. Moreover, our statistical approach enables us to estimate uncertainties – in a Bayesian context – so that we can know with some confidence whether imaged structures are real or (statistically) insignificant. Also, in the future, it will allow a more rigorous analysis of the regularity (including such scaling parameters as roughness, sharpness, type of onset, and scale-dependent impedance contrast) of the

transitions in elastic parameters detected here. This information will be gleaned from the wavenumber dependence of image gathers with scattering angle.

In contrast to methods based on forward modeling, our method imposes few *a priori* assumptions about the geometry and nature of the structures that we attempt to image. Indeed, the structures revealed in the images are entirely controlled by the data themselves. The GRT only assumes that at a predefined image point the interfaces are contiguous, but in practice this is not a serious restriction. Furthermore, we assume to have a reasonable estimate of, or reference for, the smooth elastic properties of the medium. The latter assumption is not taken lightly; indeed, our statistical analysis quantifies the extent to which it is satisfied and, moreover, enables us to model and adjust for this type of 'noise' in the image gathers. This is possible because of the pre-stack nature of our analysis: the statistical inference operates on image gathers at a large range of scatter angle, and not on the stack – or GRT image – itself.

In the general context of mixed-effect models, we have considered three particular cases: (i) a random-intercept model, which assumes that the noise in the data – or in the GRT images – is white, (ii) a depth harmonic model, which identifies and removes spurious oscillations in radial direction, and (iii) an angle-dilation depth harmonic, which considers the dilation – and concommittant reduction in radial resolution – with increasing scatter angle that is inherent in the GRT imaging under consideration. With synthetic data we demonstrate that without a rigorous statistical approach, subtle wavespeed contrast may be overlooked – or spurious ones introduced – with a traditional GRT.

We found that the performance of the random-intercept model is in many cases similar to that of the depth-harmonic and angle-dilation depth-harmonic models. The good performance of this simple model is largely due to the data driven nature of the model fitting

strategy. In particular, the use of a modified cross-validation procedure to select the tuning parameters of the model (see appendix for details) leads to robust optimal smoothing of the data that deviates significantly from the simple structure of the noise. On the other hand, the estimation procedure of the depth-harmonic models (see appendix for details) could further reduce the noise level through integration of cross-validation with estimation procedures similar to those studied in *Guo* [2002]; this is a topic of current research [*Ma and Zhong*, in preparation].

We demonstrated the feasibility of the method with an application to the *ScS* wavefield that reflects off the CMB beneath Central America. The data are described in Paper I (**Chapter II**, *Wang et al.* [2006]). Visual inspection suggests that the application of the mixed-effects models leads to an image with better lateral definition of interfaces (in addition to information as to the statistical significance of each scatterer), but that the different mixed-effect models used yield rather similar results. This may indicate that the level of harmonic noise in the real data is low or that the coherent noise is not truly harmonic. However, quantitative model validation suggests that the (angle dilation) depth harmonic models, which suppress spurious oscillations in the 2-D image profiles, produce slightly better data fits and uncertainty estimates.

The image produced by ~80,000 broad-band *ScS* data (e.g., Figure 3-11) reveals strong contrasts in elastic parameters at about 0, that is, at the depth of the CMB and, locally, near 150 and between 270-320 km above it. It is tempting to interpret the latter as the (fragments of the) 'top' of the so called D'' layer. Changes in elastic parameters near this depth have been the subject of many seismological studies (e.g., *Lay and Helmberger* [1983a]; *Tromp and Dziewonski* [1998]; *Sidorin et al.* [1999]; see *Garnero* [2000] for a comprehensive review), but there is as yet no consensus about this transition and its radial

and lateral extent. Our results suggest that the structure may not be (laterally) continuous. The images also reveal significant scatter in between the presumed top of the D'' layer and the CMB proper. Collectively, our observations suggest that the D'' region is more complex than expected from models based on simple perovskite to post-perovskite transitions. This implication will be explored elsewhere (e.g., *Van der Hilst et al.* [2007]).

3.7 Acknowledgments

We thank the Associate Editor and two anonymous reviewers for constructive comments, which we used to improve the manuscript. The data were retrieved from the Data Management Center of the Incorporated Research Institutions for Seismology. We thank the Collaborative Mathematics and Geosciences program of the US National Science Foundation for financial support (under grant EAR-0417891).

3.8 Appendix: Estimating the model parameters

(i) Random-Intercept Model

In the standard formulation of penalized least squares regression, the minimization of (3.9) is performed in a reproducing kernel Hilbert space $\mathscr{H} \subseteq \{g : J(g) < \infty\}$ in which $J(g)$ is a square semi-norm, and the solution resides in the space $\mathscr{N}_J \oplus \text{span}\{R_J(y_i, \cdot) : i = 1, \ldots, d\}$, where $\mathscr{N}_J = \{g : J(g) = 0\}$ is the null space of $J(g)$, and $R_J(\cdot, \cdot)$ is the so-called reproducing kernel in $\mathscr{H} \ominus \mathscr{N}_J$. The solution has an expression

$$g(y) = \sum_{v=1}^{k} \beta_v \phi_v(y) + \sum_{j=1}^{d} c_j R_J(y_j, y), \tag{3.12}$$

where $\{\phi_\nu\}_{\nu=1}^{k}$ is a basis of $\mathcal{N}_J$ [1].

Substituting (3.12) into (3.9), one minimizes

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(G_{ij}-\phi(y_j)^t\beta\ -\varphi(z_j)^t\mathbf{c}-b_i)^2+\frac{\sigma^2}{\sigma_s^2}\sum_{i=1}^{a}b_i^2+n\lambda\,\mathbf{c}^t\mathbf{Q}\mathbf{c} \tag{3.13}$$

with respect to $\beta\ =(\beta_1,\beta_2,\cdots,\beta_k)^t$, $\mathbf{c}=(c_1,\ldots,c_d)^t$ and $\mathbf{b}=(b_1,\ldots,b_a)^t$, where $\varphi(y)=(R_J(y_1,y),\ldots,R_J(y_d,y))$, and $\mathbf{Q}$ is $d\times d$ with the (j,k)th entry $R_J(y_j,y_k)$. Estimates of β, $\mathbf{c}$ and $\mathbf{b}$ are obtained by setting to zero the derivatives of (3.13) with respect to $\mathbf{c}$ and $\mathbf{b}$. The minimizers of (3.13) are solutions to the normal equation,

$$\begin{pmatrix}\mathbf{S}^t\mathbf{S} & \mathbf{S}^t\mathbf{R} & \mathbf{S}^t\mathbf{M}\\ \mathbf{R}^t\mathbf{R} & \mathbf{R}^t\mathbf{R}+(n\lambda)\mathbf{Q} & \mathbf{R}^t\mathbf{M}\\ \mathbf{M}^t\mathbf{S} & \mathbf{M}^t\mathbf{R} & \mathbf{M}^t\mathbf{M}+\tau\mathbf{I}\end{pmatrix}\begin{pmatrix}\beta\\ \mathbf{c}\\ \mathbf{b}\end{pmatrix}=\begin{pmatrix}\mathbf{S}^t\mathbf{G}\\ \mathbf{R}^t\mathbf{G}\\ \mathbf{M}^t\mathbf{G}\end{pmatrix}, \tag{3.14}$$

where $\mathbf{G}=(G_{11},\ldots,G_{ab})^t$, $\mathbf{S}$ is $n\times k$ with the (i,ν)th entry $\phi_\nu(y_i)$, $\mathbf{R}$ is $n\times n$ with the (i,j)th entry $\varphi_j(y_i)$, $\mathbf{M}$ is $n\times a$ block diagonal identity matrix, $\mathbf{Q}$ is $n\times n$ with the (j,k)th entry $J(\varphi_j,\varphi_k)=R_J(v_j,v_k)$, and $\tau=\sigma^2/\sigma_s^2$ and $\mathbf{I}$ is identity matrix. The normal equation of (3.14) can be solved by a Cholesky decomposition followed by backward and forward substitutions. Possible singularity of the matrix can be properly handled through pivoting in Cholesky decomposition; see, e.g., *Golub and Van Loan* [1989] and *Kim and Gu* [2004] for details.

[1] For example, take a function g defined on $[0,1]$ with $J(g)=\int(g'')^2dy$ and $\mathcal{N}_J=\{g:g(y)=\beta_1+\beta_2 y\}$. In this case we get the popular cubic splines and the reproducing kernel is $R_J(x,y)=k_2(x)k_2(y)-k_4(x-y)$, where $k_\nu=B_\nu/\nu!$ are scaled Bernoulli polynomials. See *Wahba* [1990] and *Gu* [2002] for comprehensive treatments of the subject.

The fitted values $\widehat{\mathbf{G}} = \mathbf{S}\widehat{\beta} + \mathbf{R}\widehat{\mathbf{c}} + \mathbf{M}\widehat{\mathbf{b}}$ can be written as $\widehat{\mathbf{G}} = \mathbf{U}(\lambda,\tau)\mathbf{G}$, where

$$\mathbf{U}(\lambda,\tau) = (\mathbf{S},\mathbf{R},\mathbf{M}) \begin{pmatrix} \mathbf{S}^t\mathbf{S} & \mathbf{S}^t\mathbf{R} & \mathbf{S}^t\mathbf{M} \\ \mathbf{R}^t\mathbf{R} & \mathbf{R}^t\mathbf{R}+(n\lambda)\mathbf{Q} & \mathbf{R}^t\mathbf{M} \\ \mathbf{M}^t\mathbf{S} & \mathbf{M}^t\mathbf{R} & \mathbf{M}^t\mathbf{M}+\tau\mathbf{I} \end{pmatrix}^{+} \begin{pmatrix} \mathbf{S}^t \\ \mathbf{R}^t \\ \mathbf{M}^t \end{pmatrix},$$

and $\mathbf{C}^+$ denotes a generalized inverse of $\mathbf{C}$ satisfying $\mathbf{C}\mathbf{C}^+\mathbf{C} = \mathbf{C}$, $\mathbf{C}^+\mathbf{C}\mathbf{C}^+ = \mathbf{C}^+$, $(\mathbf{C}\mathbf{C}^+)^t = \mathbf{C}\mathbf{C}^+$, and $(\mathbf{C}^+\mathbf{C})^t = \mathbf{C}^+\mathbf{C}$. This inverse is also known as Moore-Penrose inverse.

For different values of λ and τ, (3.14) defines a family of possible solutions. Optimal values of these parameters are obtained by minimizing the generalized cross-validation score

$$V(\lambda,\tau) = \frac{n^{-1}\,\mathbf{G}^t(\mathbf{I}-\mathbf{U}(\lambda,\tau))^2\mathbf{G}}{[n^{-1}\,\mathrm{tr}(\mathbf{I}-\mathbf{U}(\lambda,\tau))]^2}. \tag{3.15}$$

Gu and Ma [2005] showed that under very general conditions the minimizers of $V(\lambda,\tau)$ yield an optimal smoothing asymptotically.

(ii) Depth-Harmonic Model

To correct for harmonic components of the noise we proceed as follows. For a fixed angle, any profile $g(\theta,y)$ is a function of depth y that we will just denote by $g(y)$. These functions may be contaminated by coherent noise caused my unmodeled oscillations in the subsurface. One possible model for this oscillations is a harmonic process $\sum_i a_i cos(\omega_i y + \phi)$, where a_i and ϕ_i are, respectively, random amplitudes and phases, and the frequencies $omega_i$ are fixed but unknown. Each realization of the process is just a sum of sinusoids and each profile is contaminated by a different harmonic process. Hence, removal of a harmonic process from a profile is equivalent to correcting for sinusoid signals.

To find sinusoids hidden in the data we use a method developed by *Quinn and Fernandes* [1991]. To explain the main idea we assume a single sinusoid. A profile is modeled as $\tilde{g}(y) = a\cos(\omega y + \phi) + g(y)$, where $g(y)$ is the clean signal we want to recover. A sinusoid can be annihilated with a second order filter. Indeed, we note that $\tilde{g}(y)$ satisfies the difference equation

$$\tilde{g}(y_j) - 2\cos(\omega)\,\tilde{g}(y_{j-1}) + \tilde{g}(y_{j-2}) = g(y_j) - 2\cos(\omega)\,g(y_{j-1}) + g(y_{j-2}).$$

ARMA fitting techniques can then be used to estimate the frequency ω and then obtain amplitude and phase through least squares. The frequency estimate can be interpreted as a local maximizer of a smoothed periodogram. The case of more than one frequency is treated in a similar way; there are difference operators that annihilate all the sinusoids. For more details on this methodology see *Quinn and Fernandes* [1991].

Once the harmonic components are estimated from data, we can obtain the harmonic-component extracted profile: $\tilde{G}_{ij} = G_{ij} - \sum_k \widehat{a}_{ik}\cos(\widehat{\omega}_{ik} y_j + \widehat{\phi}_{ik})$. Then the model can be fitted as before, that is, one minimizes

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\tilde{G}_{ij} - \phi(y_j)^t\beta - \varphi(z_j)^t\mathbf{c} - b_i)^2 + \frac{\sigma^2}{\sigma_s^2}\sum_{i=1}^{a} b_i^2 + n\lambda\,\mathbf{c}^t\mathbf{Q}\mathbf{c}. \tag{3.16}$$

(iii) Angle-Dilation Depth-Harmonic Model

As in the Depth-harmonic model, the harmonic components are estimated from the data and then subtracted from to obtain the corrected profile $\tilde{G}_{ij}$. We estimate (3.10) using

penalized least squares

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\tilde{G}_{ij}-g(\alpha_i y_j)-b_i)^2+\frac{\sigma^2}{\sigma_s^2}\sum_{i=1}^{a}b_i^2+n\lambda J(g). \tag{3.17}$$

The minimization is carried out iteratively as follows:

(0) As the initial estimate we use he maximum likelihood estimates of the linear mixed-effects model $\tilde{G}_{ij}=\alpha_i y_i+b_i+\varepsilon_{ij}$.

(1) For estimated $\hat{\alpha}_i$, we minimize the following functional to obtain $\hat{\beta}$ and $\hat{\mathbf{c}}$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\tilde{G}_{ij}-\phi(\hat{\alpha}_i y_j)^t\beta\ -\varphi(\hat{\alpha}_i z_j)^t\mathbf{c})^2+\frac{\sigma^2}{\sigma_s^2}\sum_{i=1}^{a}b_i^2+n\lambda\,\mathbf{c}^t\mathbf{Q}\mathbf{c} \tag{3.18}$$

(2) For estimated $\hat{\beta}$ and $\hat{\mathbf{c}}$, we estimate α_i by minimizing

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\tilde{G}_{ij}-\phi(\alpha_i y_j)^t\hat{\beta}\ -\varphi(\alpha_i z_j)^t\hat{\mathbf{c}})^2. \tag{3.19}$$

Steps (1) and (2) are iterated until convergence.

Chapter 4

Inverse scattering with *SKKS* coda waves: imaging the core side of the CMB[†]

Abstract

In our previous studies we developed a method for imaging the heterogeneity at and near the core mantle boundary with broadband *ScS* transverse component data. Moreover we developed a statistical model to produce the image of the D'' discontinuity with variable confidence levels. In this paper, we extend our previous development in as much as that we allow for (known) discontinuities in the background model; we can then incorporate the outer core in the background and use the *SKKS* phase (radial component) and its coda to scan the D'' discontinuity from the underside. We furthermore demonstrate that the *SKKS* phase is a phase that is supplementary to *ScS* phase and that is of importance for the imaging of the D'' discontinuity. Synthetic seismograms calculated with the *WKBJ* method are used to test the performance of our method. As a proof of concept, we transform ~18,000 radial-component *SKKS* waveforms into image gathers of a CMB patch beneath Central America.

[†]Inverse scattering with *SKKS* coda waves: imaging the core side of the CMB, *Geophys. J. Int., mauscript in preparation.*

The structure of the SKKS image gathers is consistent with ScS image gathers.

4.1 Introduction

To detect heterogeneities in Earth's interior that are singular in nature, the use of scattered body wave phases is pertinent. The singular parts of these waves contain information about non-smooth variations in material and physical properties of the Earth.

Most work to date inferring the structure of the CMB region (for example, *Garnero* [2000]) has employed forward modeling to fit the waveforms of phases like *S-ScS* and *SKS-SPdKS-SKKS* on selected high-quality seismograms. For early results on modeling core phases, we refer to *Choy* [1977]. *Wang et al.* [2006] developed an inverse scattering approach based upon the generalized Radon transform (e.g., *Beylkin* [1984]; *De Hoop et al.* [1994]; *De Hoop and Bleistein* [1997]; *De Hoop et al.* [1999]; *De Hoop and Brandsberg-Dahl* [2000]; *Stolk and de Hoop* [2002]; *Brandsberg-Dahl et al.* [2003]), to image selected neighborhoods of the CMB using *ScS* data. In addition, statistical methods were applied to produce images of D'' discontinuities and estimate their uncertainty. However, using only the *ScS* data to scan the CMB area from one side (topside) and the statistical method to validate our results has its drawbacks: 1) the data coverage is insufficient for some regions, specially for regions below the center of oceans. 2) interfaces with a wavespeed drop are harder to detect than a wavespeed increase (*Flores and Lay* [2005]).

