Inverse Scattering: Theory and Application to the Imaging of the Earth's Seismic Discontinuities

by

Xuefeng Shang

M.S. Geophysics, Peking University, Beijing China, 2007
B.S. Geophysics, Peking University, Beijing China, 2004

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

February 2014

Author ....
Department of Earth, Atmospheric, and Planetary Sciences
November 8, 2013

Certified by ....
Robert D. van der Hilst
Schlumberger Professor of Earth Sciences
Thesis Supervisor

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

Accepted by . ...
Robert D. van der Hilst
Schlumberger Professor of Earth Sciences
Head, Department of Earth, Atmospheric and Planetary Sciences

Inverse Scattering: Theory and Application to the Imaging of the Earth's Seismic Discontinuities

by

Xuefeng Shang

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
on Nov. 8, 2013 in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Geophysics

Abstract

In this thesis we further develop concepts in inverse scattering, which enable higher resolution imaging with data from dense networks and arrays. We apply the new methods to studies of the crust beneath south Tibet and the core-mantle boundary (CMB) region beneath Central America and East Asia.

First, we develop a new method, which we call passive source reverse time migration (RTM), for subsurface imaging with teleseismic array data. Multi-component array data are first propagated backward by solving the elastic wave equation. After polarization separation, a modified cross-correlation imaging condition between *P* and *S* wave constitutes is applied to obtain an inverse scattering transform. From synthetic experiments it is evident that for complex geological structures RTM is superior to traditional receiver functions analysis, such as common conversion point (CCP) stacking.

Two preprocessing steps are required for RTM application on sparsely sampled teleseismic dataset: source normalization and trace interpolation. The source radiation pattern, especially the polarity of traces, is corrected by multi-channel cross-correlation technique. The unknown source signature is then estimated by principle component analysis and deconvolved from raw data by Wiener deconvolution. Curvelet interpolation with sparsity promotion is employed to interpolate irregularly and sparsely sampled traces into regular and dense grids. Synthetic and real data examples demonstrate that for typical teleseismic acquisition geometry, with 50% to 85% missing traces, the curvelet-based interpolation works remarkably well. The application on Hi-CLIMB array data in Tibetan plateau reveals clear and continuous Moho discontinuity at the depth of about 70 km, as well as fine crustal structures.

Second, we use a high-frequency approximation of inverse scattering, generalized Radon transform (GRT), to probe the lowermost mantle beneath Central America and East Asia. Inverse scattering of about 130,000 *ScS* traces and 120,000 *SKKS* traces reveals multiple reflectors above the conventional D" region. This result is inconsistent with expectations from a pure thermal response of a single isochemical post-perovskite transition but can be explained with post-perovskite transitions in differentiated slab materials. Our results imply that the lowermost mantle is more complex than hitherto thought and that the presence of interfaces and compositional heterogeneity beyond the D" region.

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

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

Acknowledgements

How time flies! The journey of my ph.D life at MIT is finally coming to the end. I would take this opportunity to thank my professors, friends and family for their enormous help and encouragement. Without their support, this thesis would be a mission impossible.

I am grateful to my advisor Prof. Rob van der Hilst, whose enthusiasm and vision about seismology always inspires me. If I have seen further it is by standing on the shoulders of a giant, figuratively and literally. I would also like to thank my co-advisor Prof. Maarten de Hoop for his patient guidance and sharp advice. He is a never outdated Wikipedia of inverse scattering and imaging. I really learn a lot from him.

Secondly, I am thankful to my thesis committee members, Prof. Alison Malcolm, Germán Prieto and Tom Herring for their valuable suggestion and comments on my thesis. In particular I would like to thank Prof. Dan Shim, who was my second project advisor, for educating me mineral physics and discussing the implication of our seismic images in the lowermost mantle.

I would appreciate tremendous help from my dear friends at MIT and Boston, and I will miss the happy days with "Chinese Mafia" in ERL. In particular I would like to thank Ms. Sue Turbak, Ms. Carol Sprague, Ms. Vicki Mckenna, Ms. Jacqueline Taylor and Ms. Linda Meinke for their consideration and remarkable support.

My deepest gratitude goes to the most two important women in my life, my mother and my wife Emily. They are always supportive. To them I dedicate my dissertation.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Inverse Scattering

In the past decades our knowledge of the Earth's interior has evolved from a radial stratified model, for instance, PREM [*Dziewoński and Anderson*, 1981] or ak135 [*Kennett et al.*, 1995], to a complex 3D view exhibiting strong lateral heterogeneities on all (resolved) length scales. The relatively smooth deviations from a 1D radially symmetric Earth can be delineated by body wave [e.g. *van der Hilst et al.*, 1997; *Grand*, 2002; *Li et al.*, 2008] or surface wave transmission tomography [e.g. *Trampert and Woodhouse*, 1995; *Boschi and Ekström*, 2002; *Yao et al.*, 2006; *Lebedev and van der Hilst*, 2008]. Seismic tomography attributes travel time anomalies to volumetric heterogeneities and is relatively insensitive to seismic sharp discontinuities due to, for instance, compositional or mineral phase changes. On the other hand, reflected [e.g. *Shearer and Flanagan*, 1999; *van der Hilst et al.*, 2007; *Cao et al.*, 2011] or scattered waves, including wave mode conversion (*P*-to-*S* or *S*-to-*P* conversion) [e.g. *Langston*, 1977; *Dueker and Sheehan*, 1997; *Bostock et al.*, 2001; *Kind et al.*, 2002; *Gilbert et al.*, 2003], can recover the topography of sharp boundaries, such as Moho, 410', 660' and D" discontinuities.

The seismic methods used so far to probe Earth's discontinuities fall into three main categories: forward modeling, 1D stacking, and inverse scattering (migration). The trial and error based forward modeling has provided evidence for strong and wide-spectrum heterogeneities in the Earth [e.g. *Lay and Helmberger*, 1983; *Zhu and Helmberger*, 1998; *Garnero*, 2000]. However, waveform modeling requires sufficient energy for the signal of interest to be clearly visible in individual traces. This requires large earthquakes and, often, analysis of data near the critical angle. In turn, this implies that only very limited regions can be studied [*Wang et al.*, 2006]. Stacking the traces over a Fresnel zone can enhance the signal-to-noise ratio, and is routinely applied in teleseismic imaging, for instance *SS* in the transition zone [e.g. *Shearer*, 1993; *Deuss and Woodhouse*, 2001], *ScS* in the lowermost mantle [e.g. *Thomas et al.*, 2004a; *Hutko et al.*, 2006], and receiver functions in crust and upper mantle [e.g. *Dueker and Sheehan*, 1997; *Zhu*, 2000]. An important drawback of stacking is the 1D model assumption, which is only valid when the discontinuity is laterally continuous and the topographic variation is small [*Zheng*, 2007; *Shang et al.*, 2012]. Compared with the two methods above, inverse scattering relies on fewer prior assumptions about the Earth's structures and can provide finer resolution images with less artifacts. It has become popular recently with available dense seismographic arrays [e.g. *Poppeliers and Pavlis*, 2003; *Chen et al.*, 2005; *Wang et al.*, 2006; *Cao et al.*, 2010; *Shang et al.*, 2012].

There are different implementations of inverse scattering. We consider the category of approaches known under the collective names of Kirchhoff migration or generalized Radon transform (GRT) [see *Beylkin*, 1985; *Miller et al.*, 1987; *Bleistein et al.*, 2001] and reverse time migration (RTM) [*Whitmore*, 1983; *McMechan*, 1983; *Baysal et al.*, 1983; *Op't Root et al.*, 2012; *Brytik et al.*, 2012]. GRT is based on the high frequency approximation and is applicable in sufficiently smooth background models. On the other hand, RTM solves the wave equation and leads to robust solutions, which admit background models of limited regularity. From data acquisition aspect, GRT is relatively flexible for source-receiver configuration whereas RTM has more stringent sampling requirement. One can choose different approaches according to the imaging targets and acquisition geometry. For instance, GRT might be adequate for deep mantle imaging due to the smoothness of available background models and sparse and irregular samplings over the globe. With dense seismic array data, RTM can be applied for high-resolution seismic images of the crust and upper mantle.

1.2 Thesis Objectives and Motivation

Increasingly dense seismographic arrays are being deployed all over the world to constrain better subsurface structures and geological processes. For conventional methods of detecting seismic discontinuities, e.g. receiver functions and common-bounce-point stacking, the gain from reducing station interval is diminishing due to the simplifying but often limiting assumptions about the Earth's structures (1D velocity model assumption, for instance) [e.g. *Rondenay*, 2009; *Shang et al.*, 2012]. In this thesis we further develop the concepts of inverse scattering for utilizing the full complexity of the wavefields recorded by dense arrays and networks. We propose passive source elastic reverse time migration, a new type of wave equation migration, for subsurface imaging with multi-component teleseismic data. Following the pioneer work of *Wang et al.* [2006] and *Wang et al.* [2008], ray theory based GRT is further refined for deep mantle probing. An efficient and robust data preprocessing workflow, including source normalization and trace interpolation, is necessary for inverse scattering applications, and is presented in this thesis as well. As real data examples, we apply passive source RTM to the images of south Tibetan lithosphere, and employ GRT to probe the lowermost mantle beneath central America and east Asia.