The approach developed and new data set used in this paper can overcome these drawbacks. With the extension presented here, we can use *SKKS* wavefield to scan the D'' from below. We can thus produce images with SKKS phases and with ScS phases of the same region; given that the relevant data are so different, a consistent result would serve as a

validation. The use of $SKKS$ has several advantages. First, unlike the S_dS phase, the amplitudes of phase SKS^dSKS are almost the same for the same amount of velocity increase or decrease (see Figure 4-5). Second, $SKKS$ provides excellent data coverage. Since the maximum epicentral distance for ScS data is about 80 degree, and because there are almost no receivers and events in large intraplate regions, such as oceans, the ScS middle point coverage is very sparse there. Indeed, Central America and eastern Eurasia are among the few regions where ScS data coverage is likely to be sufficient for successful application of the GRT with ScS data. On the other hand, the $SKKS$ middle point coverage is very good in most of the regions (see Figure 4-2). The main reason is that the epicentral distances used for $SKKS$ data are from 100-180 degrees. All these features make the $SKKS$ phase complementary phase of ScS. Of course, to use $SKKS$ data to image the CMB area from the underside has its challenges. i) the lower limit of the earthquake magnitude which can be used to image with $SKKS$ data has to be higher because the $SKKS$ rays propagate along longer raypaths, as compared with ScS rays; ii) since the $SKKS$ propagates not only in the solid mantle but the liquid outer core, one has to deal with both solid-solid and solid-liquid interfaces; iii) for ScS imaging, we only use the transverse (SH) component and we only need to deal with a relative simple system, whereas for $SKKS$ imaging, one has to deal with the coupled P-SV system. However, the potential mixture of SKS^dSKS and SKP^dPKS seems to make this approach almost infeasible. Fortunately, one can select an epicentral distance range (100-180 degree) in which there is no SKP^dPKS energy because the incidence angle of K at the CMB is beyond the post-critical angle for the mantle P-wave (see Figure 4-1).

Rost and Revenaugh [2004] found a strong arrival in the early coda of major-arc $PKKPab$ and interpreted it as an underside reflection from D''. With this arrival, they found a D'' at

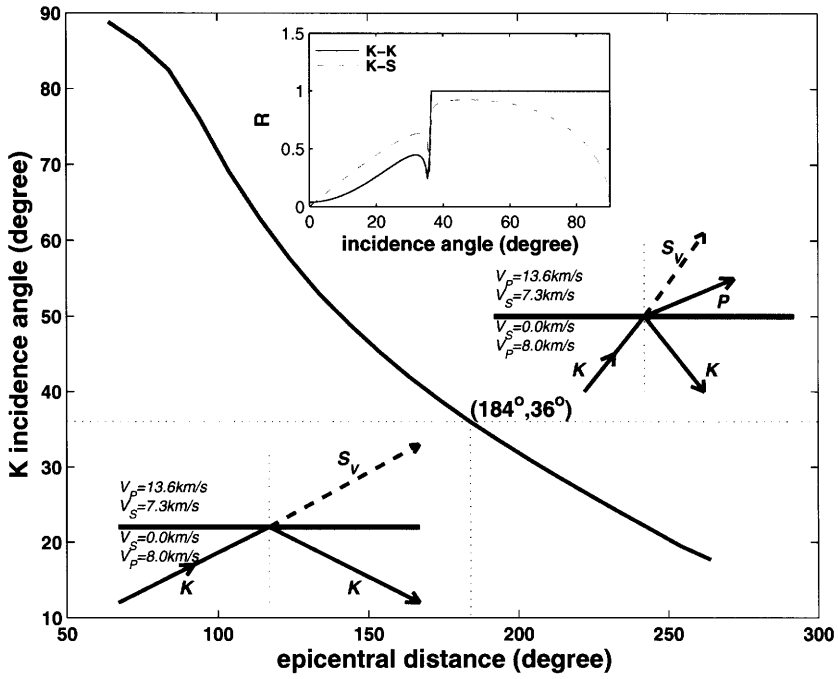

Figure 4-1: The K-wave incidence angle at the CMB vs the epicentral distance for $SKKS$ phase. The inset is the $K-K$ reflection and $K-S$ conversion coefficients vs the incidence angle at the CMB. For an epicentral distance $\Delta < 184°$, the K-wave incidence angle $\alpha > 36°$. $\alpha = 36°$ is the critical angle where the K-wave completely reflects, thus no energy of P-wave in the mantle for an epicentral distance smaller than 184°.

280 km above the CMB by converting the traveltime to depth. However, trying to use PKP^dPKP phase to find the D'' discontinuity is not easy. First as mentioned above, due to the dramatical P wavespeed change from the outer core to the mantle, it readily reach the critical angle for the incidence of K and the mantle P-wave. Therefore, the epicentral distance range where this phase shows up is highly limited. The triplication of $PKKP$ around this distance range further complicates the situation. Second, whenever there exists a PKP^dPKP phase, the PKS^dSKP phase must also appear. This may cause misinterpretation of the depth of interfaces. While the usable distance range for PKP^dPKP is limited and complicated by the triplications, the distance free of $SKKS$ triplication is large (100°-180°) and the SKS^dSKS phase is not contaminated by the SKP^dPKS phase.

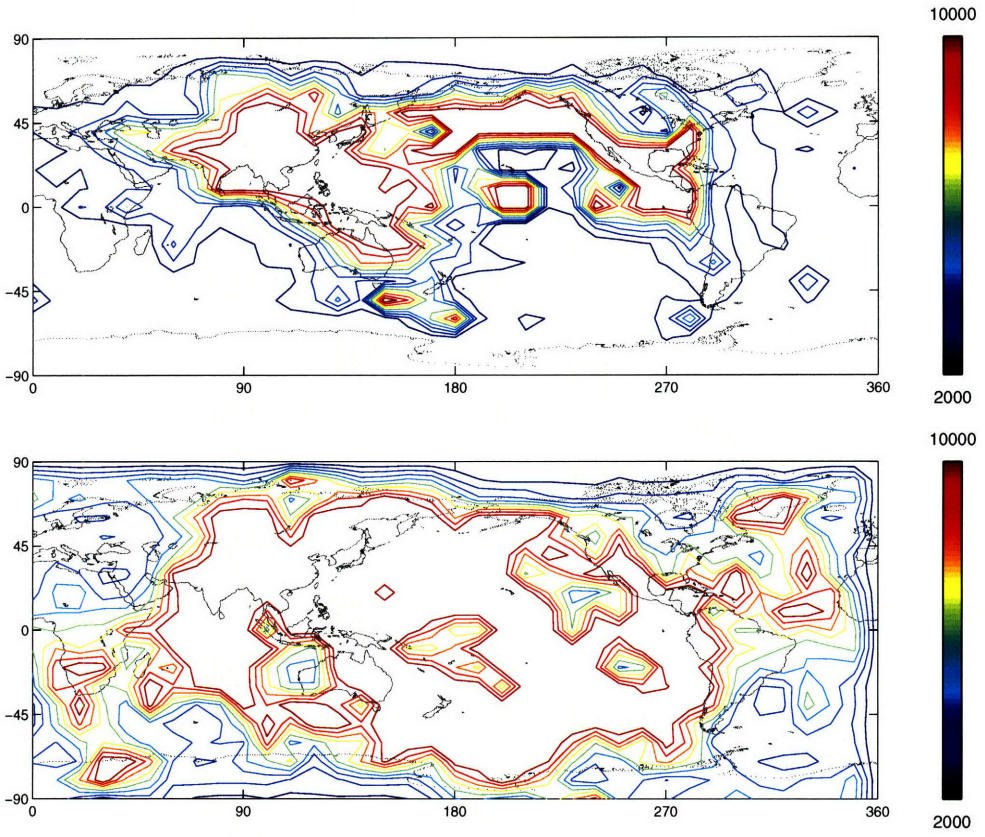

Figure 4-2: Global distribution of number of *ScS* (top) and *SKKS* (bottom) middle points in a $10^\circ \times 10^\circ$ bin. Data source is IRIS-DMC; earthquakes have $m_b > 5.2$, with origin time between 1988 and 2002.

A global energy stack of $SKKS(SV)$ shows a later, previously unidentified arrival about 1 minute after the main *SKKS* phase (see the black box in Figure 4-3). It's tempting to interpret it as the under side reflection from D'' given the fact that this arrival still shows up when using deep events only.

The main goal of this paper is to develop a mixed fluid-solid generalized Radon transform approach to inverse scattering adapted for coda waves in global seismology, and to apply this approach to a range beneath Central America. In Section 4.2 we develop the theory underlying the GRT and describe how three- component broadband data can be transform into image gathers. In Section 4.3 we use synthetic seismograms (calculated with WKBJ

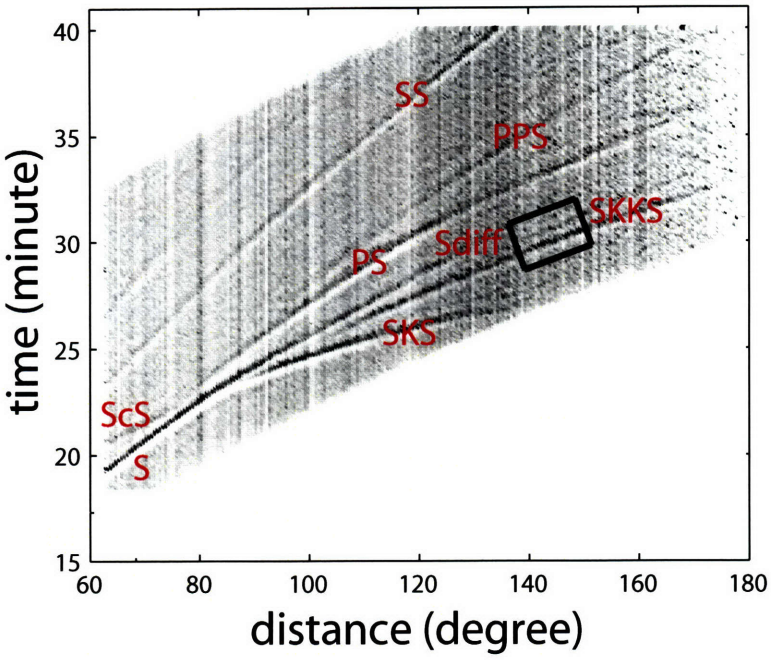

Figure 4-3: Stack of the 100,000 global $SKKS(SV)$ data. Processing details data source is IRIS-DMC; band-pass filter is 10-50 s; earthquakes have $m_b > 5.2$, with origin time between 1990-2002.

Chapman [1978]) to test the performance of our methodology. We show that in principle, the GRT with $SKKS$ phase can be used to resolve interfaces with either a wavespeed increase or wavespeed decrease, which are different to see with ScS data (*Flores and Lay* [2005]). This feature makes GRT with $SKKS$ very suitable for resolving low wavespeed layers under high wavespeed layer. We also demonstrate that the reflection coefficients, relative to $SKKS$, can be estimated. Finally, in Section 4.4 we present preliminary results of three imaging points, compare them with the ScS images, and then we show the results of joint inversion of ScS and $SKKS$.

4.2 Inverse scattering

The inverse scattering transform we present here makes essentially use of multiple scattered waves, by incorporating discontinuities in the background. The transform, nonetheless, is very closely related to the GRT and falls in the category of liberalized LS inversion for singularities.

Although the mathematics needed to develop the GRT method are quite complex, the principal idea of this method is relatively simple. The GRT method projects the scattering potential into the data as integrals over isochrone surfaces; in turn, integrating the data over isochrone surfaces recovers an image of the scattering potential.

In the sections that follow, we develop the generalized Radon transform (GRT) method by first obtaining the Green's function and Green's tensor from the governing equations in both solid and fluid media. Then we assume that the wave that propagates in a perturbed medium has the same raypath as that in a smooth background medium and the wavefield perturbation is linearly dependent on the medium perturbation. The above assumptions are known as the Born approximation (*Aki and Richards* [1980]). The general form of the problem is

$$u = F\delta c\,, \tag{4.1}$$

where u is the scattered wavefield, δc is the medium perturbation, and F is an integral operator.

Applying the adjoint operator F^* to both sides of equation 4.1 yields

$$F^*u = F^*F\delta c\,, \tag{4.2}$$

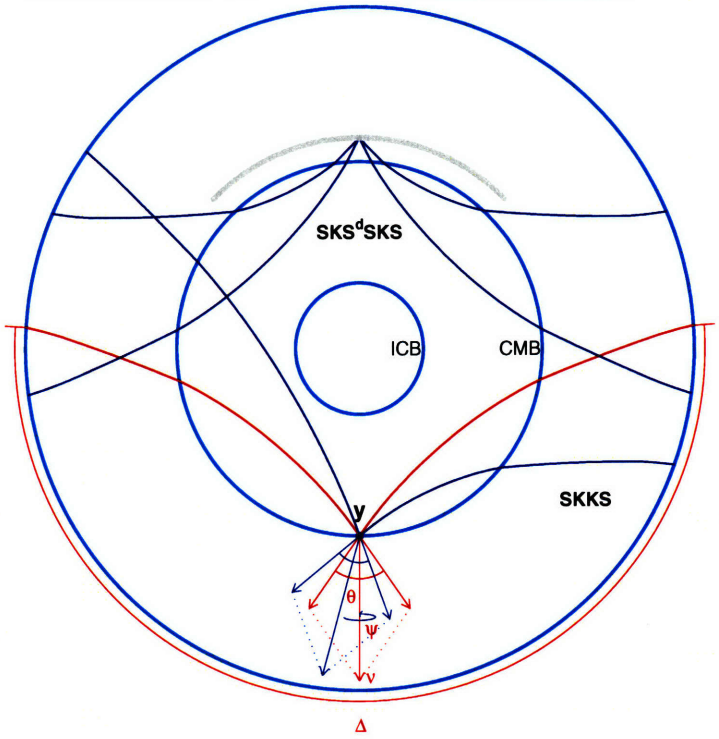

Figure 4-4: Schematic illustration of the path geometry of $SKKS$ (bottom) and SKS^dSKS (top) considered in the generalized Radon transform (GRT) of $SKKS$ data. Bottom: The source and receiver are separated by epicentral distance Δ. The image point at the CMB is denoted y. The summation of the slowness vectors of the two legs of $SKKS$ are given by ν. The scattering angle is θ and scattering azimuth is ψ. The image is, essentially, created by integration over ν.

where F^*F is the normal operator. Taking the generalized inverse of equation 4.2 produces images of the medium perturbation:

$$\delta c = (F^*F)^{-1}F^*u\,. \tag{4.3}$$

The GRT inversion defined by equation 4.3 can, under certain conditions, be subjected to restricting the inversion operator to a prescribed scattering angle and azimuth (θ, ψ). This inversion produces common image point gathers which represent an image of the Earth's interior at common locations.

4.2.1 The governing wave equations

We first look at the propagation of seismic waves in both solid and fluid media. In the solid regions of the Earth, typically shells such as the mantle, the particle displacement $u_i = u_i(x,t)$ satisfies the elastic wave equation

$$\rho\, \partial_t^2 u_i - \partial_j(c_{ijk\ell}\partial_\ell u_k) = f_i\,, \tag{4.4}$$

where $\rho = \rho(x)$ is the scalar density of mass, $c_{ijk\ell} = c_{ijk\ell}(x)$ is the elastic stiffness tensor, and $f_i = f_i(x,t)$ is the body-force source density. The stress τ_{ij} in the solid is related to the displacement as $\tau_{ij} = c_{ijk\ell}\partial_\ell u_k$ in accordance with Hooke's law. The elastodynamic Green's tensor, $G_{ip}(x,x',t)$, which is the solution of the wave equation for a point source at (x',t), satisfies the equation

$$\rho\, \partial_t^2 G_{ip} - \partial_j(c_{ijk\ell}\partial_\ell G_{kp}) = \delta_{ip}\delta(x-x')\delta(t)\,, \tag{4.5}$$

subject to the condition of causality, $G_{ip}(x,x',t) = 0$ for $t < 0$. The modes of seismic wave propagation (in the mantle) are P, SV and SH in the isotropic case, and qP, qS1 and qS2 in the anisotropic case.

In the fluid regions of the Earth, typically shells such as the outer core, the acoustic pressure $p = p(x,t)$ satisfies the acoustic wave equation

$$\kappa\, \partial_t^2 p - \partial_j(\sigma\partial_j p) = q\,, \tag{4.6}$$

where $\sigma = \sigma(x)$ is the scalar reciprocal density of mass, $\kappa = \kappa(x)$ is the compressibility or reciprocal of bulk modulus, and q is the time derivative of the volume source density of

injection rate. The scalar traction in the fluid is the opposite of the pressure; the particle velocity v_i in the fluid is related to the pressure as $v_i = -\rho^{-1}\partial_t^{-1}\partial_i p$. The acoustic Green's function, $G(x,x',t)$, satisfies the equation

$$\kappa\partial_t^2 G - \partial_j(\sigma\partial_j G) = \delta(x-x')\delta(t)\,, \tag{4.7}$$

subject to the condition of causality, $G(x,x',t) = 0$ for $t < 0$. The mode of seismic wave propagation in the outer core is denoted by K.

The frequency domain equations both for the solid and the fluid regions are obtained by replacing $-\mathrm{i}\partial_t$ by ω.

4.2.2 The source representation

Next, we need a term that describes the excitation (normally a body force) of the wavefield. We write the equivalent body force in terms of the so-called symmetric stress glut tensor S,

$$f_j = -\partial_i S_{ij}\,; \tag{4.8}$$

the stress glut tensor is related the source moment tensor as

$$(\partial_{t_0} S_{ij})(x_0,t_0) = M_{ij}\delta(x_0-s)\delta(t_0-t_s)\,. \tag{4.9}$$

Substituting equation 4.8 into 4.9 we obtain

$$f_j(x_0,t_0) = -M_{ij}\partial_i\delta(x_0-s)\,H(t_0-t_s)\,, \tag{4.10}$$

where $H(t_0 - t_s)$ is the step function. For convenience, and without loss of generality, we shift t_s to 0 for each event.

4.2.3 Geometrical ray Green's tensor

A standard method for solving a wavefield excited by a volume source is to obtain the solution of the governing equations assuming a point source, then obtain the final solution using the superposition principle.

First, we discuss the high-frequency Green's tensor restricting the analysis to the solid in the absence of a fluid region. Away from caustics at the receiver at x and at the source at x', the Green's tensor admits the oscillatory integral representation

$$G_{ip}(x,x',t) = \frac{1}{\pi} \sum_m \mathrm{Re} \int_0^\infty \xi_i(x)\xi_p(x')A^{(m)}(x,x') \exp[\mathrm{i}\omega\,(t - T^{(m)}(x,x'))]\,\mathrm{d}\omega\,, \qquad (4.11)$$

in which

$$A(x,x') = \frac{1}{4\pi[\rho(x)\rho(x')\mathscr{M}(x,x')]^{1/2}}\,, \qquad (4.12)$$

where

$$\mathscr{M}(x,x') = \|v(x')\|V(x)\frac{\partial(x^{\mathrm{W}})}{\partial(\gamma^{\mathrm{O}})} \qquad (4.13)$$

and ξ denotes the normalized polarization vector of the wave-type under consideration. The amplitude $A^{(m)}$ is obtained from A upon multiplication by the phase factor, $\exp[-\mathrm{i}(\pi/2)\iota(x,x')]$, accounting for the KMAH index ι. The index m labels the branches of the travel time function. We suppress the index m in our notation. In (4.13) v denotes the group velocity and V denotes the phase velocity; x^{W} denotes the coordinates in the wavefront at x while γ^{O}

denotes the coordinates on the slowness surface at x'. Also,

$$V(x')\mathscr{R}(x,x') := [\mathscr{M}(x,x')/(V(x)V(x'))]^{1/2} = |\det \mathbf{Q}_2(x,x')|^{1/2} = \mathscr{L}(x,x')$$

can be identified as the reciprocal of relative geometrical spreading (see, for example, [*Aki and Richards*, 1980, (9.46)] in the case of P waves); the matrix $\mathbf{Q}_2$ is a quantity, defined in [*Červený*, 2001, (4.3.5)], that is amenable to numerical computations. The travel time along the ray or path connecting x with x' is denoted by $T(x,x')$; the index m keeps track of multipathing.

Secondly, we discuss the high-frequency Green's function for the case of a analysis to the fluid in the absence of a solid region. Away from caustics at the receiver at x and the source at x', the Green's function admits the oscillatory integral representation

$$G(x,x',t) = \frac{1}{\pi} \sum_m \mathrm{Re} \int_0^\infty B^{(m)}(x,x') \exp[\mathrm{i}\omega (t - T^{(m)}(x,x'))] \, \mathrm{d}\omega \,, \tag{4.14}$$

in which

$$B(x,x') = \frac{[\rho(x)\rho(x')c_f(x)c_f(x')]^{1/2}}{4\pi \mathscr{L}(x,x')} \,, \tag{4.15}$$

with $c_f = (\sigma^{-1}\kappa)^{1/2}$ denoting the wavespeed. The amplitude $B^{(m)}$ is obtained from B upon multiplying by the phase factor accounting for the KMAH index. Thus the index m labels the branches of the travel time function. We suppress the index m in our notation.

4.2.4 Incident field

Substituting (4.10) for the inhomogeneous right-hand side of the wave equation (4.4), using Duhamel's principle, we obtain for its solution

$$u_i(x,t) \simeq \frac{1}{\pi} \sum_m \mathrm{Re} \int_0^\infty \xi_i(x) \tfrac{1}{2}(\xi_p(s)\gamma_q(s) + \xi_q(s)\gamma_p(s)) M_{pq} \\ A^{(m)}(x,s) \exp[-\mathrm{i}\omega(t - T^{(m)}(x,s))] \, \mathrm{d}\omega \,, \tag{4.16}$$

in which

$$\gamma_p(s) = (\partial_{s,p} T)(x,s) \tag{4.17}$$

is the p-component of the slowness vector at s associated with the ray connecting s with x. This equation can be compared with [*Aki and Richards*, 1980, (4.88)-(4.90)] upon reducing the elastic stiffness to the isotropic elastic case with Lamé parameters λ and μ. In these equations, $M_{pq} = \mu(\bar{u}_p \nu_q + \bar{u}_q \nu_p) A$ in which ν is the normal to the fault surface, $\bar{u}_p$ is the average displacement discontinuity, and A is the fault area. Then the source radiation pattern can be written in the form

$$F(s) = \mathsf{M}^{-1} M_{pq} \xi_p(s) V(s) \gamma_q(s) \,, \quad \mathsf{M} = \mu A \|\bar{u}\| \,.$$

4.2.5 Background fluid-solid interface

Boundary conditions

The coefficients in the above equations are associated with a background model. In such a model, the coefficients are assumed to be smooth except at a fluid-solid interface, Σ say, such as the globally estimated core-mantle boundary (CMB). At a solid-fluid interface, the

following boundary conditions apply: (i) the normal component of the particle displacement is continuous, (ii) the normal component of the surface traction in the solid and the scalar traction in the fluid are equal, (iii) the tangential components of the surface traction in the solid vanish.

Solving the boundary value problem

We couple the fluid and solid expressions developed above by solving the boundary value problem. We assume that the source is contained in the solid region while the receiver is contained in the fluid region. Asymptotically, the solution is adapted to include a transmission coefficient. This implies that the amplitude in the oscillatory integral representation becomes $\xi_p(x')C(x,x')$,

$$G_p(x,x',t) = \frac{1}{\pi}\sum_m \mathrm{Re}\int_0^\infty \xi_p(x')C^{(m)}(x,x')\exp[\mathrm{i}\omega(t-T_\Sigma^{(m)}(x,x'))]\,\mathrm{d}\omega\,. \tag{4.18}$$

With x^Σ the point of refraction on the fluid-solid interface Σ, the travel time in the phase function becomes

$$T_\Sigma(x,x') = T(x,x^\Sigma) + T(x^\Sigma,x')\,.$$

The amplitude, $C(x,x')$, is derived as follows. The incident displacement amplitude (the $-$ refers to the solid (lower mantle) side of Σ) follows to be

$$U(x^\Sigma)_- = \left[\frac{\rho(x')\|v(x')\|}{(\rho(x^\Sigma)\|v(x^\Sigma)\|)_-}\right]^{1/2}\frac{\mathscr{L}(x')}{\mathscr{L}(x^\Sigma)_-}U(x')$$

$$= \left[\frac{\rho(x')V(x')}{(\rho(x^\Sigma)V(x^\Sigma))_-}\right]^{1/2}\frac{1}{\mathscr{L}((x^\Sigma)_-,x')}\frac{1}{4\pi\rho(x')V(x')}$$

using that

$$\lim_{x''\to x'} \mathscr{L}(x'',x')U(x'') = \frac{1}{4\pi\rho(x')V(x')} ,$$

while the pressure amplitude at the receiver is related to the pressure amplitude at the fluid-solid interface (the subscript $'+'$ refers to the fluid (core) side of Σ) according to

$$P(x) = \left[\frac{\rho_f(x)c_f(x)}{(\rho_f(x^\Sigma)c_f(x^\Sigma))_+}\right]^{1/2} \frac{\mathscr{L}(x^\Sigma)_+}{\mathscr{L}(x)} P(x^\Sigma)_+ .$$

In the solid (mantle) the geometrical spreading is related to the relative geometrical spreading as [*Červený*, 2001, (4.14.44)]

$$\mathscr{L}(x^\Sigma) = |V(x^\Sigma)\|v(x^\Sigma)\|^{-1}\det(\mathbf{Q}_2(x^\Sigma,x'))\det(\mathbf{P}(x'))|^{1/2} .$$

Invoking the boundary conditions at the fluid-solid interface, accounting for its curvature, implies that

$$C(x,x') = (\rho_f(x)c_f(x))^{1/2} \frac{\mathscr{T}(x^\Sigma)}{\mathscr{L}(x,x')} \frac{1}{4\pi(\rho(x')V(x'))^{1/2}} , \tag{4.19}$$

where we have the factorization

$$\mathscr{L}(x,x') = |\det(\,\mathbf{Q}_2(x,x^\Sigma)\Pi\mathbf{Q}_2(x^\Sigma,x')\,)|^{1/2} . \tag{4.20}$$

Here (cf. [*Červený*, 2001, (4.14.71)])

$$\Pi = (\mathbf{G}_+ - \mathbf{A}_+^{\text{an}})^{-1}\mathbf{M}^{\text{F}}(x,x^\Sigma,x')[(\mathbf{G}_- - \mathbf{A}_-^{\text{an}})^{-1}]^T , \tag{4.21}$$

in which $\mathbf{M}^{\text{F}}$ is the Fresnel zone matrix at x^Σ, and $\mathbf{A}_+^{\text{an}} = \mathbf{0}$.

In general, the fluid-solid boundary conditions have to be solved numerically. This is accomplished by expressing the fluid quantities in pressure and the solid quantities in displacement and then solving the remaining system of algebraic equations. In (4.19) a transformation to a flux-normalized transmission coefficient is applied, which we write as $\mathscr{T}(x^\Sigma) = \hat{T}_\mathrm{K} \,.\, \mathscr{E}$. Here, the flux normalizing factor is given by

$$\mathscr{E} = \frac{(\rho(x^\Sigma)\|v(x^\Sigma)\|)_-^{-1/2}}{(\rho_f(x^\Sigma)c_f(x^\Sigma))_+^{1/2}} \frac{\mathscr{L}(x^\Sigma)_+}{\mathscr{L}(x^\Sigma)_-} = \frac{(\rho(x^\Sigma)\|v(x^\Sigma)\|)_-^{-1/2}}{(\rho_f(x^\Sigma)c_f(x^\Sigma))_+^{1/2}} \frac{1}{(\cos\theta_-^\Sigma)^{1/2}} \,, \tag{4.22}$$

where θ_-^Σ is the angle of incidence.

For an isotropic elastic solid-fluid interface, the following expressions can be obtained for the transmission coefficient (4.19). We write $\hat{T}_\mathrm{KS} = (-\mathrm{i}\omega)(\rho_f c_f)_+ [c_f^{-1} T_\mathrm{KS}(pc_S^{-1})^{-1}]$, in which

$$T_\mathrm{KS}(p_1,p_2) = -\frac{c_S^{-2}p^2\gamma_P\gamma_S}{\Delta_\mathrm{SCH}} \,, \quad p = (p_1^2+p_2^2)^{1/2} \,, \tag{4.23}$$

in which the slowness vector is written in interface normal components, $\gamma = (p_1, p_2, \gamma_{P,S})$ with (p_1, p_2) tangential to the interface; Δ_SCH is the Scholte denominator associated with surface waves propagating along a solid-fluid interface [1],

$$\Delta_\mathrm{SCH} = \frac{\rho_f}{4\rho_s}\gamma_P c_S^{-4} + \gamma_f[(\tfrac{1}{2}c_S^{-2} - p^2)^2 + p^2\gamma_P\gamma_S] \,.$$

Equation (4.23) appears in the generalized ray analysis of this transmission problem in a configuration of a fluid and solid half space. Analogue expressions are obtained for T_KP.

[1] We identify $\mathrm{D}/(\gamma_S)_\mathrm{core} = (2\rho_s c_S^2)^2\Delta_\mathrm{SCH}$, cf. [*Aki and Richards*, 1980, pp.436,451] upon identifying $\frac{p^2}{r_\mathrm{CMB}^2}$ with $p_1^2 + p_2^2$. The transmission coefficient for incident SV, transmitted K displacements is $\grave{S}\grave{P}$ [*Aki and Richards*, 1980, p.150] which maps to $c_f^{-1}T_\mathrm{KS}(pc_S^{-1})^{-1}$. The coefficient $\grave{S}\grave{P}$ is easily adapted to WKB calculations in a layered SNREI model [*Aki and Richards*, 1980, p.437].

4.2.6 Modeling: The short-period Born approximation

We have obtained the source and the Green's tensors in Section 4.2.2 and 4.2.3 respectively. In this section we will describe the scattered wavefield in terms of the medium contrasts.

Let u denote the scattered displacement field. Substituting the high-frequency Green's tensors evaluated in the background into the Born approximation for the scattered displacement yields

$$u_p(r,t,s) = (F\delta c)_p = -M_{qr}(s)\frac{1}{\pi}\sum_m \mathrm{Re}\int_0^\infty \int_X (-\omega^2)\tilde{\xi}_q(s)\hat{\xi}_p(r)\tilde{\gamma}_r(s)\mathscr{I}(x,s,r)$$
$$w^T(x,s,r)\delta c(x)\exp[\mathrm{i}\omega(t-T^{(m)}(x,s,r))]\mathrm{d}V(x)\,\mathrm{d}\omega\,, \qquad (4.24)$$

where we denoted the quantities associated with a ray connecting the scattering point x with an earthquake location s by $\tilde{}$ and the quantities associated with a ray connecting the scattering point x with a station at r by $\hat{}$. The quantity $\mathscr{I}(x,s,r)$ is essentially the product of amplitudes that are possibly complex through the appearance of the exponential containing the KMAH index. The travel time is the sum

$$T(x,s,r) = T(x,s) + T(r,x)\,. \qquad (4.25)$$

We employ the shorthand notation

$$W(x,s,r) = M_{qr}(s)\,\tfrac{1}{2}(\tilde{\xi}_q(s)\tilde{\gamma}_r(s) + \tilde{\xi}_r(s)\tilde{\gamma}_q(s)) = \mathsf{M}(s)V(s)^{-1}F(s)\,. \qquad (4.26)$$

We distinguish the cases where the scattering point x is contained in the solid ($x \in X_s$) and

where the scattering point x is contained in the fluid ($x \in X_f$); we have $X = X_s \cup X_f$. Thus

$$u_p(r,t,s) = u_p^s(r,t,s) + u_p^f(r,t,s)$$

where $u_p^s(r,t,s)$ – representative of topside reflection – is obtained by subsituting $X = X_s$,

$$\mathscr{I}(x,s,r) = \rho(x)\tilde{A}(x)\hat{A}(x)\,, \tag{4.27}$$

in which

$$\tilde{A}(x) = A(s,x)\,,\ \ \hat{A}(x) = A(r,x)\,, \tag{4.28}$$

and

$$w(x,s,r) = \left\{ \tilde{\xi}_i(x)\hat{\xi}_i(x), \tilde{V}_\circ(x)\hat{V}_\circ(x)\,\tilde{\xi}_i(x)\tilde{\gamma}_i(x)\hat{\xi}_k(x)\hat{\gamma}_\ell(x) \right\}\,. \tag{4.29}$$

In the isotropic case, this expression reduces to

$$w(x,s,r) = \left\{1, \cos(\theta^{\mathrm{SS}}(x,s,r))\right\} \tag{4.30}$$

for SH-to-SH scattering, and

$$w(x,s,r) = \left\{\cos(\theta^{\mathrm{SS}}(x,s,r)), \cos(2\theta^{\mathrm{SS}}(x,s,r))\right\} \tag{4.31}$$

for SV-to-SV scattering. Here

$$\cos(\theta^{\mathrm{SS}}(x,s,r)) = c_S(x)^2\tilde{\gamma}_i(x)\hat{\gamma}_i(x)\,.$$

For $u_p^f(r,t,s)$ – representative of underside reflection – we obtain the substitutions $X = X_f$,

$$\mathscr{I}(x,s,r) = \kappa(x)\tilde{C}(x)\hat{C}(x)\,, \tag{4.32}$$

in which

$$\tilde{C}(x) = C(s,x)\,,\ \ \hat{C}(x) = C(r,x)\,, \tag{4.33}$$

while

$$w(x,s,r) = \left\{1, c_f(x)^2\tilde{\gamma}_i(x)\hat{\gamma}_i(x)\right\}. \tag{4.34}$$

We write

$$\cos(\theta^{\mathrm{KK}}(x,s,r)) = c_f(x)^2\tilde{\gamma}_i(x)\hat{\gamma}_i(x)\,.$$

4.3 Resolution tests with synthetic data

We test the performance of the methodology developed above with synthetic broad-band seismograms of *SKKS*. We calculate time windows containing *SKKS* using the radially stratified wavespeed model *ak135* (*Kennett et al.* [1995]). For the calculation of the waveforms we superimpose jumps in elastic parameters at certain distances above the CMB. The seismograms in Figure 4-5A are calculated from a model with a 3% wavespeed increase at 250 km above the CMB and a 3% wavespeed decrease at 150 km above the CMB.