The most prominent feature in the lithosphere is the Mohorovičić discontinuity (Moho), which marks the boundary between the Earth's crust and mantle. Though a precise definition of Moho is absent, it is a distinct manifestation of a differentiated Earth, as physical and chemical properties change significantly across it, such as seismic wave velocity, density, rheology petrology, and mineralogy. Using Moho as a structural marker, we can infer valuable information about tectonic processes, e.g. how the oceanic slab subducts into the trench, how the lithosphere deforms during continental collision, and to which level the crust is compensated isostatically. For detailed reviews on the nature of

Moho we refer to *Jarchow and Thompson* [1989] and *Eaton* [2006] and references therein.

There are two popular seismic methods to determine the location of Moho discontinuity. One is the controlled source method. The sources explode on the surface and the reflection/refraction data are collected to constrain the Moho depth and crust velocity. The accuracy highly depends on the identification of *Pg* and *Pn* phases [e.g. *Prodehl and Mooney*, 2012]. Another method is the passive source method, such as common conversion point stacking (CCP) of receiver functions, which exploits wave mode conversions between *P* and *S* at discontinuities [e.g. *Langston*, 1977; *Ammon*, 1991; *Dueker and Sheehan*, 1997; *Gilbert et al.*, 2003]. This method fails in presence of small-scale structures (comparable with wavelength) due to the 1D model assumption. Under complex geological circumstances inverse scattering, such as passive source RTM, can improve the image quality with less artificial structures.

Since the collision between the Indian and Eurasian plates started ~50 million years ago, the Tibet plateau has been uplifted more than 5 km and the crust is twice as thick as average continental crust (~70 km) [e.g. *Yin and Harrison*, 2000; *Royden et al.*, 2008]. Evident deformation occurred in the crust and lithosphere during the continental collision. Rich features of Moho in Tibet are reported, such as abrupt steps (up to 20 km vertical offset) [e.g. *Hirn et al.*, 1984; *Zhu and Helmberger*, 1998; *Wittlinger et al.*, 2004], double Moho [e.g. *Kind et al.*, 2002; *Wittlinger et al.*, 2004; *Nábělek et al.*, 2009; *Wittlinger et al.*, 2009], and disrupted Moho [e.g. *Nowack et al.*, 2010]. In this thesis we use passive source RTM to delineate the Moho discontinuity in south Tibet with Hi-CLIMB array data.

The structure atop the core-mantle boundary (CMB) is as enigmatic as that of the Earth's lithosphere. One prominent feature in the lowermost mantle from global tomography is a belt of high seismic velocity along the circum-Pacific and two large low velocity provinces (LLVPs) beneath the middle Pacific and west Africa [e.g. *Li and Romanowicz*, 1996; *Grand*, 2002; *Ishii and Tromp*, 2004; *Li et al.*, 2008]. Due to the good spatial correlation with subduction zones and hot spots, a popular hypothesis is that the high velocity anomaly may be related to subducted slab dregs [e.g. *Richards and Engebretson*, 1992; *Christensen and Hofmann*, 1994; *Grand et al.*, 1997; *van der Hilst et al.*, 1997], while the low velocities are often interpreted as dense thermochemical piles or super mantle plumes [e.g. *Kellogg et al.*, 1999; *Ni et al.*, 2002; *Jellinek and Manga*, 2004; *McNamara and Zhong*, 2005]. However, the thermal and chemical attributes of such heterogeneities in the lowermost mantle are still not well constrained.

Since a mantle discontinuity is firstly documented at the top of D" region [*Wright and Lyons*, 1980; *Lay and Helmberger*, 1983], the D" discontinuity (~100 to 450 km above CMB) is reported at dozens of locations all over the world [for a review, see *Wysession et al.*, 1998]. Global existence of the D" discontinuity has, however, not been established. It is partially due to the source-receiver geometry restriction imposed by the prevailing method, trial-and-error based waveform modeling, and only fairly limited D" regions have been studied [*Wang et al.*, 2006]. On the other hand, with inverse scattering, such as

GRT, it is plausible to image the fine structures on a large (potentially continental) scale [e.g. *Wang et al.*, 2006; *van der Hilst et al.*, 2007; *Wang et al.*, 2008]. The nature of the D" discontinuity is still under debate. It could represent a thermal, mineralogical or chemical boundary [e.g. *Wysession et al.*, 1998; *Sidorin et al.*, 1999; *Sun et al.*, 2006]. Recent studies reveal more than one seismic discontinuity within this mysterious lowermost mantle, such as in the large low-shear-velocity province (LLSVP) beneath central Pacific ocean [*Lay et al.*, 2006], north Asia [*Gaherty and Lay*, 1992; *Thomas et al.*, 2004b], and central America [*Thomas et al.*, 2004a; *van der Hilst et al.*, 2007; *Hutko et al.*, 2008].

The discovery of $MgSiO_3$ perovskite (Pv) to post-perovskite (pPv) phase transition offers new opportunities for understanding the structures near CMB [*Murakami et al.*, 2004; *Oganov and Ono*, 2004; *Shim et al.*, 2004]. Due to a large Clapeyron slope for Pv-pPv phase transition and large thermal gradients expected in the D" region (presumably a thermal boundary layer), the Pv-pPv phase boundary could be crossed twice by the geotherm, as pPv might transfer back to Pv in the deeper location [e.g. *Hernlund et al.*, 2005; *Lay et al.*, 2006; *van der Hilst et al.*, 2007]. The existence of such a post-perovskite lens has been proposed to explain the double layers in D" region, which if correct, enables the estimation of the thermal structure at the lowermost mantle [see *Hernlund et al.*, 2005; *Lay et al.*, 2006; *van der Hilst et al.*, 2007]. Recent high pressure-temperature studies, however, suggest that pPv transition is sensitive to the mantle chemical composition [e.g. *Catalli et al.*, 2009; *Andrault et al.*, 2010; *Grocholski et al.*, 2012]. The pPv transition pressure for standard bulk composition, such as pyrolite, may be too high to occur in the lower mantle. Candidate compositions for a seismically detectable pPv transition at pressures less than the CMB include mid-oceanic ridge basalt (MORB) and harzburgite components of differentiated oceanic lithosphere transported to the lowermost mantle by subduction. Motivated by the hypothesis above, we conduct high-resolution and large-scale seismic images of the lowermost mantle along the circum-Pacific region, east Asia and central America.

1.3 Thesis Outline

In **Chapter II** we present passive source reverse time migration, a new type of wave equation prestack depth migration, to image crust and mantle structures. Theory and synthetic tests are carried out to demonstrate the viability and promise of the new method. In **Chapter III** we propose a semi-automatic preprocessing workflow suitable for massive volume of teleseismic data. Synthetic and real data show the efficiency and robustness of the workflow. This workflow is used in the real data applications of Chapter V and VI. In **Chapter IV** we investigate the concept of wavefield interpolation, which is an important preprocessing step for reverse time migration. We employ the curvelet transform and sparsity promoting inversion to reconstruct the wavefield from sparsely sampled dataset. In **Chapter V** passive source RTM is adopted to the imaging of the lithosphere beneath south Tibet with Hi-Climb array data. The new method reveals complex and rich structures in this prominent continental collision region. In **Chapter VI** we apply ray theory based GRT to image the lowermost mantle beneath central America and east Asia. The seismic images are interpreted in terms of mineral physics and geodynamics. In **Chapter VII** we summarize the key results of our studies and discuss possible directions for the future work.

References

Ammon, C. J. (1991), The isolation of receiver effects from teleseismic P waveforms, *Bull. Seism. Soc. Am.*, *81*(6), 2504-2510.

Andrault, D., M. Munoz, N. Bolfan-Casanova, N. Guignot, J.-P. Perrillat, G. Aquilanti, and S. Pascarelli (2010), Experimental evidence for perovskite and post-perovskite coexistence throughout the whole D" region, *Earth Planet. Sci. Lett.*, *293*(1), 90-96.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood (1983), Reverse time migration, *Geophysics*, *48*(11), 1514-1524.

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

Bleistein, N., J. K. Cohen, and J. W. Stockwell (2001), *Mathematics of multidimensional seismic imaging, migration, and inversion*, Springer Verlag.

Boschi, L., and G. Ekström (2002), New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. Geophys. Res.*, *107*(B4), 2059.

Bostock, M., S. Rondenay, and J. Shragge (2001), Multiparameter two-dimensional inversion of scattered teleseismic body waves. I. Theory for oblique incidence, *J. Geophys. Res.*, *106*, 771-782.

Brytik, V., M. V. de Hoop, and R. D. van der Hilst (2012), Elastic-wave inverse scattering based on reverse time migration with active and passive source reflection data, in *Inside Out - Inverse Problems, MSRI Publications*, edited by G. Uhlmann, pp. 411-453, Cambridge University Press.

Cao, Q., P. Wang, R. van der Hilst, M. de Hoop, and S.-H. Shim (2010), Imaging the upper mantle transition zone with a generalized Radon transform of SS precursors, *Phys. Earth Planet. Inter.*, *180*(1), 80-91.