SKKS and SKS^dSKS are so-called min-max phases and their waveform is distorted by a $\pi/2$ phase shift (see Figure 4-5A). This can be modeled with a Hilbert transform, but the most straightforward way is to deconvolve the phases by themselves. To achieve this goal, we followed the procedures explained in **Chapter II**. That is, we first use multi-channel cross correlation (MCCC) to align the synthetic data, followed by a principal component

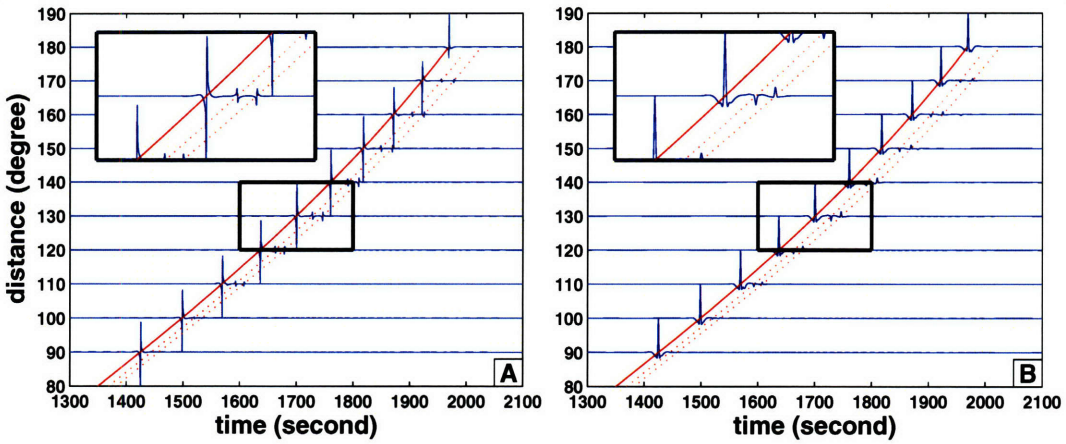

Figure 4-5: A. Record section of synthetic data for model with two contrasts (one increase and one decrease) above the CMB, calculated with WKBJ. The red solid lines are the travel time curves of $SKKS$ phase and the red dotted lines are the travel time curves of SKS^dSKS phases. B: Record section of synthetic data: after deconvolved by the PCA estimated $SKKS$ phase. The top black boxes in A and B is the blow-up of the bottom black boxes in A and B respectively.

analysis (PCA) to estimate the $SKKS$ phase, and then deconvolve the estimation from the full synthetic seismograms. The resulting seismograms (after deconvolution) are shown in Figure 4-5B.

In one series of tests we use an idealized (geographical) distribution of specular reflections (Figure 4-6A,B); in another we use the actual earthquake-station distribution (Figure 4-6C). We show results for (synthetic) data bandpass filtered between 10-50 s (Figure 4-6A,C).

The 28 traces in the left of Figure 4-6A,C are extracted trace by trace from synthetic data with different epicentral distances (scatter angles). Aliasing is visible in Figure 4-6C due to sparse and uneven sampling.

For the wavespeed models and associated ray geometries considered we can calculate the reflection coefficient R relative to $SKKS$ as a function of scatter angle. We then compare it with the relative reflection coefficient inferred from the GRT. These two coefficients generally agree well with each other (see Figure 4-6B).

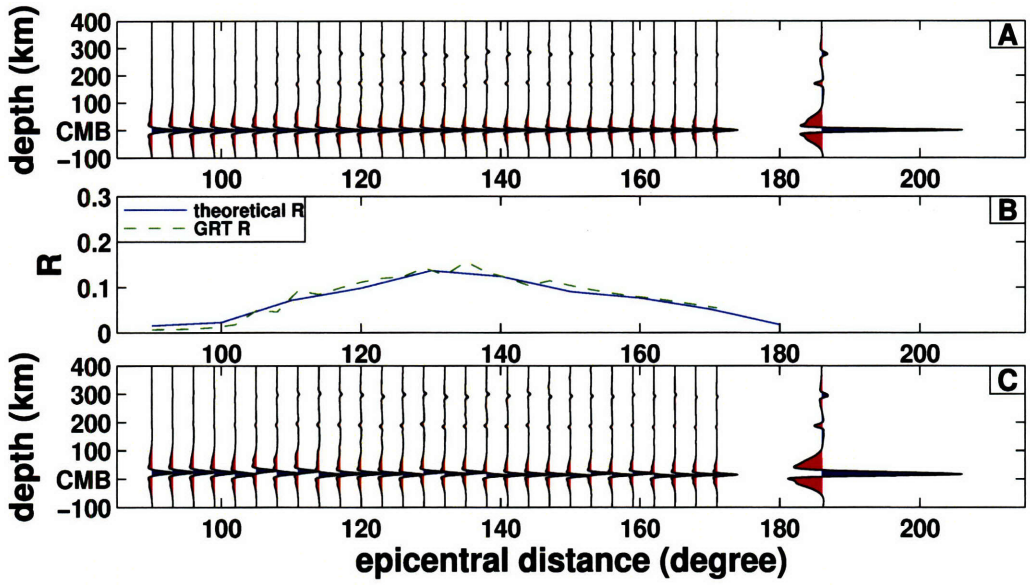

Figure 4-6: Illustration of the construction of GRT stacks (images) from image gathers at different scatter angles. The traces on the left of panels (A) and (C) are image gathers at scattering angles produced from the synthetic data as in Figure 4-5. The traces on the right are stacks over scattering angles. The gathers and stacks in (A) are produced from an artificial (regular) source-receiver distribution; the results in (C) were computed using the data coverage depicted in Figure 4-8.

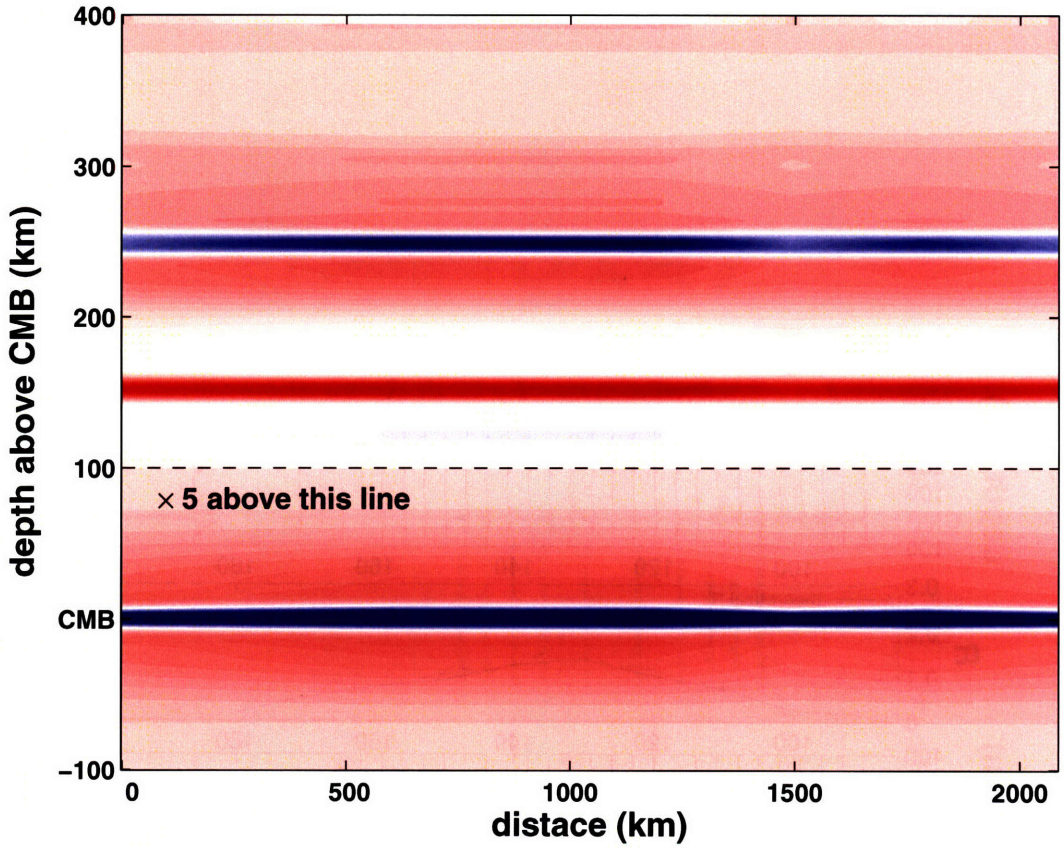

Figure 4-7: The amplitude above the dashed line are multiplied by a factor of 5 to make it comparable to that at the CMB. This figure shows the recovery of the input model with a wavespeed decrease at 150 km above the CMB and a wavespeed increase 250 km above the CMB.

We demonstrate that our method is able to detect multiple interfaces with opposite velocity changes. For this purpose, forty-one imaging points are chosen along the great circle transect from (-105W, 0) to (-75W, 30N). Figure 4-7 is generated by the lateral juxtaposition of IGs stacked over all angles.

4.4 Imaging the CMB beneath Central America

We apply the GRT method to a broadband wavefield formed by *SKKS* (and its precursors and coda) that sample a $50^\circ \times 50^\circ$ core mantle boundary (CMB) beneath Central America (Figure 4-8). This region has been studied intensively and several investigators have found evidence for structural complexity within D'' (e.g., *Garnero* [2000]; *Buffett et al.* [2000]; *Thomas et al.* [2004]). Here we present three image gathers of lowermost mantle structure.

4.4.1 Data preprocessing and analysis

The generalized Radon transform is applied to data from many events and seismic stations. One could view this transform as a focusing procedure using 'arrays' of sources and receivers, searching for singularities in the Earth's interior, here the lowermost mantle.

The data are collected from the following sources. Estimates of origin times and source locations (hypocenters) are obtained from the EHB (*Engdahl et al.* [1998]) data base for all events. Three-component broadband waveforms for all events in our data set are obtained from IRIS (Incorporated Research Institutions for Seismology). The minimum magnitude considered in this study is set at $m_b > 5.5$. The range of angular epicentral distances was chosen from $100 - 180^\circ$ to void the potential mixture of SKS^dSKS and SKP^dPKS. The preprocessing sequence is similar to the one we explained in **Chapter II** and the published paper (*Wang et al.* [2006]). We use *SKKS* phase as reference phase to apply the principal component analysis (PCA). The band-pass filter we used in this study is $10 - 50$ s.

The application of the generalized Radon transform requires binning in scattering angle θ. From a physical point of view, we should match the bin size with the reflected-wave Fresnel zone. This match will only be possible for selected neighborhoods of the CMB.

The Fresnel volume can be determined by ray tracing through the condition

$$\min\{T(y,s)+T(x^{\Sigma},y,r)-T(r,s), s \leftrightarrow r\} = \tfrac{1}{2} f_0^{-1}$$

where f_0 represents the peak frequency and y lies on the boundary of the Fresnel volume.

The binning in θ requires a traveltime correction: Data from intervals $\Delta_\varepsilon(\theta)$, for given ('common') ν and ψ, are stacked prior to which traveltime variations with $\Delta = \Delta(\nu, \theta, \psi)$ are corrected for using Grand's spherically symmetric model.

To form the final gathers, the outcome of the generalized Radon transform is stacked over azimuth ψ, with the appropriate weights. The transform is essentially an integration over dip directions (and implicitly over isochrones); the cone of associated dip directions has typically an opening of about 24 degrees. The detectable structural dip directions must be contained in this cone. The CMB topography is covered by this cone assuming it is mild (*Morelli and Dziewonski* [1987]; *Sze and van der Hilst* [2003]).

4.4.2 Preliminary results: three image gathers

In this section we demonstrate the performance of our method to infer interfaces in the vicinity of the CMB. Presumably, any (local) reflector should show up at least at (close to) the same radius for any processed angle. However, due to the difference in coverage, quality for different epicentral distances, this is not always the case. Thus, the statistical analysis developed in **Chapter III** is very much needed to extract the coherent signals.

To illustrate the concepts developed so far, we draw from ~18,000 *SKKS* displacement records in Figure 4-8 to construct GRT image gathers of three locations beneath Central America (Figure 4-8, blowup) [figure needs to be updated]. We estimate the image gather

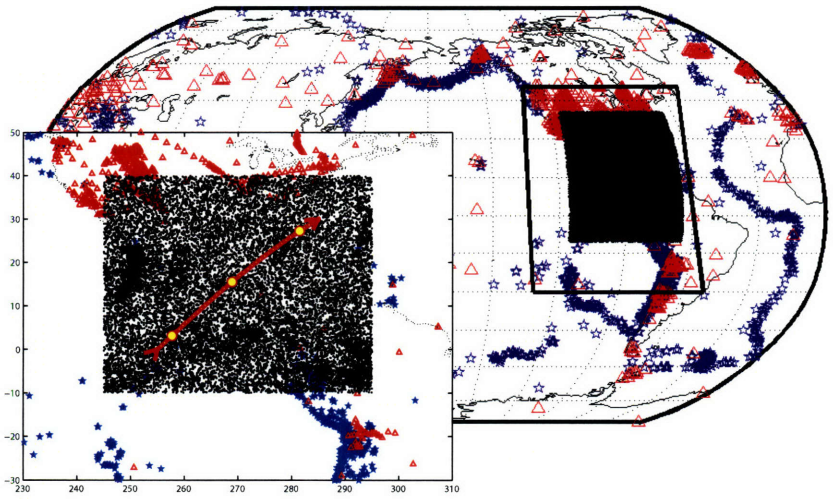

Figure 4-8: Geographic map of the region under study, depicting the epicenters of the ~2,200 earthquakes (blue stars) and the locations of the ~1,200 stations (inverted red triangles) that yielded the data used in the construction of the common image-point gathers. The $50° \times 50°$ CMB bin is indicated by the densely sampled rectangle: small black dots mark specular CMB reflection points of the ~18,000 *ScS* data. The inset in the left is the blowup of the black block in the right. The red arrow is profile B-B' in Figure 5-1. The image gathers of the three yellow dots from left to right are shown in Figure 4-9A, C, and B.

using all radial resolution bands corresponding to different scattering angles (epicentral distances). (Figure 4-9. Unlike imaging with *ScS* data, where S_dS changes its polarity upon crossing the intramission angle, S'^dS' keeps a fixed polarity for the epicentral distance range from 100-180°. Therefore, it's reasonable to make the estimate over all the angles to obtain the image gathers. The resulting three image gathers are shown in Figure 4-9A-C. The right two traces in Figure 4-9A-C are the image gathers for the same image points with *SKKS* (left) and *ScS* (right) data. The CMB is resolved using both *ScS* and *SKKS* data. Strong refectors are observed for all the three image points with *SKKS* study which is close to the L1 structure, the top of ppv lens, in Figure 5-4. We can also identify the L2 structure which presumably is the base of ppv lens (*Van der Hilst et al.* [2007]) in

Figure 5-4A. In Figure 5-4B, L2 structure is within the side lobe of CMB reflection due to the long frequency band chosen in this study. However, further investigation is required to understand why the locations for all the reflectors are not quite the same for *SKKS* and *ScS* studies.

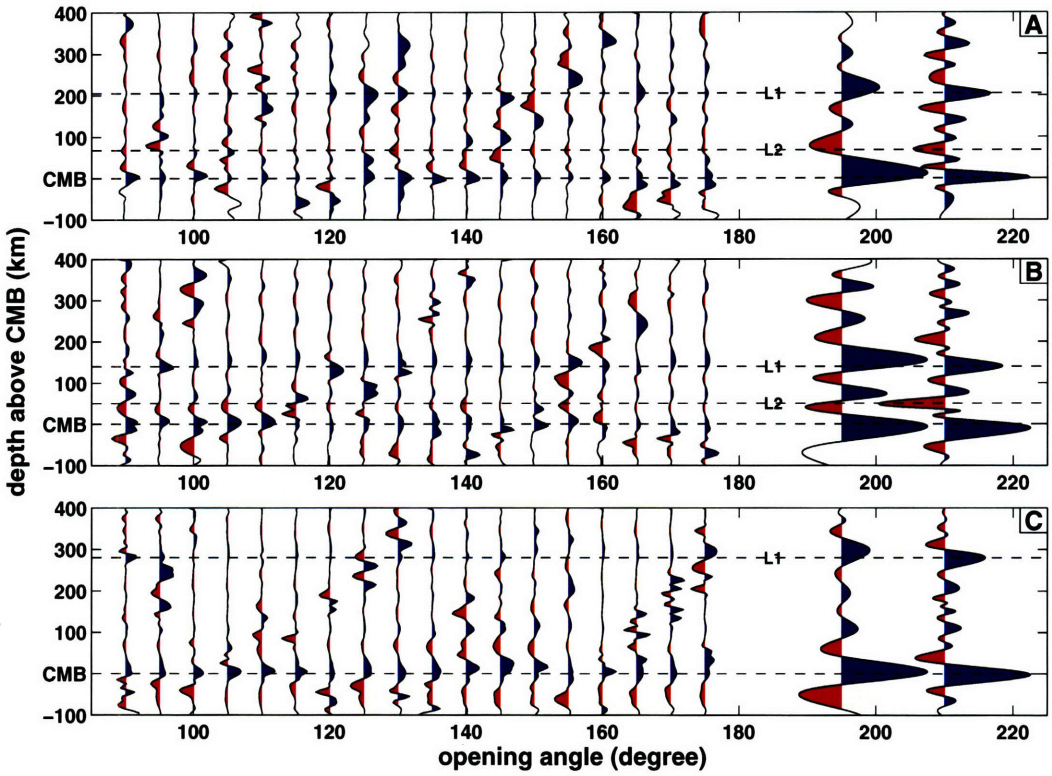

Figure 4-9: Illustration of angle gathers and angle stacks with real data for three image points marked in the blowup in Figure 4-8. L1 and L2 follow the meaning in *Van der Hilst et al.* [2007]. The right two traces are the angle stacks of SKKS data and ScS data (from Figure 3.B in *Van der Hilst et al.* [2007]).

4.5 Discussion and concluding remarks

The generalized Radon transform generates out of data a collection of angle common-image-point gathers. The generalized Radon transform maps data as a function of 5 vari-

ables globally to a set of images (of best resolved parameter combinations) parametrized by scattering angle and azimuth, which again yields a function of 5 variables. We stack over scattering azimuth. The mapping may generate artifacts in the presence of caustics that are well understood and can be suppressed. The gathers associated with a common image point admit a refined process of denoising and detecting singularities associated with scatters, the unfolding of caustics, and an analysis of (ir)regularity of variation in material properties at multiple scales. Potentially, by combining a sufficient number of bins, the local curvature of the CMB can be estimated.

The GRT is a comprehensive theory/framework. Kirchhoff migration can be viewed as a special case, typically, assuming the absence of caustics; in Kirchhoff migration one often uses surface offset as the redundant variable, which leads to fundamental artifacts in image gathers in the presence of caustics. (*Revenaugh* [1995]) is an example of classical Kirchhoff migration using all the data but loosing their sign information. Kichhoff time migration is a special case of Kirchhoff migration – it assumes straight rays in an effective medium that changes with the depth of image point (*Simon et al.* [1996]). Double beamforming (*Scherbaum et al.* [1997]) provides the input to so-called map migration (*Hedlin et al.* [1991]) revealing, geometrically, the propagation of singularities by the imaging operators. Applying the GRT to a single data point generates an image distribution with as its singular support an isochrone. If one has only a few isolated data points, one overlays the corresponding isochrones (*Lay and Young* [1996]).

In principle, *ScS* and *SKKS* can be combined (through compensation of the outer core) to form a GRT integration over migration dip directions covering entire (unit) sphere. However, one should do the integration with some caution: The integration will be sensitive to anisotropy. Conversely, we could use the integration to detect the presence of anisotropy.

The major interfaces observed with *ScS* and *SKKS* are consistent with each other (Figure 4-9). The difference in locations of the interfaces may indicate that the lower mantle is anisotropic. We expect that a joint inversion of *ScS* and *SKKS* wavefields and their precursors and coda will generate more accurate images and provide further insight on the anisotropy in the lower mantle. Figure 4-2 indicates that the best region for such a study is Eurasia, from Urals to Japan, and perhaps into SW Pacific to New Zealand.

Chapter 5

Implications of a post-perovskite lens[†]

Abstract

With the imaging technology discussed in **Chapters II** and **III** we can explore the deep mantle on an unprecedented spatial scale. As a first example we used about 80,000 broadband core-related *ScS* waves were used to image the structure at and near the core mantle boundary beneath North and Central America through a generalized Radon transform approach. Mixed-effect statistical models were applied to produce the images and estimate uncertainties. We present three profiles beneath Central and North America. Multiple, piecewise continuous interfaces in the D'' layer are observed. The observed depth variations of a widespread interface $\sim$150-300 km above the CMB correlate with tomographic S-wavespeed perturbations. This interface, which we interpret as the top of the D'' layer, is consistent with the so-called post-perovskite transition. A deeper interface, associated with a wavespeed drop, may represent the back-transition to perovskite and, thus, the base of a lowermost mantle lens rich in post-perovskite. The complexity within this lens may be due to multiple phase transitions. With the assumption that these interfaces are associated with the perovskite to post-perovskite and post-perovskite to perovskite transformations, respectively, and that the Clapeyron slope of these transitions is $\sim$10 MPaK^{-1}, we estimate that the temperature near the top of the D'' layer beneath Central America is $\sim$ 700 K colder than ambient mantle. We inferred that the temperature change across the thermal boundary layer is 1500 ± 100 K and that the temperature of the mantle at the CMB is $3,950 \pm 200$ K. The inferred local heat flux is 80-160 mWm^{-2}. The estimated average heat flux of 50-100

[†]Based on *Van der Hilst et al*, Science, 315, 1813-1817, 2007.

mWm^{-2} implies a global core flux of 7.5-15 terawatts.

5.1 Introduction

The interface between the lowermost mantle and the outer core, the core-mantle boundary (CMB), represents the most dramatic contrast in (mass) density and visco-elastic properties within the Earth's interior. It marks the change from solid silicates to liquid metals. The D'' was initially introduced by *Bullen* [1949] as a division of the lower mantle (Layer D) and was later treated as the top of the discontinuous velocity increase about 200-300 km above the CMB. Interpretation of the D'' in terms of geodynamics and mineral physics remains a huge challenge due to its unreachability and complexity (*Wysession et al.* [1998]; *Garnero* [2000]). But the recent discovery of the phase transition from perovskite (pv) to post-perovskite (ppv) in $(Mg,Fe)SiO_3$ explains many previously unexplained seismic features of the D'' layer, such as the presence of a seismic discontinuity, its Clapeyron slope, the contrast in elastic properties across it, and the anisotropy and bulk-shear velocity anti-correlation just above the CMB (e.g., *Oganov and Ono* [2004]; *Murakami et al.* [2004]; *Shim et al.* [2004]; *Iitaka et al.* [2004]; *Merkel et al.* [2006]; *Wentzcovitch et al.* [2006]).

If one can identify the seismologically inferred interfaces associated with the ppv transition, using the pressure-temperature dependence of this transition, one can, in principle, estimate the temperature variations along the interfaces. Estimating the temperature gradient – and, thus the core heat flux – is possible if we can identify seismic interfaces due to double-crossing of the phase boundary (*Hernlund et al.* [2005]; *Lay et al.* [2006]).

Seismic (transmission and normal mode) tomography has been successful in revealing long-period (smooth) changes in wavespeed (e.g., *Dziewonski* [1984]; *Van der Hilst et al.*

[1997]; *Romanowicz* [2003]). Structures at length scales far smaller than can be resolved by tomography cause wavefield scattering, including reflections and phase conversions. The scattered wavefield has been used in many studies, for instance, to estimate stochastic properties of deep mantle heterogeneity (e.g., *Hedlin et al.* [1997]; *Margerin and Nolet* [2003]), to determine variations in depth to and reflectivity of known mantle discontinuities (e.g., *Paulssen* [1988]; *Van der Lee et al.* [1996]; *Shearer and Flanagan* [1999]; *Shearer et al.* [1999]; *Gu and Dziewonski* [2002]; *Deuss and Woodhouse* [2002]; *Chambers et al.* [2005]), to explore the lowermost mantle (*Garnero* [2000]; *Castle and Van der Hilst* [2000], and many others), and to search for previously unknown interfaces (e.g., *Lay and Helmberger* [1983b]; *Revenaugh and Jordan* [1991]; *Kawakatsu and Niu* [1994]; *Vinnik et al.* [2001]; *Castle and van der Hilst* [2003]). But the most detailed constraints on D'' structure to date have come from forward modeling of shear waves reflected at or near the CMB (*Lay and Garnero* [2004]; *Helmberger and Ni* [2005]).

Waveform modeling has its drawbacks, however. It is, as yet, only practical for relatively simple 1- or 2-D structural geometries, the uniqueness of which is not easily established. Furthermore, it mainly uses signal associated with near and post critical incidence because the reflection coefficient is then sufficiently large to produce signal that can be seen (in raw data or after stacking) and modeled. However, the necessary restriction to narrow epicentral distance ranges discards most of the available seismic data and strongly limits the geographical regions beneath which the D'' can be studied. Kirchhoff migration stacks have revealed a local step in the D'' discontinuity (*Hutko et al.* [2006]), but strict selection and visual inspection of data from near-critical reflections hinders the type of D'' exploration that is needed to know whether such steps are unique phenomena, ubiquitous features, or – perhaps – partial recoveries of multiple interfaces.

5.2 Methodology and data

In this study, we first use the GRT approach (**Chapter II**, *Wang et al.* [2006]) and statistical methods (**Chapter III**, *Ma et al.* [2006]) to obtain images of two-dimensional (2-D) sections. Subsequently, we combine the information thus obtained with results from mineral physics to estimate temperatures at and above the CMB.

Three steps – illustrated in Figure 3-2 – are involved in obtaining our images. First, after careful data selection, the time windows containing the (transverse-component) *ScS* wavefields are used to construct radial image gathers of the same image point for different opening (or scatter) angles. Second, we use statistical models to estimate a single radial profile for this particular one image point from these 'common image point gathers' (see Figure 2-3). Third, lateral juxtaposition of these 1-D profiles for different image points in one particular cross-section generates a 2-D images for this section.

Estimating the temperature requires the following information. First, we need to identify a particular phase transition with a seismological interface; Second, we need to estimate the *in situ* pressure at that interface from its depth; Third, we need to know the thermodynamic (P-T) conditions, that is, the Clapeyron slope, of the stability of the mineral constituents at the interface. The pressure estimates then yield temperature. Furthermore, temperature estimates at different depths constrain the thermal gradient and, thus, under assumption of thermal conductivity, the radial heat flow across the D'' layer. For this purpose we assume a (steady state) thermal diffusion across a thermal boundary layer into a half space represented by the lower mantle. We take the first diffusion length as the thermal boundary layer (TBL) and the average temperature gradient across this layer as the thermal

gradient of the D'' layer. The temperature at depth z ($z_{cmb} = 0$) is

$$T(z) = T_{cmb} - \Delta T_{HS} erf(z/H), \tag{5.1}$$

where T_{cmb} is the temperature at the CMB and the error function (erf) describes heat diffusion into a half space and ΔT_{HS} the total temperature contrast across the half space. From ΔT_{HS} we estimate the contrast across the TBL as

$$\Delta T_{TBL} = erf(1)\Delta T_{HS} = 0.84\Delta T_{HS}. \tag{5.2}$$

For different Clapeyron slopes we calculate the temperature at positions B3 and B4 (i.e., $T(B3)$ and $T(B4)$). For a given diffusion length H, we can then find a unique solution of T_{cmb} and ΔT_{TBL} (Figure 5-5).

Compared to the studies in **Chapter II** and **III** (*Wang et al.* [2006]; *Ma et al.* [2006]), here we included data from more recent earthquakes and from another bin (of $30^\circ \times 30^\circ$; see Figure 5-1). About 80,000 transverse-component *ScS* seismograms, from $\sim 1,500$ earthquakes ($m_b > 5.2$, origin time 1988-2006) recorded at one or more of a total number of $\sim 1,200$ seismic stations, were used in this study (Figure 5-1B). For all events considered here, origin times and source locations (hypocenters) were obtained from *Engdahl et al.* [1998], hereinafter referred to as EHB. For all events in our data set, three-component broadband waveforms were retrieved from the Data Management Center of the Incorporated Research Institutions for Seismology (IRIS). For details of the data processing we refer to **Chapter II** (*Wang et al.* [2006]). The most important aspects are: band-pass filtering (accepting periods between 1-20 s), suppression of the effects of different recording systems through removal of the instrument response, estimation (and removal – by

deconvolution– from the data) of the source signature using a polarity check and a principal component analysis, amplitude normalization to the ScS reflection avoid predominance of a few large earthquakes, and correction for Earth's ellipticity.

5.3 D'' structure and temperature

Three image profiles, interpolated in 3-D space, are presented in Figure 5-2. A clear CMB is revealed, but since the existence of this boundary is not disputed, we will focus the discussion on the complex pattern of scatters and interfaces up to 400 km above it. We illustrate the GRT image after statistical analysis in the top panels in Figure 5-4. For comparison with the smooth wavespeed variations, in the bottom panels we display the tomographically inferred variations in S-wavespeed and superimpose the scatters at more than 75% confidence level. In the rest of this chapter, we focus on three structures labeled L1, L2, and L3.

5.3.1 Post-perovskite transition

Structure L1 is the most obvious structure in Figure 5-4 (except for the CMB itself). It shows up in all the three profiles. This piecewise continuous structure is unambiguous because it is laterally continuous and because it has an amplitude that is a significant fraction of the CMB peak. Many localized studies of the D'' discontinuity beneath Central America have been performed and the results were listed in Table 1 in *Thomas et al.* [2004]. Studies 1 (*Lay and Helmberger* [1983b]), 2 (*Kendall and Nangini* [1996]), 5 (*Kendall and Nangini* [1996]), and 6 (*Kendall and Shearer* [1994]) revealed a D'' elevation from 250-290 km, which are remarkably consistent with our result in profile A-A'. Studies 7 (*Ding and Helm-*

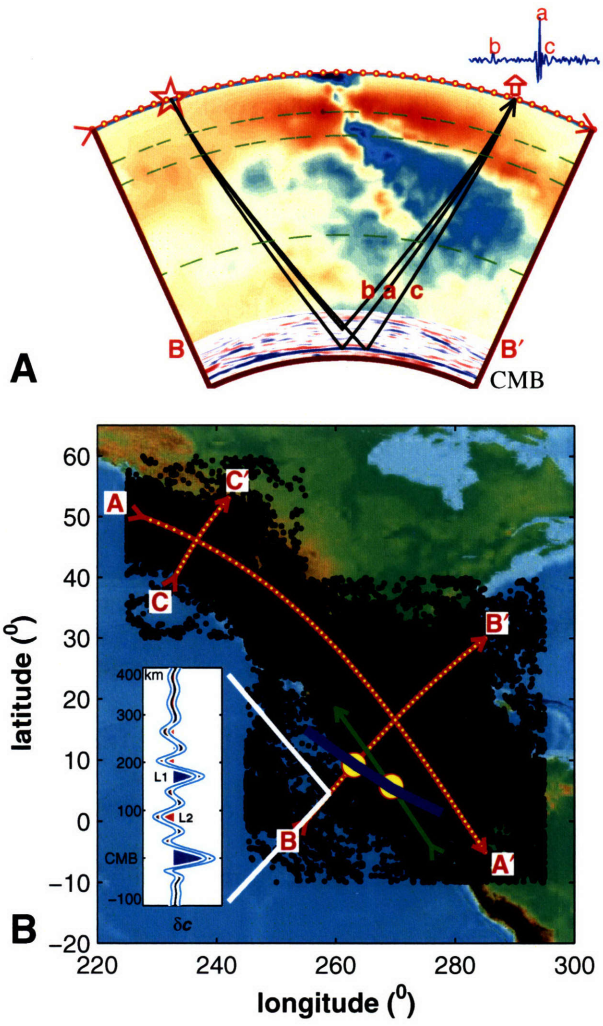

Figure 5-1: A: The upper part is the tomographic P-wavespeed variations (*Kárason and van der Hilst* [2001]). Deep subduction is observed beneath Central America. The lower 400 km is obtained with our GRT method. Superimposed on the tomography/scattering image are schematic ray paths of ScS waves reflecting at and above the CMB: a depicts specular CMB reflections, which contribute to the main ScS arrival in the seismogram shown, b depicts scattering above the CMB, which produces precursors, and c depicts non-specular reflection (at CMB or above it), which arrive mainly in the coda of ScS . B: Geographic map of the study region with bounce-points of the 80,000 ScS data used in our inverse scattering study (black dots), along with the lines of cross section for the 2-D images shown in Figures 5-2 and 5-4. At each yellow dot, the generalized Radon transform produces radial profiles of contrasts in elastic properties (inset lower left). Only structure outside the 75% confidence level (thin lines) is discussed here. The green line depicts the profile in *Hutko et al.* [2006]. The two large yellow dots are where the steep D'' topography are observed by this study and *Hutko et al.* [2006]. The thick blue line delineates the possible boundary between the cold mantle (right) and the hot mantle (left).

berger [1997]), 8 (*Kendall and Shearer* [1994]), and 10 (*Reasoner and Revenaugh* [1999]) reveals a D'' elevation from 180-200 km, much smaller than other studies. These observations are consistent with the steep D'' topography change observed in profile B-B'. While most of the previous studies emphasized only on selected (discrete) points or small CMB patches beneath Caribbean, the GRT described here constrains L1 over thousands of kilometers (and millions of kilometer squared). We only present three profiles here, but other sections can be generated in order to illustrate the 3-D structure of the D'' discontinuity.