Cao, Q., R. van der Hilst, M. de Hoop, and S.-H. Shim (2011), Seismic imaging of transition zone discontinuities suggests hot mantle west of Hawaii, *Science*, *332*(6033), 1068-1071.

Catalli, K., S. H. Shim, and V. Prakapenka (2009), Thickness and Clapeyron slope of the post-perovskite boundary, *Nature*, *462*(7274), 782-U101.

Chen, L., L. Wen, and T. Zheng (2005), A wave equation migration method for receiver function imaging: 1. Theory, *J. Geophys. Res.*, *110*(B11), B11309.

Christensen, U. R., and A. W. Hofmann (1994), Segregation of subducted oceanic crust in the convecting mantle, *J. Geophys. Res.*, *99*(B10), 19867-19884.

Deuss, A., and J. Woodhouse (2001), Seismic observations of splitting of the mid-transition zone discontinuity in Earth's mantle, *Science*, *294*(5541), 354-357.

Dueker, K. G., and A. F. Sheehan (1997), Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track, *J. Geophys. Res.*, *102*, 8313-8327.

Dziewoński, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*(4), 297-356.

Eaton, D. W. (2006), Multi-genetic origin of the continental Moho: insights from Lithoprobe, *Terra Nova*, *18*(1), 34-43.

Gaherty, J. B., and T. Lay (1992), Investigation of laterally heterogeneous shear velocity structure in D" beneath Eurasia, *J. Geophys. Res.*, *97*(B1), 417-435.

Garnero, E. (2000), Lower mantle heterogeneity, *Annu. Rev. Earth Planet. Sci*, *28*, 509-537.

Gilbert, H. J., A. F. Sheehan, K. G. Dueker, and P. Molnar (2003), Receiver functions in the western United States, with implications for upper mantle structure and dynamics, *J. Geophys. Res*, *108*, 2229.

Grand, S. P., R. D. van der Hilst, and S. Widiyantoro (1997), High resolution global tomography: a snapshot of convection in the Earth, *GSA Today*, *7*(4).

Grand, S. P. (2002), Mantle shear-wave tomography and the fate of subducted slabs, *Phil. Trans. R. Soc. A*, *360*(1800), 2475-2491.

Grocholski, B., K. Catalli, S.-H. Shim, and V. Prakapenka (2012), Mineralogical effects on the detectability of the postperovskite boundary, *Proc. Natl. Acad. Sci.* , *109*(7), 2275-2279.

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*(7035), 882-886.

Hirn, A., A. Nercessian, M. Sapin, G. Jobert, X. Z. Xin, G. E. Yuan, L. D. Yuan, and T. J. Wen (1984), Lhasa block and bordering sutures-a continuation of a 500-km Moho traverse through Tibet, *Nature*, *307*, 25-27.

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*(7091), 333-336.

Hutko, A. R., T. Lay, J. Revenaugh, and E. J. Garnero (2008), Anticorrelated seismic velocity anomalies from post-perovskite in the lowermost mantle, *Science*, *320*(5879), 1070.

Ishii, M., and J. Tromp (2004), Constraining large-scale mantle heterogeneity using mantle and inner-core sensitive normal modes, *Phys. Earth Planet. Inter.*, *146*(1), 113-124.

Jarchow, C. M., and G. A. Thompson (1989), The nature of the Mohorovicic discontinuity, *Annu. Rev. Earth Planet. Sci.* , *17*, 475.

Jellinek, A. M., and M. Manga (2004), Links between long-lived hot spots, mantle plumes, D", and plate tectonics, *Rev. Geophys.*, *42*(3).

Kellogg, L. H., B. H. Hager, and R. D. van der Hilst (1999), Compositional stratification in the deep mantle, *Science*, *283*(5409), 1881-1884.

Kennett, B., E. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Inter.*, *122*(1), 108-124.

Kind, R., X. Yuan, J. Saul, D. Nelson, S. Sobolev, J. Mechie, W. Zhao, G. Kosarev, J. Ni, and U. Achauer (2002), Seismic images of crust and upper mantle beneath Tibet: Evidence for Eurasian plate subduction, *Science*, *298*(5596), 1219-1221.

Langston, C. A. (1977), Corvallis, Oregon, crustal and upper mantle receiver structure from teleseismic P and S waves, *Bull. Seism. Soc. Am.*, *67*(3), 713-724.

Lay, T., and D. Helmberger (1983), A lower mantle S-wave triplication and the shear velocity structure of D", *Geophys. J. R. Astron. Soc*, *75*, 799-838.

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*(5803), 1272-1276.

Lebedev, S., and R. D. van der Hilst (2008), Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms, *Geophys. J. Inter.*, *173*(2), 505-518.

Li, C., R. D. van der Hilst, E. R. Engdahl, and S. Burdick (2008), A new global model for P wave speed variations in Earth's mantle, *Geochem. Geophy. Geosyst.*, *9*(5), Q05018.

Li, X.-D., and B. Romanowicz (1996), Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. Geophy. Res.*, *101*(B10), 22245-22272.

McMechan, G. A. (1983), Migration by Extrapolation of Time-Dependent Boundary Values, *Geophys. Prospect.*, *31*(3), 413-420.

McNamara, A. K., and S. Zhong (2005), Thermochemical structures beneath Africa and the Pacific Ocean, *Nature*, *437*(7062), 1136-1139.

Miller, D., M. Oristaglio, and G. Beylkin (1987), A new slant on seismic imaging: Migration and integral geometry, *Geophysics*, *52*(7), 943-964.

Murakami, M., K. Hirose, K. Kawamura, N. Sata, and Y. Ohishi (2004), Post-perovskite phase transition in $MgSiO_3$, *Science*, *304*(5672), 855-858.

Nábělek, J., G. Hetényi, J. Vergne, S. Sapkota, B. Kafle, M. Jiang, H. Su, J. Chen, and B. S. Huang (2009), Underplating in the Himalaya-Tibet collision zone revealed by the Hi-CLIMB experiment, *Science*, *325*(5946), 1371-1374.

Ni, S., E. Tan, M. Gurnis, and D. Helmberger (2002), Sharp sides to the African superplume, *Science*, *296*(5574), 1850-1852.

Nowack, R. L., W. P. Chen, and T. L. Tseng (2010), Application of gaussian-beam migration to multiscale imaging of the lithosphere beneath the Hi-CLIMB array in Tibet, *Bull. Seism. Soc. Am.*, *100*(4), 1743-1754.

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*(6998), 445-448.

Op't Root, T., C. C. Stolk, and M. V. de Hoop (2012), Linearized inverse scattering based on seismic reverse time migration, *J. Math. Pures Appl.*, *98*(2), 211-238.

Poppeliers, C., and G. L. Pavlis (2003), Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 1. Theory, *J. Geophys. Res.*, *108*(B2), 2112.

Prodehl, C., and W. D. Mooney (2012), *Exploring the Earth's Crust: History and Results of Controlled-Source Seismology*, Geological Society of America.

Richards, M. A., and D. C. Engebretson (1992), Large-scale mantle convection and the history of subduction, *Nature*, *355*(6359), 437-440.

Rondenay, S. (2009), Upper mantle imaging with array recordings of converted and scattered teleseismic waves, *Surv. Geophys.*, *30*(4), 377-405.

Royden, L. H., B. C. Burchfiel, and R. D. van der Hilst (2008), The geological evolution of the Tibetan Plateau, *Science*, *321*(5892), 1054-1058.

Shang, X., M. V. de Hoop, and R. D. van der Hilst (2012), Beyond receiver functions: Passive source reverse time migration and inverse scattering of converted waves, *Geophys. Res. Lett.*, *39*, L15308.

Shearer, P. M. (1993), Global mapping of upper mantle reflectors from long-period SS precursors, *Geophys. J. Inter.*, *115*(3), 878-904.

Shearer, P. M., and M. P. Flanagan (1999), Seismic velocity and density jumps across the 410-and 660-kilometer discontinuities, *Science*, *285*(5433), 1545-1548.

Shim, S., T. Duffy, R. Jeanloz, and G. Shen (2004), Stability and crystal structure of $MgSiO_3$ perovskite to the core-mantle boundary, *Geophys. Res. Lett*, *31*, 1029-1039.

Sidorin, I., M. Gurnis, and D. V. Helmberger (1999), Evidence for a ubiquitous seismic discontinuity at the base of the mantle, *Science*, *286*(5443), 1326-1331.

Sun, D., T. R. A. Song, and D. Helmberger (2006), Complexity of D" in the presence of slab-debris and phase changes, *Geophys. Res. Lett.*, *33*(12), L12S07.

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

Thomas, C., J. M. Kendall, and J. Lowman (2004b), Lower-mantle seismic discontinuities and the thermal morphology of subducted slabs, *Earth Planet. Sci. Lett.*, *225*(1-2), 105-113.

Trampert, J., and J. H. Woodhouse (1995), Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Inter.*, *122*(2), 675-690.

van der Hilst, R., S. Widiyantoro, and E. Engdahl (1997), Evidence for deep mantle circulation from global tomography, *Nature*, *386*(6625), 578-584.

van der Hilst, R., M. de Hoop, P. Wang, S. H. Shim, P. Ma, and L. Tenorio (2007), Seismostratigraphy and thermal structure of Earth's core-mantle boundary region, *Science*, *315*(5820), 1813-1817.