The bottom panels of Figure 5-4 reveal a strong correlation between the depth variation of L1 with the change in S-wavespeed. The origin of the wavespeed variations is not well known, and it is likely that chemical heterogeneity plays a role (*Van der Hilst and Kárason* [1999]). If the S-wavespeed change is mainly contributed from a thermal origin, however, our L1 structure indicates a positive Clapeyron slope in almost everywhere beneath Central America. *Sidorin et al.* [1999] assumed that the tomographic S-wavespeed variations have a thermal origin only. Starting from a reference point where the position of D'' discontinuity was best constrained, assuming a Clapeyron slope 6 MPaK^{-1}, they integrate over the whole Earth and obtained the map of the D'' elevation. In a recent analysis, using similar approach *Sun et al.* [2006]) chooses a different reference point and results in a D'' elevation deeper by ~ 95 km. But our L1 structures is remarkably close to the original prediction (solid blue lines in Figure 5-4, lower panels). Our L1 structure has a larger topography than predicted from $\gamma = 6$ MPaK^{-1}. This suggests that the actual Clapeyron slope, γ, must be larger than 6 MPaK^{-1}, which would be consistent with the estimates from mineral physics (*Oganov and Ono* [2004]; *Tsuchiya et al.* [2004]; *Iitaka et al.* [2004]; *Hirose et al.* [2006]; *Hernlund and Labrosse* [2007]), that the amplitude of the real S-wavespeed variations was tomographically underestimated, or that compositional effects on shear wavespeed are more important

than assumed by *Sidorin et al.* [1999]. But – at least qualitatively – L1 is consistent with predictions from the post-perovskite (ppv) transition. If this interpretation is correct, we can estimate the (lateral variations in) temperature from lateral changes in interface depth. For $\gamma_{ppv} = 10$ MPaK^{-1}, the temperature difference between A1 and A2 is ~ 600 K, and the difference between B1 and B2 is ~ 700 K. In turn, for Grand's S-model (*Grand* [2002]) these values imply an empirical $\mathrm{dln}V_S/\mathrm{d}T \approx 1.5 \times 10^{-5} K^{-1}$. (NB this conversion factor implicitly accounts for effects of compositional heterogeneity and can be used to estimate 3-D thermal anomalies from the shear wavespeed tomography.)

The position where a steep topography is observed in our study along section B-B' is close to where a step-like jump is reported by *Hutko et al.* [2006]. Interestingly, even though our profile B-B' is perpendicular to theirs a topography jump is observed in both. This suggests that these 2-D sections sample a NW-SE trending structure This also demonstrates that a 3-D scan of the area is important for a better understanding of the D'' topography. Seismic tomography (Figure 5-1) reveals that some cold debris sinks all the way down to the CMB beneath Central and south America (*Grand* [2002]). *Hutko et al.* [2006] suggest that the step-like jump is due to the folding of the subducted slab. *Tan et al.* [2002] and *Ribe et al.* [2007] showed that slab buckling is, indeed, likely to occur in this region.

5.3.2 Post-perovskite lens

The GRT results suggest the presence of complex structures in between interface L1 and the CMB. Of particular interest are the ones labeled L2. The red color is used to mark a S-wavespeed decrease (see L2 in Figure 5-1). *Thomas et al.* [2004] also reported a wavespeed decrease above CMB, but the location was different from our results pertaining to L2 (it was deeper, that is, closer to the CMB, and more parallel to the D'' interface).

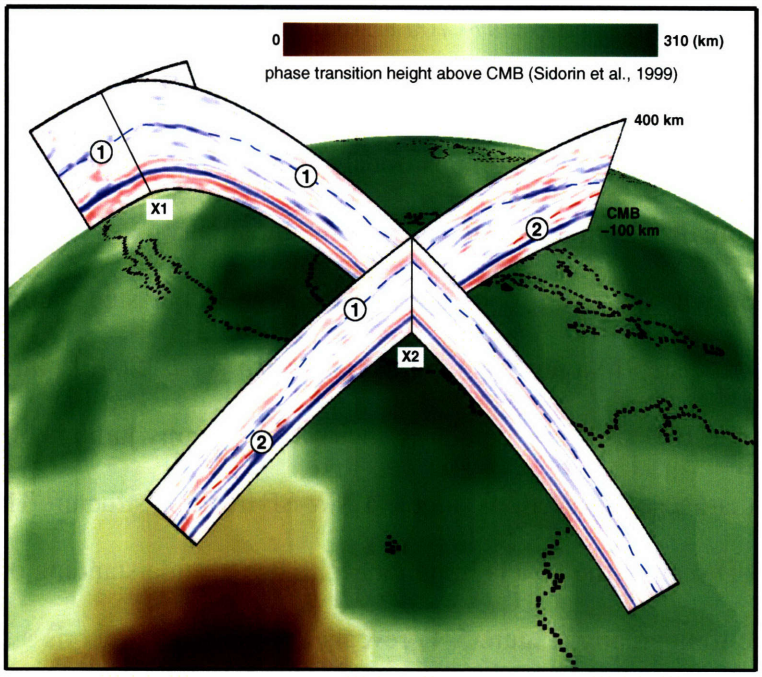

Figure 5-2: Three-dimensional exploration seismology of the lowermost mantle. Seismic images of the lowermost mantle (CMB to 400 km above it) are produced by lateral juxtaposition of radial general Radon transform profiles (Figure 3-2) calculated at image points along the section lines shown in Figure 5-1. Structure outside 75% confidence bands (*Ma et al.* [2006]) includes the CMB (at 0 km) and several scatter interfaces above it. Thinly dashed lines indicate scatter interfaces (L1, L2) highlighted in Figure 5-4. This 3-D rendition illustrates the large spatial scales over which inverse scattering with the *ScS* wavefield can be used to explore the lowermost mantle. The intersection points of sections are labeled X1 and X2. The background color is the D'' elevation above the CMB predicted by *Sidorin et al.* [1999].The correlation between L1 and the predicted values is very good.

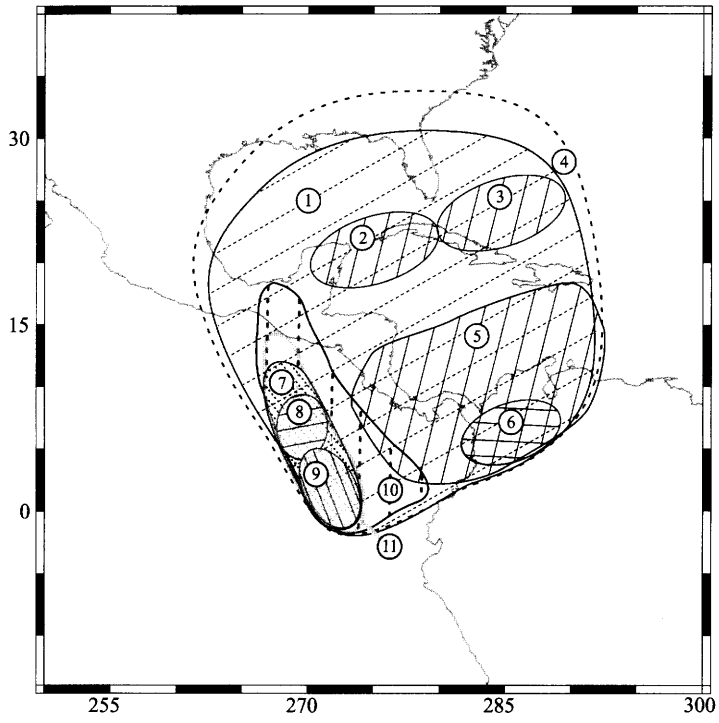

Figure 5-3: After *Thomas et al.* [2004]: Shaded regions and numbers correspond to the specific studies listed in Table 1 in *Thomas et al.* [2004].

A wavespeed drop that is too close to the CMB cannot be resolved due to the presence of side lobes of the CMB reflection. Two dipping interfaces on both sides of profile B-B' are observed. They gradually dip towards and approach the CMB where the shear wavespeed is high and the temperature is – assumed to be – low. Even though it can't be observed directly with our method, we argue that L2 is, effectively, at the CMB between 1000-2000 km along B-B'. Since the Clapeyron slope for ppv transition is positive, the lower the mantle temperature the deeper the back transition (ppv $\rightarrow$ pv), if exists. However, the previous estimated CMB temperature, T_{cmb}, varies form 3,750 K to 4,800 K (*Stacey* [1992]; *Alfe et al.* [2002]; *Knittle and Jeanloz* [1991]), which is higher than $T_{ppv,cmb}$. This indicates that ppv is instable at the CMB and a back transition must occur somewhere above the CMB, and close to it in regions with very low lowermost mantle temperatures. The

second crossing is hard to detect with *ScS* data (*Flores and Lay* [2005]), not only because it could be masked by the CMB side lobes but because the amplitude of seismic signal (in particular, S_dS with large epicentral distances) is often weaker (NB: this is not true for the underside reflection S'^dS', see **Chapter IV**). But recent studies provide evidence for it near the CMB beneath the Pacific (*Lay et al.* [2006]) and the Cocos plate (*Sun et al.* [2006]). The topography of the L2 is anti-correlated with that of L1, which supports the interpretation in terms of the back transformation and provides tantalizing evidence for the existence of a lowermost mantle lens that is rich in post-perovskite. From the images we estimate that this lens has a length of > 2300 km and a thickness of ~ 250 km.

5.3.3 Mantle temperature near CMB

In Section 5.3.1 we showed how we can use interface L1 to estimate the temperature variations laterally at 150–300 km above the CMB. Unlike previous studies in which double crossing points were used to infer the thermal structure above the CMB (*Hernlund et al.* [2005]; *Lay et al.* [2006]), our images provide double crossing points over laterally contiguous parts of the D''. In section 5.2 we explained how one could use such double crossings to estimate radial temperature variations. However, in equation 5.1 we have three unknown variables T_{cmb}, ΔT_{HS}, and H, while we have only two data points for every double crossing pair. Therefore, we need to explore a range of parameter combinations to obtain insight into the trade-offs between them. For one double crossing pair we first find a solution of T_{cmb} and ΔT_{HS} for a given TBL thickness, H, and a fixed Clapeyron slope, $\gamma_{ppv} = 10$ MPaK^{-1}. Subsequently, we varied H and obtained a series of T_{cmb} and ΔT_{HS}. The results are shown in Figure 5-5 for double crossing pair (B3, B4). For other pairs of double crossing points, we obtained similar results. Not surprisingly, H increases with the increasing

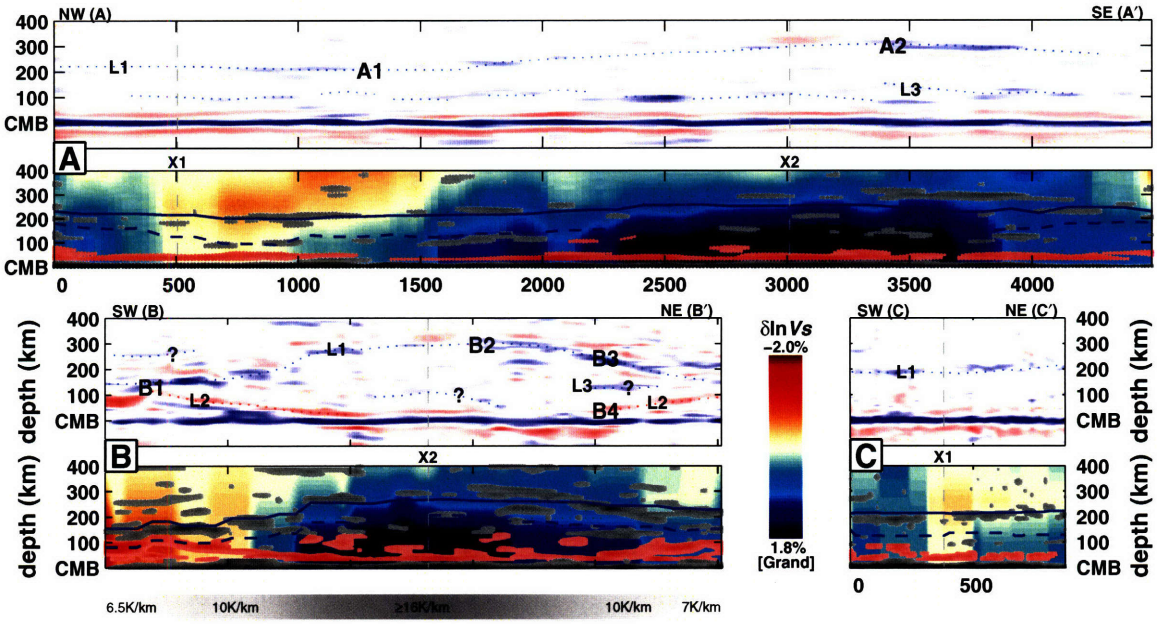

Figure 5-4: Reflectivity from inverse scattering, at more than 75% confidence (*Ma et al.* [2006]) (top) and S-speed (dlnVS) from tomography (*Grand* [2002]) (bottom). Scatter images are obtained by interpolation between GRT profiles (Figure 5-1B, inset) calculated every 1° (60 km at CMB) along sections A-A', B-B', and C-C' in Figure 5-1. For the frequencies and incidence angles used, the radial resolution is 10 km. The color scale for tomography is given between B-B' and C-C'. In the top panels, L1, L2, L3 label the scatter interfaces (thinly dashed) discussed in the text: in the bottom panels, the associated scatterers (visually enhanced) are superimposed on the tomography profiles, with dark grey depicting positive reflections and pink/red negative contrasts. Interface L1 aligns increases of wavespeed with increasing depth; L2 delineates a decrease; L3 is more ambiguous but generally coincides with a wavespeed increase. Whereas L1 and L2 are piecewise continuous, L3 has an intermittent, "en echelon" appearance. The solid (dashed) blue lines in the bottom panels depict the phase transition location predicted by (*Sidorin et al.* [1999]) ((*Sun et al.* [2006])). Points A1-2 and B1-4 on L1 and L2 are used for temperature calculations (Figure 5-5). The grey scale below B-B' depicts the lateral variation in temperature gradient along the CMB (for $\gamma_{ppv} = 10$ MPa/K). In the central portion of the section dT/dz cannot be determined directly because the occurrence of the double crossing cannot be resolved.

of temperature, that is, from colder to warmer D'' regions. T_{cmb} and ΔT_{HS} are relatively stable for $H = 100$ to 200 km. Thus with an assumption of $\gamma_{ppv} = 10$ MPaK^{-1}, we infer $T_{cmb} = 3950 \pm 200$ K and $\Delta T_{TBL} = 1500 \pm 100$ K (Figure 5-5C). These estimates depend on the actual pressure-temperature of ppv transition. The uncertainties for such kind measurements can reach 200-400 K or 5-100 GPa. But the following estimate depends only on the Clapeyron slope. For a Clapeyron slope $\gamma_{ppv} = 10$ MPaK^{-1}:

$$T_{cmb} \approx T_{ppv,cmb} + 290K. \tag{5.3}$$

For a Clapeyron slope $\gamma_{ppv} = 10$ MPaK^{-1}, $T_{ppv,cmb} \approx 3700$ K. Within error, our estimate of T_{cmb} is consistent with the estimate of 4100 K in *Lay et al.* [2006].

Since the out-core is liquid, the temperature just below the CMB must be larger than the melting temperature of core materials there. It is interesting to find that experimental and theoretical research has yielded estimates of the melting temperatures of pure Fe in a wide range of values, from 3200 ± 200 K (*Boehler* [1993]) to 4800 ± 200 K (*Williams et al.* [1987]). The addition of minor elements could reduce the melt temperature by up to 1000 K, so that the melt temperature of the alloy could be as low as 2200 K, but *Boehler et al.* [1995] proves that the effect of oxygen and sulfur, the major trace elements in the outer-core, is fairly small. If our estimate of the temperature of the CMB is correct, the melting temperature in *Williams et al.* [1987] may be too high.

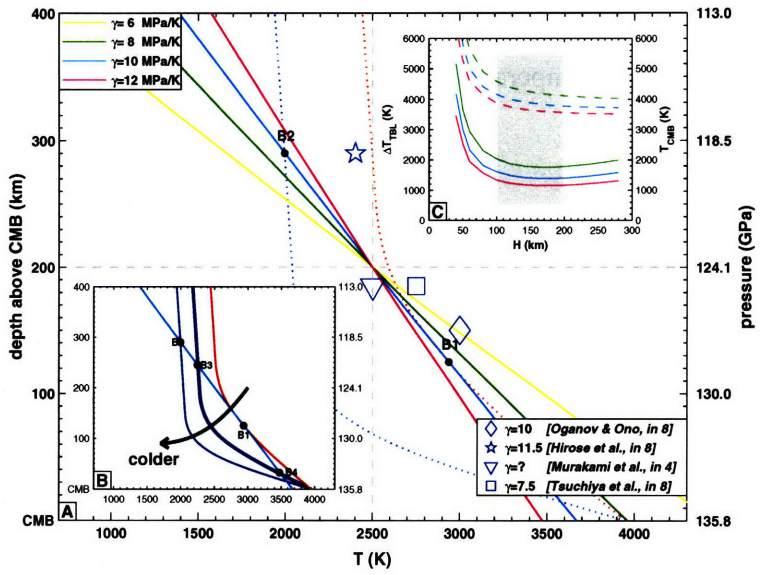

Figure 5-5: (A) Temperature (abscissa) versus distance above CMB (ordinate). As a point of reference for producing absolute temperature and depth (pressure), we assume Pppv=124 GPa and Tppv=2,500K, and we consider Clapeyron slopes γ_{ppv}=6, 8, 10, and 12 MPa/K. If L1 in Figure 5-4 represents the ppv transition, the temperature at points B1,2 can be estimated: for γ_{ppv}=10 MPa/K, T_{ppv},B2=2,000 K and Tppv,B1=2,900 K. Dotted lines depict estimated geotherms through B1 and B2: the shallow part is adiabatic but toward the CMB the conductive geotherm is described as an error function (the change of adiabat to conductive is not known, however). (B) Cold, intermediate, and hot geotherms. If the mantle temperature decreases a ppv transition occurs at increasing distances above CMB and a second crossing would occur at decreasing height above CMB. The geotherm through the ppv transition (B3) and the back-transformation (B4) is calculated using a thickness of the hypothetical boundary layer, H, of 150 km. (C) Mantle temperature, T_{cmb}, at the CMB (dashed lines) and the temperature change, T_{TBL}, across the thermal boundary layer (solid lines) as a function of TBL thickness, H, and for γ_{ppv}=8, 10, and 12 MPa/K. Stable estimates are obtained for H=100-200 km (grey shading).