Wang, P., M. de Hoop, R. van der Hilst, P. Ma, and L. Tenorio (2006), Imaging of structure at and near the core mantle boundary using a generalized radon transform: 1. Construction of image gathers, *J. Geophys. Res.*, *111*.

Wang, P., M. V. de Hoop, and R. D. van der Hilst (2008), Imaging the lowermost mantle (D") and the core-mantle boundary with SKKS coda waves, *Geophys. J. Inter.*, *175*(1), 103-115.

Whitmore, N. D. (1983), Iterative depth migration by backward time propagation, *SEG Technical Program Expanded Abstracts*, *2*(1), 382-385.

Wittlinger, G., J. Vergne, P. Tapponnier, V. Farra, G. Poupinet, M. Jiang, H. Su, G. Herquel, and A. Paul (2004), Teleseismic imaging of subducting lithosphere and Moho offsets beneath western Tibet, *Earth Planet. Sci. Lett.*, *221*(1), 117-130.

Wittlinger, G., V. Farra, G. Hetényi, J. Vergne, and J. Nábělek (2009), Seismic velocities in southern Tibet lower crust: A receiver function approach for eclogite detection, *Geophys. J. Inter.*, *177*(3), 1037-1049.

Wright, C., and J. Lyons (1980), Further evidence for radial velocity anomalies in the lower mantle, *Pure Appl. Geophys.*, *119*(1), 137-162.

Wysession, M. E., T. Lay, J. Revenaugh, Q. Williams, E. Garnero, R. Jeanloz, and L. H. Kellogg (1998), The D" discontinuity and its implications, in *The core-mantle boundary region*, edited by M. Gurnis, M. E. Wysession, E. Knittle and B. A. Buffett, pp. 273-298, American Geophysical Union.

Yao, H., R. D. van der Hilst, and M. V. de Hoop (2006), Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis, I. Phase velocity maps, *Geophys. J. Inter.*, *166*(2), 732-744.

Yin, A., and T. M. Harrison (2000), Geologic evolution of the Himalayan-Tibetan orogen, *Annu. Rev. Earth Planet. Sci.* , *28*(1), 211-280.

Zheng, Y. (2007), Imaging upper mantle discontinuities and Earth's small scale heterogeneities, phD thesis, Univerisity of California, Santa Cruz.

Zhu, L., and D. V. Helmberger (1998), Moho offset across the northern margin of the Tibetan Plateau, *Science, 281*(5380), 1170-1172.

Zhu, L. (2000), Crustal structure across the San Andreas Fault, southern California from teleseismic converted waves, *Earth Planet. Sci. Lett., 179*(1), 183-190.

Chapter 2

Passive Source Reverse Time Migration[†]

In this chapter, we present a wave equation prestack depth migration to image crust and mantle structures using multi-component earthquake data recorded at dense seismograph arrays. Transmitted *P* and *S* waves recorded on the surface are back propagated using an elastic wave equation solver. The wave modes are separated after the reverse-time continuation of the wavefield from the surface, and subjected to a (cross-correlation type) imaging condition forming an inverse scattering transform. Reverse time migration (RTM) does not make assumptions about the presence or properties of interfaces – notably, it does not assume that interfaces are (locally) horizontal. With synthetic experiments, and different background models, we show that passive source RTM can reconstruct dipping and vertically off-set interfaces even in the presence of complex wave phenomena (such as caustics and point diffraction) and that its performance is superior to

---

[†] Published as: Shang, X., M. V. de Hoop and R. D. van der Hilst (2012), Beyond receiver functions: Passive source reverse time migration and inverse scattering of converted waves, *Geophys. Res. Lett.*, 39, L15308.

traditional receiver function analysis, e.g. common conversion point (CCP) stacking, in complex geological environments.

2.1 Introduction

Increasingly dense seismographic arrays are being deployed all over the world in attempts to constrain subsurface structures and geological processes in greater detail. Reducing the spacing between seismograph stations will, ultimately, have diminishing returns unless we can apply imaging methods that exploit the full complexity of the recorded broadband wavefields and that do not rely on simplifying – but often profoundly limiting – prior assumptions about Earth's structure.

Common conversion point (CCP) stacking techniques are now routinely applied in the receiver function workflow to image interfaces in the crust and mantle beneath the stations [e.g. *Dueker and Sheehan*, 1997; *Gilbert et al.*, 2003; *Zhu*, 2000]. For smoothly varying structures this can produce good results [e.g. *Zhai and Levander*, 2011], but the horizontal interface assumption in CCP stacking prevents the accurate imaging of geologically complex structures, such as dipping and laterally discontinuous interfaces (e.g. strong interface topography, steep faults, steps in Moho). Moreover, stacking data from individual stations cannot adequately suppress scattering or diffraction "artifacts" [e.g. *Chen et al.*, 2005; *Rondenay*, 2009; *Sheehan et al.*, 2000].

With access to data from dense arrays we can avoid stacking in the spatial domain, e.g. over a presumed horizontal interface, and propagate scattered energy back to the subsurface point where scattering occurs. Reverse time migration (RTM) improves image quality remarkably. Migration is based on the single scattering approximation and does not distinguish between refraction and reflection. Its application to refraction data, however, is more recent than its application to reflection data; see *Levander et al.* [2005].

There are different types of seismic imaging method. We consider reverse time migration [*Baysal et al.*, 1983; *McMechan*, 1983; *Whitmore*, 1983] and the category of approaches known under the collective names of Kirchhoff migration [*Bleistein et al.*, 2001] or generalized Radon transform inversion. For example, *Beylkin* [1985] and *Stolk and De Hoop* [2002] analyze inverse scattering using asymptotic methods and *Op't Root* et al. [2012] use the full-wave equation. Applications to global seismology, with waveform data from earthquakes, which (relative to active source exploration) introduces complications due to uncertainties in hypocenter location and source signature, can be found in *Rondenay et al.* [2001], *Poppeliers and Pavlis* [2003], and *Chen et al.* [2005]. Ray theory yields high frequency solutions of the wave equation and essentially is applicable in sufficiently smooth background models. Solving the full wave equation leads to robust solutions, which admit background models of limited regularity.

Brytik et al. [2012] developed a comprehensive theory for RTM-based (elastic) inverse scattering with converted waves in anisotropic media. Building on that analysis we present here a wave equation method for the migration of converted waves to image crust and mantle structures using teleseismic array data. Here, we use a bilinear imaging operator acting on the data, which is essentially a cross-correlation operator of all receiver pairs. Passive source RTM of converted waves differs fundamentally from single station receiver function (RF) analysis and also in several important ways from more traditional RF migration by, for instance, *Chen et al.* [2005]. Firstly, it concerns reverse-time continuation, in which the raw data from individual earthquakes are back propagated, whereas *Chen et al.* [2005] uses CCP stacked RFs, which degrades spatial resolution because of the implicit 1D assumption in RF construction that the interfaces are (locally) horizontal. Moreover, it solves the elastic wave equations explicitly and accounts for wave phenomena such as the formation of caustics underneath the array.

2.2 Methodology

Teleseismic *P* (or S) waves impinging on an interface from below can convert to other modes (e.g. *P*-to-*S* or *S*-to-*P*), which then propagate with different wave speeds to seismographs at the Earth's surface (Fig. 2-1). With assumed wave propagation speeds and angles of incidence, traditional receiver functions simply convert the arrival time difference between transmitted and converted phases as measured at a single station down to the depth where the conversion occurs (Fig. 2-1B), and signal-to-noise is enhanced by stacking over data with common conversion points. As is illustrated in Fig. 2-1B, for single station RF analysis the points where refraction (e.g., *P-P*) and mode conversion (e.g., *P-S*) occur are not the same, which introduces the need to assume that – at least locally (that is, within the first Fresnel zone of the incident wave) – the interface is horizontal. In contrast, RTM exploits the entire wave field, as sampled by a dense array, to locate the points where scattering (including mode conversion) occurs (Fig. 2-1A). Noise suppression in RTM occurs through the continuity property of the underlying inverse scattering transform on the one hand and through stacking over different sources (or events) on the other hand. The resolution can be expressed in terms of the Rayleigh diffraction limit (e.g. $\lambda/4$, λ is the incident wave wavelength).

Snapshots of the wave field can be reconstructed by back propagation of the recorded array data. Due to the difference in P and S wavespeed, the relevant parts of the transmitted and converted wave fronts were in the same location only at the time of conversion. While strictly correct only in the absence of multiple scattering, this observation was used in developing the basic "imaging condition" [e.g., *Claerbout*, 1971] in exploration seismology. The conversion time (and the corresponding location of the point where conversion occurred) can be found by applying the imaging condition (here, a cross-correlation function between different wave modes).

There are three main steps in teleseismic, passive source RTM. Firstly, for each earthquake, using reverse-time continuation, the snapshots of the elastic wave field are reconstructed from the recorded multicomponent array data. Secondly for each snapshot P and S constituents are separated by polarization decomposition (detailed later). Finally the imaging condition, derived from a cross-correlation between P and S wave constituents, is applied. The final image is then obtained by summation of the images from individual events.

A major challenge for passive source RTM is the uncertainty in source signature and the fact that all sources are different. For *P* (*S*) wave incidence we estimate the source signature from the vertical (radial) component array data by multichannel analysis, e.g. multichannel cross correlation (MCCC) and principle component analysis (PCA) [*Rondenay and Fischer*, 2003; *VanDecar and Crosson*, 1990]. The elastic Green's functions are then estimated by deconvolution of the source signature from all three components, and the wave field is reconstructed by reverse-time continuation from the surface boundary data [e.g. *McMechan*, 1983]. Here we solve the elastic wave equation in the time domain using a staggered grid finite difference propagator [*Virieux*, 1986].

In 3-D (isotropic elastic media), for each time step, the reconstructed displacement wavefield $\mathbf{u}_r$ can be decoupled into *P*, *SV* and *SH* components by projection operators Q^* [*Brytik et al.*, 2011]:

$$u_{r,\mathrm{P}} = Q_{\mathrm{p}}^{*}\,\mathbf{u}_r \quad u_{r,\mathrm{SV}} = Q_{\mathrm{SV}}^{*}\,\mathbf{u}_r \quad \text{and} \quad u_{r,\mathrm{SH}} = Q_{SH}^{*}\,\mathbf{u}_r\ , \tag{2.1}$$

which are defined as

$$\begin{aligned} Q_{\mathrm{P}}^{*} &= (-\Delta)^{-1/2}\frac{1}{i}\begin{pmatrix} \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \end{pmatrix} \\ Q_{\mathrm{SV}}^{*} &= (-\Delta)^{-1/2}(-\Delta)^{-1/2}\begin{pmatrix} \dfrac{\partial^2}{\partial x_1 \partial x_3} & \dfrac{\partial^2}{\partial x_2 \partial x_3} & -\dfrac{\partial^2}{\partial x_1^2}-\dfrac{\partial^2}{\partial x_2^2} \end{pmatrix} \\ Q_{\mathrm{SH}}^{*} &= (-\Delta')^{-1/2}\frac{1}{i}\begin{pmatrix} -\dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_1} & 0 \end{pmatrix} \end{aligned} \tag{2.2}$$

Here, Δ is the Laplacian operator, Δ' is $\partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$, and i is the imaginary unit. In 2-D these operators degenerate to:

$$\begin{aligned} Q_{\mathrm{P}}^{*} &= (-\Delta)^{-1/2}\frac{1}{i}\begin{pmatrix}\dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2}\end{pmatrix} \\ Q_{\mathrm{S}}^{*} &= (-\Delta')^{-1/2}\frac{1}{i}\begin{pmatrix}-\dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_1}\end{pmatrix} \end{aligned} \tag{2.3}$$

After decoupling of *P* and *SV* components, the imaging condition for *P* wave incidence from a single source is applied [*Brytik et al.*, 2012]:

$$\begin{aligned} I(\mathbf{x}) = \int_0^T \Bigg\{ & \left(\frac{1}{i}\partial_t u_{r,\mathrm{p}}(\mathbf{x},t)\right)\left(\frac{1}{i}\partial_t u_{r,\mathrm{sv}}(\mathbf{x},t)\right) \\ & + \left(\frac{1}{i}\nabla\right)u_{r,\mathrm{p}}(\mathbf{x},t)\cdot\left(\frac{1}{i}\beta(\mathbf{x})(-\Delta)^{-1/2}\partial_t\nabla\right)u_{r,\mathrm{sv}}(\mathbf{x},t)\Bigg\}dt \end{aligned} \tag{2.4}$$

where *T* is the total recorded time, and $\beta(\mathbf{x})$ is *S* wave velocity. For *S* wave incidence, one can exchange $u_{r,\mathrm{p}}$ with $u_{r,\mathrm{sv}}$ and replace $\beta(\mathbf{x})$ by $\alpha(\mathbf{x})$ (*P* wave velocity). At sufficiently high frequencies the operator $(-\Delta)^{-1/2}\partial_t$ can be approximated (in the asymptotic limit) as $\beta(\mathbf{x})$ – or $\alpha(\mathbf{x})$ for *S* wave incidence.

The transmission (reflection) coefficient $\mathrm{T}_{\mathbf{ps}}$ ($\mathrm{R}_{\mathbf{ps}}$) is an odd function of *P* wave incidence angle [*Aki and Richards*, 2002; *Balch and Erdemir*, 1994]. In absence of elaborate corrections for radiation patterns (see *Brytik et al.* [2012]), at least we need to correct for the sign of converted *SV* waves in order to constructively sum all partial images. We determine the sign by evaluating the Poynting vector [e.g. *Červený*, 2001; *Dickens and Winbow*, 2011] to find the incidence angle of *P* wave and transmission angle of converted *SV* wave. The sign of *SV* waves is then reversed, if necessary.

2.3 Numerical examples

We demonstrate the performance of converted wave RTM with numerical experiments, using synthetic data generated by a fourth order finite difference scheme. We also investigate the effect of background models on image quality and compare the performances of RTM with CCP stacks of RFs. Two models with lateral heterogeneities are employed to generate synthetic data: a lens-kink model and a layer-kink model (Fig. 2-2C and 2-2D). Both models contain a discontinuous interface. In the lens-kink model a Gaussian low velocity lens is used to investigate effects of wave phenomena caused by a low velocity anomaly in the crust. Caustics can be observed in the data recorded at the surface (two components of the particle velocity) (Fig. 2-2A, inset). Interior (explosive) point sources are placed at the bottom of the models (Fig. 2-2C and 2-2D) and injected as *P* arrivals. We use a total of 22 events, but the contribution of individual sources is also illustrated. The central frequency of the source function is 2 Hz. The spacing between receivers (at the surface) is 300 m, which is smaller than the "spatial" Nyquist frequency and avoids "spatial" aliasing.

To assess the effect of background model, we use two types of model for the inversion step (that is, the back propagation): a 1-D model with a linear increase of wave speed with depth (Fig. 2-3A) and a 2-D smooth model with (smooth) spatial variations in wave speed that resemble the actual structures (Fig. 2-3D). We note that the latter could be obtained, for instance, through tomography. Comparing the RTM imaging results for the lens-kink model (Fig. 2-3B and 2-3E), the kink structure is, as expected, better recovered if we use the 2-D background model. Absence of the low velocity lens in the background model introduces artificial topography on the horizontal layer between 20~40 km horizontal distance (Fig. 2-3E), and inclusion of the lens (Fig. 2-3D, inset) yields near perfect recovery (Fig. 2-3E, inset). In the layer-kink model, the 2D smooth model improves the reconstruction of the vertical structure (Fig. 2-3F) and the flat interface below the kink. The amplitude along this flat interface is not balanced (e.g. amplitude decrease can be observed for x>50 km range) due to illumination effects; this common phenomenon can be corrected in angle domain [e.g. *Wu et al.*, 2004] but that is not done here. A partial image from one earthquake (Fig. 2-3F, inset) suggests that the main structures can be revealed by only a few events as long as they provide good illumination.

In Fig. 2-4 we compare converted wave RTM with CCP stacking results. The models used to generate the synthetic wavefields are the same as in Fig. 2-3 except the edges of the kink discontinuity are smoothed to suppress the corner diffraction that would otherwise overwhelm the CCP stacks even more. Three localized plane waves (plane waves tapered by a Gaussian window) with different position and incident angles are used as incident waves (Fig. 2-4A and 2-4D). For this comparison, we use the same 1-D background model for CCP stacking and RTM. Even for this relatively simple model, the wave field appears to be too complicated for CCP stacking to reconstruct the input model (Fig. 2-4B and 2-4E). Indeed, the horizontal interfaces can barely be discerned among the image artifacts, and the vertical structure is not recovered at all. In contrast, despite the complexity of the wavefield, the RTM images reveal clearly the interfaces, even with data from only three sources.

2.4 Discussion

Linearized imaging methods, either ray-based or wave equation based, are sensitive to the background model. We demonstrate here that a background model with lateral heterogeneity can greatly improve the image quality, especially near vertical structures. A smooth background model can be estimated from geological models, from travel time or surface wave tomography – either in active or passive (e.g., ambient noise) studies, or by wave equation (WE) reflection tomography [*Burdick et al.*, 2013]. Indeed, in the future we aim to combine passive source RTM and WE reflection tomography in an explicit (non-linear) joint inversion of teleseismic wavefields for wavespeed and interface location.

With the simple models used here, the kinked interface violates the flat layer assumption in CCP stacking and the low velocity lens causes the formation of caustics that cannot be accounted for with CCP stacking. The diffraction artifacts in the CCP stacking image can be suppressed somewhat if more events are included, but breakdown of (translational) symmetry assumptions cannot be avoided. In contrast, RTM – a wave equation method based on an artifact-free imaging condition – accounts for wavefield complexity and effectively migrates corner diffraction energy back to the proper position in the final image (Fig. 2-4C and 2-4F). Another indication of the promise of passive source RTM is that only a few earthquakes with good illumination coverage and high signal-to-noise ratio are sufficient, even in geological environments that render CCP stacking ineffective. Given the irregular distribution of naturally occurring earthquakes, our method can improve seismic imaging in large areas, provided, of course, that data from dense seismograph arrays are available.