5.3.4 Core heat flux

In the previous subsection, we estimated the temperature drop, ΔT_{TBL}, across the TBL for given H. Then the average temperature gradient across the TBL can be estimate as

$$\frac{dT}{dz} = \frac{\Delta T_{TBL}}{H}. \tag{5.4}$$

For $H = 100 - 200$ km, we estimate $\frac{dT}{dz}$=7-16 Kkm^{-1}. *Lay et al.* [2006] gave values of 8.5 ± 2.5 Kkm^{-1}. After obtaining the thermal gradient, we can estimate the heat flux

$$q = -\kappa \frac{dT}{dz} = -\kappa \frac{\Delta T_{TBL}}{H}, \tag{5.5}$$

where κ is the thermal conductivity.

The estimates of the regional and global heat flux are highly uncertain for the following reasons: First, the estimate of $\frac{dT}{dz}$, even if it's accurate, is a point estimation. We took the higher bound (16 Kkm^{-1}) as the representative of cold D'' area and the lower bound (7 Kkm^{-1}) as the globe average. However, the actual gradient would be larger than these values if the wavespeed perturbations are (partially) due to chemical variations (*Van der Hilst and Kárason* [1999]). Second, the value of thermal conductivity, κ, isn't well constrained. A $\kappa = 10$ Wm^{-1}K^{-1} (*Stacey* [1992]) has been used by many people. However, one cannot rule out the possibility that κ is as small as 5 Wm^{-1}K^{-1} (*Hofmeister* [1999]; *Van den Berg et al.* [2005]). Nevertheless, we infer the global heat loss of 7.5 to 15 terawatts, which is higher than conducted along the core adiabat.

5.3.5 Other D'' interfaces?

While structure L1 and L2 are semi-continuous, the structure labeled L3 is a lot more scattered. With the current study, it's difficult to argue if it's a partial detection of interfaces or due to point scatters. The possible buckling of the slabs, along with preserved compositional heterogeneities, can perhaps provide some of the complexity detected here. Alternatively, local changes in temperature or chemical composition (e.g., in iron content (*Williams et al.* [1987]) or partitioning between pv and magnesiowüstite (*Kobayashi et al.* [2005]) can readily create multiple crossings fo the phase boundary. Alternating, irregular ppv and pv lenses (or layers) may be more realistic than a single ppv lens (*Shim* [2005]) and can explain some the multiple, but intermittent, scatter interfaces. This calls for further investigation by means of joint imaging and geodynamical and petrological modeling.

5.4 Summary

The observed D'' interfaces allow us to estimate the lateral temperature variations in the lowermost mantle beneath Central and North America. At a site of deep subduction, Central America, the lowermost mantle is ~ 700 K colder than the surrounding mantle. Double crossing points from the seismically computed post-perovskite lens give determinations of the temperature of the CMB and the thermal gradient across the thermal boundary layer (TBL). We infer the temperature of the CMB $T_{cmb} = 3950 \pm 200$ K, which is consistent with the Iron melting temperature given in *Boehler* [1993]. We allow the thermal conductivity to change from 5-10 $\mathrm{Wm^{-1}K^{-1}}$ and estimate the global average heat flux of 50-100 $\mathrm{mWm^{-2}}$ and a total heat loss of 7-15 TW. The latter is lower than the value obtained by *Lay et al.* [2006] but higher than conducted along the core adiabat. This estimate is consistent

with the suggestion by *Nimmo et al.* [2004], made to satisfy the current inner-core size and the high heat flow, that the core contains 400 ppm potassium.

5.5 Future work

The three image profiles, interpolated in 3-D space, show the potential of using our methods for the systematic, 3-D, high resolution investigation of the lowermost mantle over larger regions. The global *ScS* and *SKKS* (see Figure 4-2) midpoint distribution indicates that similar high resolution D'' imaging should be possible in many places, for example, beneath a large area of Eurasia. Admittedly, the first images resulting from our methods are ambiguous in some places. We expect that the methods will be improved in many ways, such as further refinement of inverse scattering, better data pre-processing techniques, and the addition of data. We can also use waveform modeling for targets of particular interest to give new understanding of the D" layer. We expect that a joint inversion of *ScS* and *SKKS* wavefields and their precursors and coda will generate more accurate images and provides further insight on the anisotropy in the lower mantle.

Chapter 6

Concluding remarks

6.1 Summary

The generalized Radon transform was initially introduced in exploration seismology as a method for detecting hydrocarbon reservoirs. In this thesis, we developed a mixed fluid-solid generalized Radon transform approach to inverse scattering adapted for coda waves in global seismology. A theory for the statistical inference of singularities (discontinuities) was also developed to produce images of interfaces and estimate their uncertainty.

With the outlined methods, we scan the lowermost mantle beneath Central and North America from both topside (*ScS*) and underside (*SKKS*). About 80,000 transverse-component wavefields were used for *ScS* study and $\sim 18,000$ radial-component wavefields were used for *SKKS* study. The major findings of this thesis are: (1) A clear CMB was revealed. The existence of this boundary is not disputed, but this is the first image of the CMB with a length scale of $> 3,500$ km; (2) The D'' discontinuity interface was constrained over thousands of kilometers beneath Central and North America. The depth of this interface varies from $\sim 150 - 300$ km. Combined with the results obtained from mineral physics, we es-

timated the lateral temperature variation along the D'' discontinuity and found that the D'' beneath Central American, a site of deep subduction, is $\sim$ 700 K colder than the ambient mantle; (3) Steep topography of the top of D'' layer was observed. This observation along with the result by *Hutko et al.* [2006] suggests an existence of a NW-SE trending structure; (4) A lowermost mantle lens that is rich in post-perovskite was observed. With the double crossing points from this lens, assuming a Clapeyron slope of 10 MPa^{-1}, I inferred the temperature of the CMB as 3950 ± 200 K, the thermal gradient across the thermal boundary layer from $7 - 16$ Kkm^{-1}. Using the inferred thermal gradient, assuming a thermal conductivity $\kappa = 5$ $Wm^{-1}K^{-1}$, I estimated that global heat loss is from 7.5 to 15 terawatts; (5) A rich pattern of scatterers and interfaces was observed between the top of the ppv lens and the CMB. This may indicate multiple phase transitions within the D'' layer; (6) The major interfaces observed with *ScS* and *SKKS* are consistent with each other. The difference in locations of the interfaces may indicate that the lower mantle is radially anisotropic.

The D'' discontinuity interfaces observed agree with various previous studies (*Lay and Helmberger* [1983b]; *Kendall and Shearer* [1994]; *Kendall and Nangini* [1996]; *Ding and Helmberger* [1997]; *Reasoner and Revenaugh* [1999]; *Sidorin et al.* [1999]; *Hernlund et al.* [2005]). The estimated temperature of the CMB is consistent with iron melting temperature obtained by *Boehler* [1993]. The inferred global heat loss of 7.5 to 15 terawatts is lower than the value obtained by *Lay et al.* [2006] but higher than conducted along the core adiabat. This estimate is consistent with the suggestion by *Nimmo et al.* [2004], made to satisfy the current inner-core size and the high heat flow, that the core contains 400 ppm potassium.

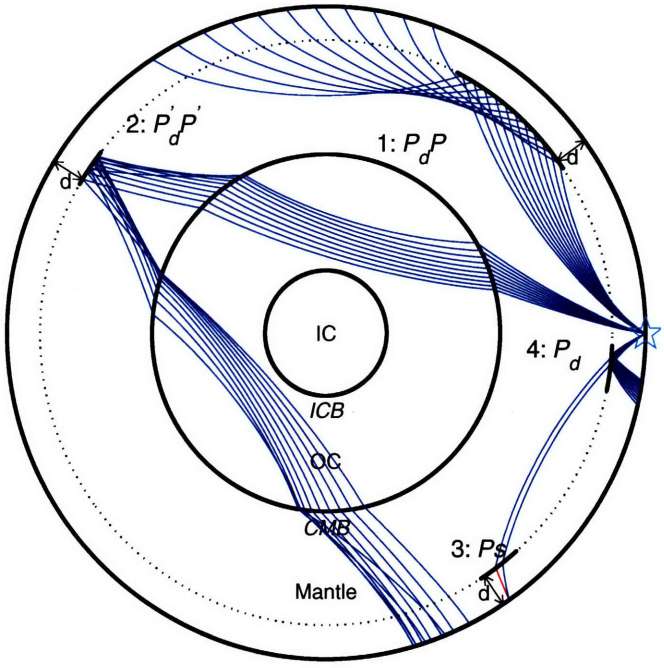

Figure 6-1: Schematic ray geometry of a variety of phases that can be used in mantle discontinuity studies (here shown for *P*-waves).

6.2 Future Work

The three image profiles, interpolated in 3-D space, show the potential of using our methods for the systematic, 3-D, high resolution investigation of the lowermost mantle over larger regions. The global *ScS* and *SKKS* (see Figure 4-2) midpoint distribution indicates that similar high resolution D'' imaging should be possible in many places, for example, beneath a large area of Eurasia. Admittedly, the first images resulting from our methods are ambiguous in some places. We expect that the methods will be improved in many ways, such as further refinement of inverse scattering, better data pre-processing techniques, and the addition of data. We can also use waveform modeling for targets of particular interest to give new understanding of the *D*" layer. We expect that a joint inversion of *ScS* and *SKKS* wavefields and their precursors and coda will generate more accurate images and

provide further insight on the anisotropy in the lower mantle. Figure 4-2 indicates that the best region for such a study is Eurasia, from Urals to Japan, and perhaps into SW Pacific to New Zealand.

As presented in Figure 3-2, the GRT always produces a series of radial profiles for the same target image point. This redundancy provides not only the possibility to statistical analysis but feasibility to characterize the interfaces (e.g., the type of jump across interfaces). Such regularity estimation is a topic of research ongoing.

Figure 6-1 shows different types of seismic arrivals which can be used for a GRT study. With one or more of the phases presented here, and with only slight modifications of the GRT method, we can, in principle, study all the possible discontinuities in the mantle. We can also apply a similar approach to imaging the inner-core boundary (ICB).

List of Figures

Bibliography

Aki, K., and P. G. Richards (1980), *Quantitative seismology: Theory and methods*, vol. 1, Freeman, San Francisco.

Alfe, D., M. J. Gillan, and G. D. Price (2002), Composition and temperature of the Earth's core constrained by combining ab initio calculations and seismic data, *Earth Planet. Sci. Lett.*, *195*, 91–98.

Avants, M., T. Lay, and E. J. Garnero (2006a), A new probe of ULVZ *S*-wave velocity structure: Array stacking of *ScS* waveforms, *Geophys. Res. Lett.*, *33, L073145*, doi:10.1029/2005GL024,989.

Avants, M., T. Lay, S. A. Russell, and E. J. Garnero (2006b), Shear velocity variation within the D'' region beneath the central Pacific, *J. Geophys. Res.*, *111, B05305*, doi:10.1029/2004JB003,270.

Beylkin, G. (1985), Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform, *J. of Math. Phys.*, *26*, 99–108.

Beylkin, G., and R. Burridge (1990), Linearized inverse scattering problems in acoustics and elasticity, *Wave Motion*, *12*, 15–52.

Beylkin, K. (1984), The inversion problem and applications of the generalized Radon transform, *Comm. Pure Appl. Math.*, *37*, 579–599.

Bleistein, N. (1987), On imaging of reflectors in the Earth, *Geophysics*, *52*(6), 931–942.

Boehler, R. (1993), Temperatures in the Earth's core from melting-point measurements of iron at high static pressures, *Nature*, *363*, 534–536.

Boehler, R., A. Chopelas, and A. Zerr (1995), Temperature and chemistry of the core-mantle boundary, *Chem. Geol.*, *120*, 199–205.

Bostock, M. G., S. Rondenay, and J. Shragge (2001), Multiparameter two-dimensional inversion of scattered teleseismic body waves, *J. Geophys. Res.*, *106*, 30,771–30,782.

Brandsberg-Dahl, S., M. V. de Hoop, and B. Ursin (2003), Focusing in dip and ava compensation on scattering-angle/azimuth common image gathers, *Geophysics*, *68*, 232–254.

Buffett, B. A., E. J. Garnero, and R. Jeanloz (2000), Sediments at the top of the Earth's core, *Science*, *290*, 1338–1342.

Bullen, K. E. (1949), Compressibility-pressure hypothesis and the Earth's interior, *Month. Not. R. Astr. Soc., Geophys. Suppl.*, *5*, 5355–5368.

Burridge, R., M. V. de Hoop, D. Miller, and C. Spencer (1998), Multiparameter inversion in anisotropic media, *Geophys. J. Int.*, *134*, 757–777.

Castle, J. C., and R. D. Van der Hilst (2000), The core-mantle boundary under the gulf of alaska: No ULVZ for shear waves, *Earth Planet. Sci. Lett.*, *176*, 311–321.

Castle, J. C., and R. D. van der Hilst (2003), Searching for seismic scattering off deep mantle interfaces between 800 and 2000 km depth, *J. Geophys. Res.*, *108*, doi:10.1029/2001JB000,286.

Červený, V. (2001), *Seismic ray theory*, Cambridge University Press, Cambridge.

Chambers, K., J. H. Woodhouse, and A. Deuss (2005), Topography of the 410-km discontinuity from *PP* and *SS* precursors, *Earth Planet. Sci. Lett.*, *235*, 610–622.

Chapman, C. H. (1978), New method for computing synthetic seismograms, *Geophys. J. Int.*, *54*, 481–518.

Choy, G. L. (1977), Theoretical seismograms of core phases calculated by frequency-dependent full wave theory and their interpretation, *Geophys. J. Roy. Astron. Soc.*, *51*, 275–311.

Clayton, R. W., and R. H. Stolt (1981), A Born-WKBJ inversion method for acoustic reflection data, *Geophysics*, *46*, 1559–1567.

De Hoop, M. V., and N. Bleistein (1997), Generalized Radon transform inversions for reflectivity in anisotropic elastic media, Inverse Problems, *13*, 669–690.

De Hoop, M. V., and S. Brandsberg-Dahl (2000), Maslov asymptotic extension of generalized radon transform inversion in anisotropic elastic media: A least-squares approach, *Inverse Problems*, *16*, 519–562.

De Hoop, M. V., and C. Spencer (1996), Quasi Monte-Carlo integration over $S^2 \times S^2$ for migration $\times$ inversion, *Inverse Problems*, *12*, 219–239.

De Hoop, M. V., R. Burridge, C. Spencer, and D. Miller (1994), Generalized Radon transform amplitude versus angle (GRT/AVA) migration/inversion in anisotropic media, in *Proc SPIE 2301*, pp. 15–27, SPIE.

De Hoop, M. V., C. Spencer, and R. Burridge (1999), The resolving power of seismic amplitude data: An anisotropic inversion/migration approach, *Geophysics*, *64*, 852–873.

Deuss, A., and J. H. Woodhouse (2002), A systematic search for upper mantle discontinuities using *SS* precursors, *Geophys. Res. Lett.*, *29*, doi:10.1029/2002GL014,768.

Ding, X. M., and D. V. Helmberger (1997), Modeling D'' structure beneath Central America with broadband seismic data, *Phys. Earth Planet. Inter.*, *101*, 245–270.

Douma, H., and M. V. de Hoop (2006), Explicit expressions for pre-stack map time-migration in isotropic and VTI media and the applicability of map depth-migration in heterogeneous media, *Geophysics*, p. in print.

Dziewonski, A. M. (1984), Mapping the lower mantle - determination of lateral heterogeneity in *P*-velocity up to degree and order-6, *J. Geophys. Res., 89 (NB7)*, 5929.

Engdahl, E. R., R. D. van der Hilst, and R. P. Buland (1998), Global teleseismic earthquake relocation from improved travel times and procedures for depth determination, *Bull. Seis. Soc. Am., 88 (3)*, 722–743.

Flores, C., and T. Lay (2005), The trouble with seeing double, *Geophys. Res. Lett, 24*, L24,305.

Garnero, E. J. (2000), Heterogeneity of the lowermost mantle, *Ann. Rev. Earth Planet. Sci., 28*, 509–537.

Golub, G., and C. Van Loan (1989), *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, Baltimore, MD.