In our demonstration of the concept and promise of passive source RTM we suppressed several practical challenges. Firstly, RTM has stringent sampling requirements. The sampling theorem suggests that the spatial interval Δx, in principle, should be less than $(\lambda_a)_{min}/2$, where $(\lambda_a)_{min}$ is the minimum horizontal apparent wavelength for a given depth. For a typical teleseismic study, with frequencies around or above 1Hz, a station interval of 5 km should be sufficient [see *Chen et al.*, 2005 for details]. In the real acquisition, irregular and sparse (aliased) sampled data can introduce significant "noise" in RTM. Such problems can be partially mitigated by interpolation in frequency-wavenumber [*Spitz*, 1991; *Zwartjes and Sacchi*, 2007] or curvelet domain [*Naghizadeh and Sacchi*, 2010]. Secondly, the effect of limited illumination needs to be compensated in passive source RTM. In our experiments, the 22 point sources at the bottom provide near full illumination coverage. In real data, the incident angle of teleseismic wave is constrained in a narrow range (~ 15°- 40°) due to the sparsity and irregularity of the earthquake distribution. This means that geological structures can only be partially imaged. The illumination (aperture) effect can be corrected by computing the (diagonal of) normal operator and its approximate inverse [e.g. *De Hoop et al.*, 2009]. Another problem is the presence of multiples, for instance the reverberation in shallow sedimentary layers. Certain techniques from exploration seismology can to some extent suppress the multiples [e.g. *Berkhout and Verschuur*, 1997]. These challenges are not unique to RTM, however, and thus beyond the scope of this paper.

2.5 Summary

We present a method for subsurface imaging with multi-component data from dense seismograph arrays that does not rely on simplifying (e.g., 1-D) assumptions about the geometry of the geological structures of interest. The array data are backward propagated by solving the elastic wave equation directly. After polarization separation, a modified cross correlation imaging condition between *P* and *S* wave constitutes is applied to obtain an inverse scattering transform. From synthetic experiments, it is evident that for complex geological structures the new method is superior to the traditional CCP receiver function stacking, provided that data from dense seismograph arrays are available. At present, few arrays are suitable for applications of teleseismic RTM without substantial preprocessing (including interpolation), but in view of the trend to deploy increasingly dense arrays we expect that passive source RTM will become feasible – and even routine – in the near future.

Acknowledgements

We thank Alan Levander and an anonymous reviewer for their valuable comments.

References

Aki, K., and P. G. Richards (2002), *Quantitative seismology*, University Science Books.

Balch, A., and C. Erdemir (1994), Sign change correction for prestack migration of P-S converted wave reflections, *Geophys. Prospect.*, *42*(6), 637-663.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood (1983), Reverse time migration, *Geophysics*, *48*(11), 1514-1524.

Berkhout, A. J., and D. J. Verschuur (1997), Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations, *Geophysics*, *62*(5), 1586-1595.

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

Bleistein, N., J. K. Cohen, and J. W. Stockwell (2001), *Mathematics of multidimensional seismic imaging, migration, and inversion*, Springer Verlag.

Brytik, V., M. V. de Hoop, and R. D. van der Hilst (2012), Elastic-wave inverse scattering based on reverse time migration with active and passive source reflection data, in *Inside Out - Inverse Problems, MSRI Publications*, Cambridge University Press.

Brytik, V., M. V. de Hoop, H. F. Smith, and G. Uhlmann (2011), Decoupling of modes for the elastic wave equation in media of limited smoothness, *Commun. Part. Diff. Eq.*, *36*(10), 1683-1693.

Burdick, S., M. de Hoop, S. Wang, and R. van der Hilst (2013), Teleseismic wave-equation reflection tomography using free surface reflected phases, *Geophys. J. Inter. in revision.*

Červený, V. (2001), *Seismic Ray Theory*, Cambridge University Press.

Chen, L., L. Wen, and T. Zheng (2005), A wave equation migration method for receiver function imaging: 1. Theory, *J. Geophys. Res.*, *110*(B11), B11309.

De Hoop, M., H. Smith, G. Uhlmann, and R. Van der Hilst (2009), Seismic imaging with the generalized Radon transform: a curvelet transform perspective, *Inverse Problems*, *25*, 025005.

Dickens, T. A., and G. A. Winbow (2011), RTM angle gathers using Poynting vectors, *SEG Technical Program Expanded Abstracts*, 2011-3109.

Dueker, K. G., and A. F. Sheehan (1997), Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track, *J. Geophys. Res.*, *102*, 8313-8327.

Gilbert, H. J., A. F. Sheehan, K. G. Dueker, and P. Molnar (2003), Receiver functions in the western United States, with implications for upper mantle structure and dynamics, *J. Geophys. Res*, *108*, 2229.

Levander, A., F. Niu, and W. W. Symes (2005), Imaging teleseismic P to S scattered waves using the Kirchhoff integral, *Geophys. Monograph-American Geophys. Union*, *157*, 149.

McMechan, G. A. (1983), Migration by Extrapolation of Time-Dependent Boundary Values, *Geophys. Prospect.*, *31*(3), 413-420.

Naghizadeh, M., and M. D. Sacchi (2010), Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data, *Geophysics*, *75*(6), WB189-WB202.

Op't Root, T. J. P. M., C. C. Stolk, and M. V. De Hoop (2012), Linearized inverse scattering based on seismic reverse time migration, *J. Math. Pures Appl.*, *98*(2), 211-238.

Poppeliers, C., and G. L. Pavlis (2003), Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 1. Theory, *J. Geophys. Res.*, *108*(B2), 2112.

Rondenay, S. (2009), Upper mantle imaging with array recordings of converted and scattered teleseismic waves, *Surv. Geophys.*, *30*(4), 377-405.

Rondenay, S., and K. M. Fischer (2003), Constraints on localized coremantle boundary structure from multichannel, broadband SKS coda analysis, *J. Geophys. Res*, *108*, 2537.

Rondenay, S., M. G. Bostock, and J. Shragge (2001), Multiparameter two-dimensional inversion of scattered teleseismic body waves 3. Application to the Cascadia 1993 data set, *J. Geophys. Res.*, *106*(B12), 30795-30807.

Sheehan, A. F., P. M. Shearer, H. J. Gilbert, and K. G. Dueker (2000), Seismic migration processing of P-SV converted phases for mantle discontinuity structure beneath the Snake River Plain, western United States, *J. Geophys. Res.*, *105*(B8), 19055-19065.

Spitz, S. (1991), Seismic trace interpolation in the *F-X* domain, *Geophysics*, *56*(6), 785-794.

Stolk, C. C., and M. V. de Hoop (2002), Microlocal analysis of seismic inverse scattering in anisotropic elastic media, *Commun. Pur. Appl. Math*, *55*(3), 261-301.

VanDecar, J., and R. Crosson (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seismol. Soc. Am.*, *80*(1), 150-159.

Virieux, J. (1986), P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method, *Geophysics*, *51*(4), 889-901.

Whitmore, N. D. (1983), Iterative depth migration by backward time propagation, *SEG Technical Program Expanded Abstracts*, *2*(1), 382-385.

Wu, R. S., M. Luo, S. Chen, and X. B. Xie (2004), Acquisition aperture correction in angle domain and true amplitude imaging for wave equation migration, *SEG Technical Program Expanded Abstracts*, *23*(1), 937-940.

Zhai, Y., and A. Levander (2011), Receiver function imaging in strongly laterally heterogeneous crust: Synthetic modeling of BOLIVAR data, *Earthquake Science*, *24*(1), 45-54.

Zhu, L. (2000), Crustal structure across the San Andreas Fault, southern California from teleseismic converted waves, *Earth Planet. Sci. Lett.*, *179*(1), 183-190.

Zwartjes, P., and M. Sacchi (2007), Fourier reconstruction of nonuniformly sampled, aliased seismic data, *Geophysics*, *72*(1), V21-V32.

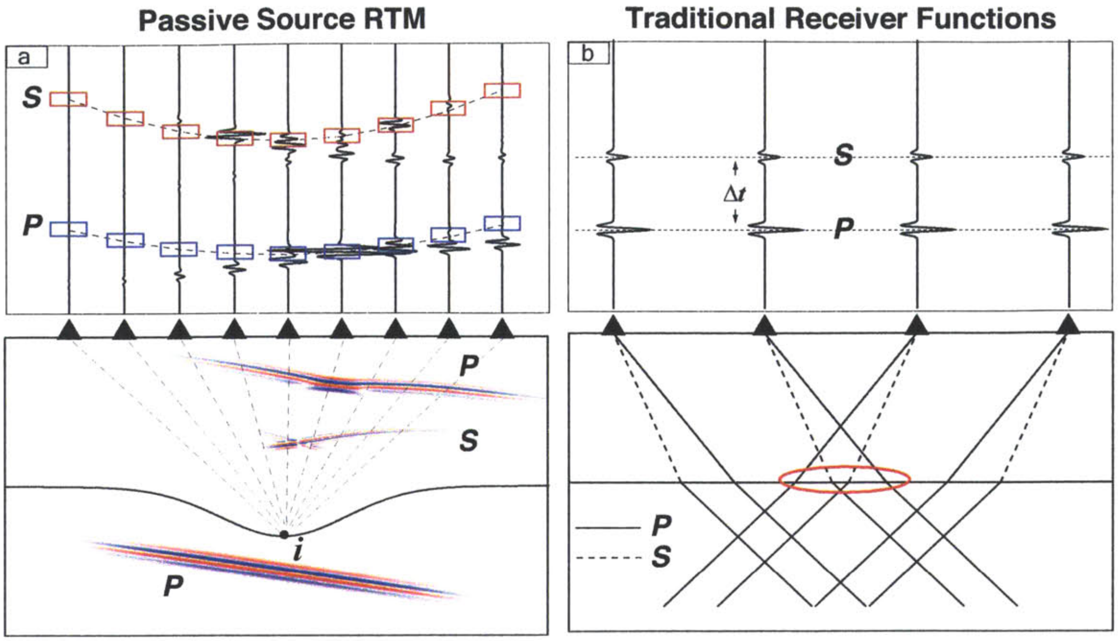


Figure 2-1 Schematic illustration of (A) passive-source reverse time migration and (B) common conversion point (CCP) stacking of traditional receiver functions. (A) Bottom: *P* waves impinging on a contrast produces direct *P* and converted *S* waves, which can be recorded at an array of seismograph stations at the surface (black triangles). Top: *P* and *S* energy due to scattering at, say, point *i* arrives at different times (blue and red boxes, respectively). In inverse sense, the location point *i* can be reconstructed by optimization of the correlation between the back projected *P* and *S* wavefields. (B) Bottom: in traditional receiver functions, the *P-SV* conversion is assumed to occur at an interface that is (locally) horizontal. Top: The travel time difference Δ*t* between transmitted *P* and converted *S* is a measure of interface depth, and data redundancy is obtained by stacking over common conversion points (CCP), shown as a red oval.

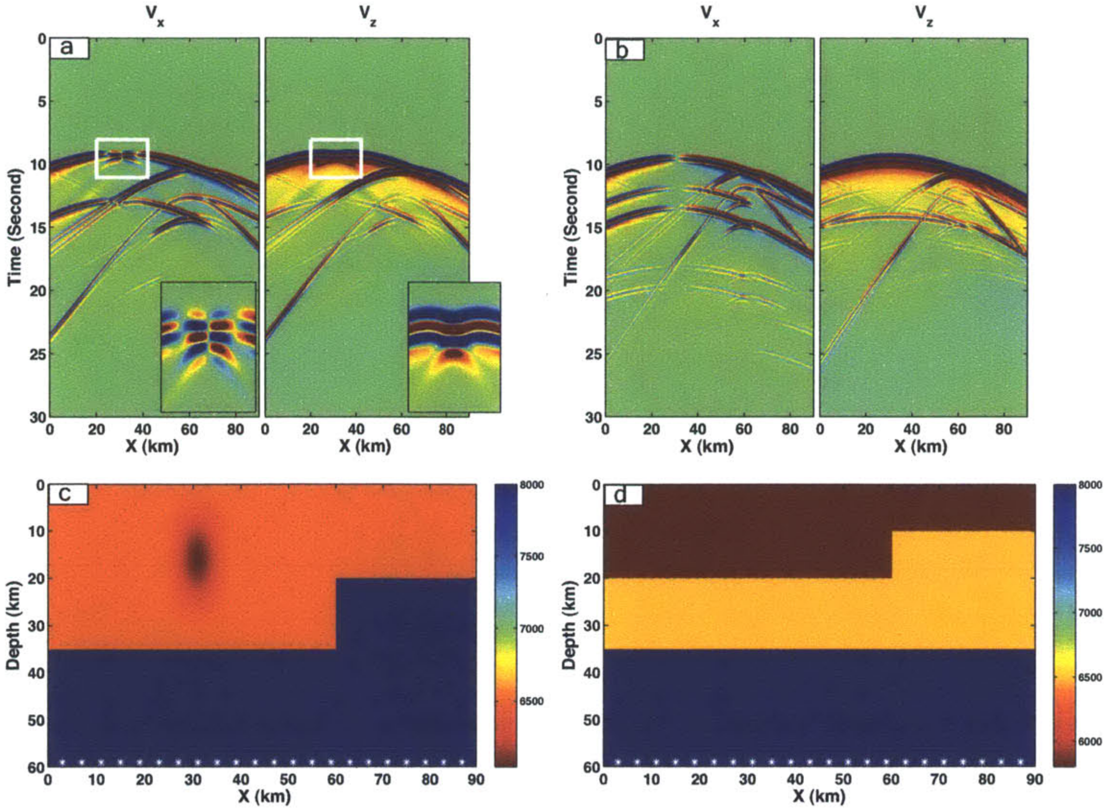


Figure 2-2 Synthetic data (A, B) computed for two test models – a lens-kink model (C) and a layer-kink model (D) – with v_x and v_z the horizontal and vertical components, respectively, of particle velocity. The dimension of the models is 90 km by 60 km. Receivers are at the surface and up to 22 (explosive) sources are located at the bottom of the models, shown as white stars in (C) and (D). The source central frequency is 2 Hz, and a Ricker wavelet is chosen as the source time function. In the lens-kink model, a 7% low velocity Gaussian lens forms a crustal low velocity anomaly; the caustics produced by this lens are visible in the inset of (A). Corresponding *S* and mass density models are obtained through scaling of the *P* models shown here.

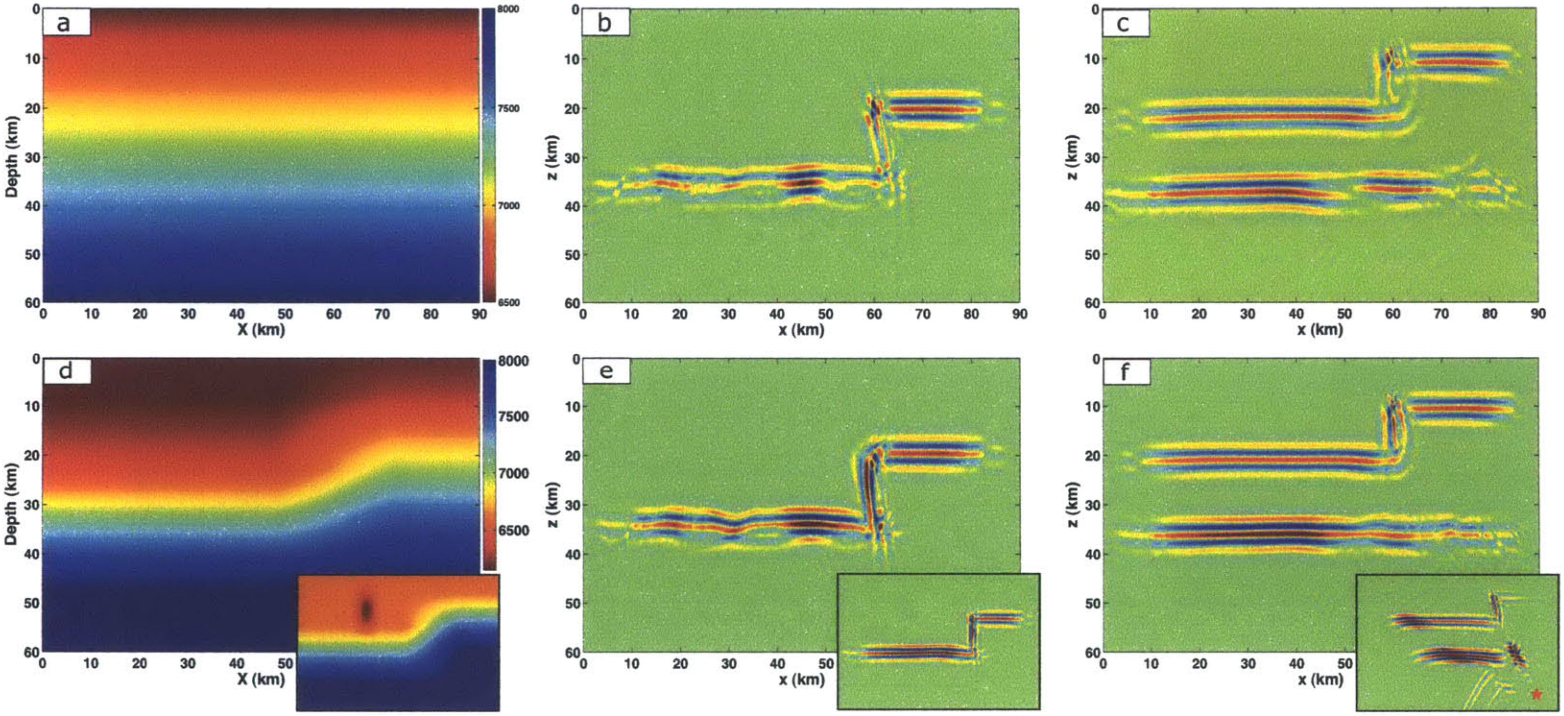


Figure 2-3 Background models (A, D) and converted wave RTM imaging results (B, C and E, F). (A) Gradual increase of wavespeed with depth. (B, C) Recovery of the lens-kink model (data shown in Fig. 2-2A) and layer-kink model (data in Fig. 2-2B) using the 1-D background model shown in (A). (D) Smooth 2-D background (inset: with low velocity lens). (E) Recovery of the lens-kink model with the 2-D background model (inset: recovery using 2-D model that includes the low velocity lens). (F) Recovery of the layer-kink model with the 2-D background model (inset: partial image obtained with data from a single source, red star).