Grand, S. P. (2002), Mantle shear-wave tomography and the fate of subducted slabs, *Phil. Trans. Roy. Soc., A 360*, 2475–2491.

Gu, C. (2002), *Smoothing Spline ANOVA Models*, Springer-Verlag, New York.

Gu, C., and P. Ma (2005), Optimal smoothing in nonparametric mixed-effect models, *Ann. Statist., 33*, 1357–1379.

Gu, Y. J., and A. M. Dziewonski (2002), Global variability of transition zone thickness, *J. Geophys. Res., 107*.

Guillemin, V. (1985), *Pseudodifferential operators and applications (Notre Dame, Ind., 1984)*, chap. On some results of Gel'fand in integral geometry, pp. 149–155, Amer. Math. Soc., Providence, RI.

Guo, W. (2002), Functional mixed effects models, *Biometrics*, *58*(1), 121–128.

Hagedoorn, J. G. (1954), A process of seismic reflection interpretation, *Geophysical Prospecting*, *2*, 85–127.

Hedlin, M. A., J. B. Minster, and J. Orcutt (1991), Beam-stack imaging using a small aperture array, *Geophys. Res. Lett.*, *18*, 1771–1774.

Hedlin, M. A. H., P. M. Shearer, and P. S. Earle (1997), Seismic evidence for small-scale heterogeneity throughout the Earth's mantle, *Nature*, *387*, 145–150.

Helmberger, D. V., and S. Ni (2005), Seismic modeling constraints on the South African superplume, *Geophys. Monograph. Ser., Am. Geophys. Uni.*, *160*, 63–82.

Hernlund, J. W., and S. Labrosse (2007), Geophysically consistent values of the perovskite to post-perovskite transition Clapeyron slope, *Geophys. Res. Lett.*, *34*, L05,309, doi:10.1029/2006GL028,961.

Hernlund, J. W., C. Thomas, and P. J. Tackley (2005), A doubling of the post-perovskite phase boundary and structure of the Earth's lowermost mantle, *Nature*, *434*, 882–886.

Hirose, K., R. Sinmyo, N. Sata, and Y. Ohishi (2006), Determination of post-perovskite phase transition boundary in $MgSiO_3$ using Au and MgO pressure standards, *Geophys. Res. Lett.*, *33*, L01,310.

Hofmeister, A. M. (1999), Mantle values of thermal conductivity and the geotherm from phonon lifetimes, *Science*, *283*, 1699–1706.

Hutko, A. R., T. Lay, E. J. Garnero, and J. Revenaugh (2006), Seismic detection of folded, subducted lithosphere at the core-mantle boundary, *Nature*, *441*, 333–336, doi:10.1038/nature04,757.

Iitaka, T., K. Hirose, K. Kawamura, and M. Murakami (2004), The elasticity of the $MgSiO_3$ post-perovskite phase in the lowermost mantle, *Nature*, *430*, 442–445.

Kárason, H., and R. D. van der Hilst (2001), Tomographic imaging of the lowermost mantle with differential times of refracted and diffracted core phases (*PKP*, P_{diff}), *J. Geophys. Res.*, *106*, 6569–6588.

Kawakatsu, H., and F. L. Niu (1994), Seismic evidence for a 920-km discontinuity in the mantle, *Nature*, *371*, 301–305.

Kendall, J.-M., and C. Nangini (1996), Lateral variations in D'' below the Caribbean, *Geophys. Res. Lett.*, *23*, 399–402.

Kendall, J.-M., and P. M. Shearer (1994), Lateral variations in D'' thickness from long period shear wave data, *J. Geophys. Res.*, *99*, 1575–1590.

Kennett, B. L. N., and O. Gudmundsson (1996), Ellipticity corrections for seismic phases, *Geophys. J. Int.*, *127*, 40–48.

Kennett, B. L. N., E. R. Engdahl, and R. P. Buland (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, *122*, 108–124.

Kim, Y.-J., and C. Gu (2004), Smoothing spline Gaussian regression: More scalable computation via efficient approximation, *J. Roy. Statist. Soc. Ser. B*, *66*, 337–356.

Kleyn, A. (1977), On the migration of reflection time contour maps, *Geophys. Prosp.*, *25*, 125–140.

Knittle, E., and R. Jeanloz (1991), The high-pressure phase diagram of $Fe_{0.94}O$: A possible constituent of the Earth's core, *J. Geophys. Res.*, *96*, 16,169–16,180.

Kobayashi, Y., T. Kondo, E. Ohtani, N. Hirao, N. Miyajima, T. Yagi, T. Nagase, and T. Kikegawa (2005), Fe-Mg partitioning between (Mg, Fe)SiO_3 post-perovskite, perovskite, and magnesiowüestite in the Earth's lower mantle, *Geophys. Res. Lett.*, *32*, L19,301.

Lay, T., and E. J. Garnero (2004), *The state of the planet*, Geophys. Monogr. Ser., AGU, Washington, D.C.

Lay, T., and D. V. Helmberger (1983a), A shear velocity discontinuity in the lower mantle, *Geophys. Res. Lett.*, *10*, 63–66.

Lay, T., and D. V. Helmberger (1983b), A lower mantle S-wave triplication and the shear velocity structure of D'', *Geophys. J. R. Astron. Soc.*, *75*, 799–837.

Lay, T., and C. J. Young (1996), Imaging scattering structures in the lower mantle by migration of long-period S waves, *J. Geophys. Res.*, *101 (B9)*, 20,023–20,040.

Lay, T., J. Hernlund, E. J. Garnero, and M. S. Thorne (2006), A post-perovskite lens and D'' heat flux beneath the Central Pacific, *Science*, *314*, 1272–1276.

Lumley, D. E., J. F. Claerbout, and D. Bevc (2001), Anti-aliased Kirchhoff 3-D migration, *SEP*, *80*.

Ma, P. (2003), Nonparametric mixed-effects models, Ph.D. thesis, Purdue University, West Lafayette, IN.

Ma, P., P. Wang, M. V. de Hoop, L. Tenorio, and R. D. van der Hilst (2006), Imaging of structure at and near the core mantle boundary using a generalized Radon transform: II-Statistical inference of singularities, *J. Geophys. Res.*, p. in press.

Margerin, L., and G. Nolet (2003), Multiple scattering of high-frequency seismic waves in the deep Earth: *PKP* precursor analysis and inversion for mantle granularity, *J. Geophys. Res.*, *108 (B11), Art. No. 2514*.

Merkel, S., A. Kubo, L. Miyagi, S. Speziale, T. S. Duffy, H.-K. Mao, and H.-R. Wenk (2006), Plastic deformation of $MgGeO_3$ post-perovskite at lower mantle pressures, *Science*, *311*, 644–646.

Miller, D., M. Oristaglio, and G. Beylkin (1987), A new slant on seismic imaging: Migration and integral geometry, *Geophysics*, *52*, 943–964.

Morelli, A., and A. M. Dziewonski (1987), *Seismic tomography*, D. Reidel, Norwell, Mass.

Murakami, M., K. Hirose, K. Kawamura, N. Sata, and O. Y. (2004), Post-perovskite phase transition in $MgSiO_3$, *Science*, *304*, 855–858.

Nimmo, F., G. D. Price, J. Brodholt, and D. Gubbins (2004), The influence of potassium on core and geodynamo evolution, *Geophys. J. Int.*, *156*, 363–376, doi: 10.1111/j.1365–246X.2003.02,157.x.

Norton, S. G., and M. Linzer (1981), Ultrasonic scattering potential imaging in three dimensions: Exact inverse scattering solutions for plane, cylindrical, and spherical apertures, *IEEE Trans. on Biomedical Engineering*, *BME-28*, 202–220.

Nychka, D. (1988), Bayesian confidence intervals for smoothing splines, *J. Amer. Statist. Assoc.*, *83*, 1134–1143.

Oganov, A. R., and S. Ono (2004), Theoretical and experimental evidence for a post-perovskite phase of $MgSiO_3$ in Earth's D'' layer, *Nature*, *430*, 445–448.

Paulssen, H. (1988), Evidence for a sharp 670-km discontinuity as inferred from *P*-to-*S* converted waves, *J. Geophys. Res.*, *93*, 10,489–10,500.

Pinheiro, J. C., and D. M. Bates (2000), *Mixed-Effects Models in S and S-PLUS*, Springer-Verlag, New York.

Poppeliers, C., and G. L. Pavlis (2003), Three-dimensional, prestack, planewave migration of teleseismic *P*-to-*S* converted phases I: Theory, *J. Geophys. Res.*, *108*, 2112.

Quinn, B. G., and J. M. Fernandes (1991), A fast efficient technique for the estimation of frequency, *Biometrika*, *78*, 489–497.

Reasoner, C., and J. Revenaugh (1999), Short-period *P* wave constrains on D'' reflectivity, *J. Geophys. Res.*, *104*, 955–961.

Revenaugh, J. (1995), A scattered-wave image of subduction beneath the transverse ranges, *Science*, *268*, 1888–1892.

Revenaugh, J., and T. H. Jordan (1991), Mantle layering from *ScS* reverberations .2. the transition zone, *J. Geophys. Res.*, *96*, 19,763–19,780.

Ribe, N., E. Stutzmann, Y. Ren, and R. D. van der Hilst (2007), Buckling instabilities of subducted lithosphere beneath the transition zone, *Earth Planet. Sci. Lett.*, *254*, 173–179.

Robinson, G. K. (1991), That BLUP is a good thing: The estimation of the random effects, *Statist. Sci.*, *6*, 15–51 (with discussions).

Romanowicz, B. (2003), Global mantle tomography: Progress status in the past 10 years, *Ann. Rev. Earth Planet. Sci.*, *31*, 303–328.

Rondenay, S., and K. M. Fischer (2003), Constraints on localized CMB structure from multichannel, broadband *SKS*-coda analysis, *J. Geophys. Res.*, *108*, doi:10.1029/2003JB002,518.

Rost, C., and C. Thomas (2002), Array seismology: Methods and applications, *Rev. Geophys.*, *40*, in print.

Rost, S., and J. Revenaugh (2004), Small-scale changes of core-mantle boundary reflectivity studied using core reflected *PcP*, *Phys. Earth Planet. Inter.*, *145*, 19–36.

Rydberg, T., and M. Weber (2000), Receiver function array: A reflection seismology approach, *Geophys. J. Int.*, *141*, 1–11.

Scherbaum, F., F. Krüger, and M. Weber (1997), Double beam imaging: Mapping lower mantle heterogeneities using combinations of source and receiver arrays, *J. Geophys. Res.*, *102 (B1)*, 507–522.

Schimmel, M., and H. Paulssen (1997), Noise reduction and detection of weak, coherent signals through phase-weighted stacks, *Geophys. J. Int.*, *130*, 497–505.

Schleicher, J., M. Tygel, and P. Hubral (1993), 3-D true-amplitude finite-offset migration, *Geophysics*, *58*, 1112–1126.

Schneider, W. A. (1978), Integral formulation for migration in two and three dimensions, *Geophysics*, *43*, 49–76.

Shearer, P. M., and M. P. Flanagan (1999), Seismic velocity and density jumps across the 410- and 660-kilometer discontinuities, *Science*, *285*, 1545–1548.

Shearer, P. M., M. P. Flanagan, and M. A. H. Hedlin (1999), Experiments in migration processing of *SS* precursor data to image upper mantle discontinuity structure, *J. Geophys. Res.*, *104*, 7229–7242.

Shim, S.-H. (2005), Stability of $MgSiO_3$ perovskite in the lower mantle, *Geophys. Monograph., Am. Geophys. Un.*, *160*, 261–282.

Shim, S.-H., T. S. Duffy, R. Jeanloz, and G. Shen (2004), Stability and chrystal structure of $MgSiO_3$ perovskite to the core-mantle boundary, *Geophys. Res. Lett.*, p. doi:10.1029/2004GL019639.

Sidorin, I., M. Gurnis, and D. V. Helmberger (1999), Evidence for a ubiquitous seismic discontinuity at the base of the mantle, *Science*, *286*, 1326–1331.

Simon, M., H. Gebrande, and M. Bopp (1996), Pre-stack migration and true-amplitude processing of DEKORP near-normal incidence and wide-angle reflection measurements, *Tectonophysics*, *264*, 381–393.

Stacey, F. D. (1992), *Physics of the Earth*, 3rd ed., Brookfield, Kenmore, Brisbane 4069, Australia.

Stolk, C. C., and M. V. de Hoop (2002), Microlocal analysis of seismic inverse scattering in anisotropic, elastic media, *Comm. Pure Appl. Math.*, *55*, 261–301.

Stolk, C. C., and M. V. de Hoop (2004), Seismic inverse scattering in the downward continuation approach, *SIAM J. Appl. Math.*, submitted.

Stolt, R. H., and A. B. Weglein (1985), Migration and inversion of seismic data, *Geophysics*, *50*, 2456–2472.

Sun, D., T.-R. Song, and D. V. Helmberger (2006), Complexity of D'' in the presence of slab-debris and phase changes, *Geophys. Res. Lett.*, *33*, *L12S07*, doi:10.1029/2005GL025,384.

Sze, E., and R. D. van der Hilst (2003), Core mantle boundary topography from short period *PcP*, *PKP*, and *PKKP* data, *Phys. Earth Planet. Inter.*, *135*, 27–46.

Tan, E., M. Gurnis, and L. Han (2002), Slabs in the lower mantle and their modulation of plume formation, *Geochem. Geophys. Geosyst.*, *3*, 1067, doi:10.1029/2001GC000,238.

Thomas, C., E. J. Garnero, and T. Lay (2004), High-resolution imaging of lowermost mantle structure under the Cocos plate, *J. Geophys. Res.*, *109*, B08,307.

Tromp, J., and A. M. Dziewonski (1998), Geoscience – Two views of the deep mantle, *Science*, *281*, 655–656.

Tsuchiya, T., J. Tsuchiya, K. Umemoto, and R. M. Wentzcovitch (2004), Phase transition in $MgSiO_3$-perovskite in the Earth's lower mantle, *Earth Planet. Sci. Lett.*, *224*, 241.

Ulrych, T. J., M. D. Sacchi, and S. L. M. Freire (1998), Eigenimage processing of seismic sections: In *Covariance Analysis of Seismic Signal Processing*, edited by R.L. Kirlin, and W.J. Done, *SEG Monograph*, p. Tulsa.

Van den Berg, A. P., E. S. G. Rainey, and D. A. Yuen (2005), The combined influences of variable thermal conductivity, temperature- and pressure-dependent viscosity and core-mantle coupling on thermal evolution, *Phys. Earth Planet. Inter.*, *149*, 259–278.

Van der Hilst, R. D., and H. Kárason (1999), Compositional heterogeneity in the bottom 1000 km of Earth's mantle: Towards a hybrid convection model, *Science*, *283*, 1885–1888.

Van der Hilst, R. D., S. Widyantoro, and E. R. Engdahl (1997), Evidence for deep mantle circulation from global tomography, *Nature*, *386*, 578–584.

Van der Hilst, R. D., M. V. de Hoop, P. Wang, S.-H. Shim, P. Ma, and L. Tenorio (2007), Seismo-stratigraphy and thermal structure of Earth's core-mantle boundary region, *Science*, *315*, 1813–1817.

Van der Lee, S., H. Paulssen, and G. Nolet (1996), Variability of $P_{660}S$ phases as a consequence of topography of the 660 km discontinuity, *Phys. Earth Planet. Inter.*, *86*, 147–164.

VanDecar, J. C., and R. S. Crosson (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seismol. Soc. Am.*, *80*, 150–159.

Vinnik, L., M. Kato, and H. Kawakatsu (2001), Search for seismic discontinuities in the lower mantle, *Geophys. J. Int.*, *147*, 41–56.

Vonesh, E. F., and V. M. Chinchilli (1997), *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, Marcel Dekker, New York.

Wahba, G. (1983), Bayesian "confidence intervals" for the cross-validated smoothing spline, *J. Roy. Statist. Soc. Ser. B*, *45*, 133–150.

Wahba, G. (1990), *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 59, SIAM, Philadelphia.

Wang, P., M. V. de Hoop, R. D. van der Hilst, P. Ma, and L. Tenorio (2006), Imaging of structure at and near the core mantle boundary using a generalized Radon transform: I- construction of image gathers, *J. Geophys. Res.*, *111*, B1230, doi:10.1029/2005JB004,241.

Wang, Y. (1998), Mixed-effects smoothing spline ANOVA, *J. Roy. Statist. Soc. Ser. B*, *60*, 159–174.

Wentzcovitch, R. M., T. Tsuchiya, and J. Tsuchiya (2006), $MgSiO_3$ post-perovskite at D'' conditions, *PNAS*, *103*, 543–546.

Williams, Q., R. Jeanloz, J. Bass, B. Svendsen, and T. J. Ahrens (1987), The melting curve of iron to 250 GPa: A constraint on the temperature at Earth's center, *Science*, *236*, 181–182.

Wysession, M. E., T. Lay, J. Revenaugh, Q. Williams, E. J. Garnero, R. Jeanloz, and L. H. Kellogg (1998), The D'' discontinuity and its implications, *In: "The core mantle boundary region, geodynamics series, Gurnis et al. (Editors)"*, *28*, 273–298.