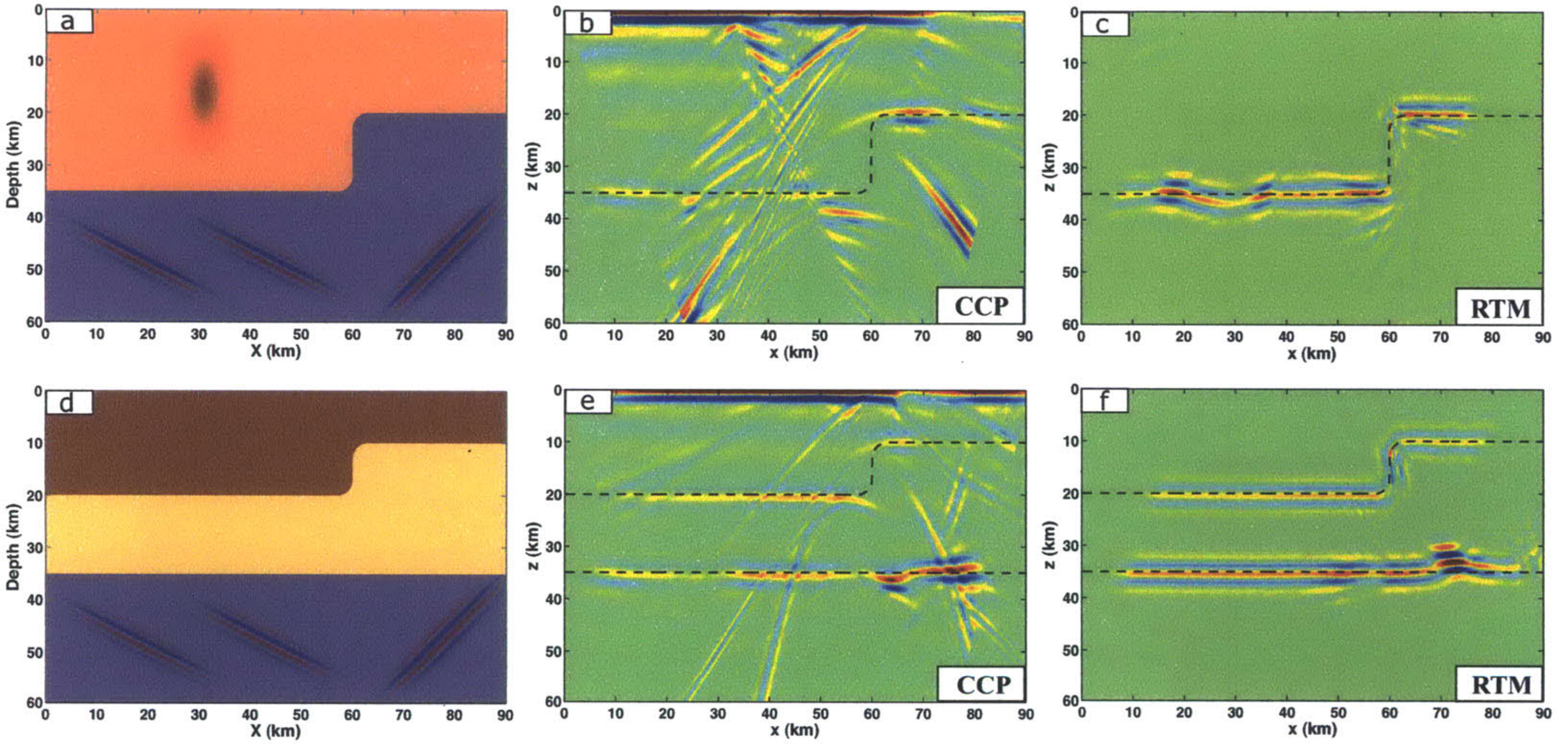


Figure 2-4 Comparison between CCP stacking and RTM. (A) Lens-kink model. Three localized *P* plane waves are used as incident sources to generate synthetic data. The localized plane waves with different incident angles are shown in the bottom layer. (B) CCP stacking results using 1D linearly increased model. (C) RTM result based on the same 1D model as in (B). (D) Layer-kink model. The same localized *P* plane waves as of (A) are used for the generation of synthetic data. (E) CCP stacking results using 1D linearly increased model. (F) RTM result based on the same 1D model as in (E). The dashed lines in (B), (C), (E) and (F) show the interfaces in true models (A) and (D).

Chapter 3

Teleseismic Data Preprocessing for Inverse Scattering Applications

In this chapter, we describe in detail the preprocessing of raw seismic data and retrieval of the Earth's impulse response (Green's function) for our inverse scattering applications. There are three main steps: instrument response removal, source signature estimation, and source deconvolution. Furthermore, noisy and dead traces need to be discarded. We propose a semi-automatic workflow that is suitable for applications to large datasets. Firstly, the instrument response is removed. The traces are then aligned with a reference phase by multi-channel cross-correlation. The noisy traces are marked and removed automatically at the same time. The polarity of traces is corrected according to the sign of cross-correlation coefficients. After alignment, the source signature is estimated by principle component analysis. Finally, the source signature is deconvolved by a Wiener inverse filter, which is optimal, data-adaptive, and, therefore, superior to traditional water level deconvolution.

3.1 Introduction

With rapidly increasing broadband seismic data, especially deployment of dense arrays, an efficient and robust preprocessing workflow is necessary for teleseismic imaging. The recorded raw data are a convolution of earthquake source, Earth's impulse response (Green's function), and instrument response. The main purpose of data preprocessing is to extract the Green's function and increase single-to-noise (SNR) ratio. Normally the workflow includes removal of bad (or dead) traces, suppression of random noise, and deconvolution of the instrument response and source terms. Compared with the controlled seismic experiments, one essential challenge is that the source mechanism of natural earthquakes is highly non-uniform and usually not well known. Therefore, a key step in the preprocessing workflow is source normalization, involving source magnitudes, unknown source time functions, and intrinsic source radiation patterns [*Zheng*, 2007].

Multiple earthquakes illuminate the target region from different directions. The final seismic image is obtained by stacking all of the partial images from individual events. If the source magnitude were not taken into account, large earthquakes would dominate the final image. Such a problem can be solved by introducing reference phases, which are easily identified in the data, e.g. direct *S* wave [*Wang et al.*, 2006], surface multiple *SS* wave [*Cao et al.*, 2010], and depth phase *pP* [*Zheng*, 2007]. The traces then can be normalized with respect to the reference phase. The precursors or coda of the reference phase can be used for fine structure imaging.

The earthquake mechanism (source time function and radiation pattern) varies across events. It can be estimated by long-period waveform inversion [e.g. *Dziewoński et al.*, 1981; *Sipkin*, 1982; *Nábělek*, 1984]. The centroid-moment-tensor (CMT) solution for $M_w \geq 5.0$ events can be retrieved from the Internet website www.globalcmt.org [*Ekström et al.*, 2012]. With recordings from dense arrays, for instance USArray and dense arrays in Europe and China, the source signature can be estimated by multi-channel techniques, e.g. multi-channel cross-correlation (MCCC), and principle component analysis (PCA) [e.g. *Rondenay et al.*, 2005; *VanDecar and Crosson*, 1990].

In the following sections, we introduce an efficient and robust workflow combining source normalization and data selection. There are three main steps: (i) align the data with respect to the incident wavefield. Meanwhile, the polarity of particle motion is corrected and bad traces are removed automatically; (ii) estimate the source signature; (iii) deconvolve the source signature.

3.2 Source Estimator and Data Selector

3.2.1 Shift and Sum

The coherency of teleseismic waveforms leads naturally to the use of correlation methods. We assume that for a given earthquake each receiver in a seismic array contains a time-shifted signal as well as noise. Receiver *i* records the data as,

$$d_i(t) = f(t - t_i) + n_i(t) \ , \tag{3.1}$$

where t_i is relative delay time and $n_i(t)$ is the noise term. One can consider the coherent signal $f(t)$ as an effective source wavelet (or source signature). Note that the source signature may include source side reverberations that have similar moveout curves as the incident wave, while the scattered wavefield with distinct moveout is regarded as noise [e.g. *Rost and Thomas*, 2002; *Rondenay*, 2009].

We follow *Rost and Thomas* [2002] but proceed in the frequency domain. Eq. (3.1) then becomes:

$$d_i(\omega) = f(\omega)e^{j\omega t_i} + n_i(\omega) \tag{3.2}$$

or

$$f(\omega) = d_i(\omega)e^{-j\omega t_i} - n_i(\omega)e^{-j\omega t_i} \ . \tag{3.3}$$

The source signature can be estimated by "shift and sum" as

$$\hat{f}(\omega) = \frac{1}{M}\sum_{i=1}^{M} d_i(\omega)e^{-j\omega t_i} \ , \tag{3.4}$$

where M is the number of receivers. It is an unbiased estimator of $f(\omega)$, given the noise term (including white noise and correlated noise with distinct moveout) is stacked out

$$\frac{1}{M}\sum_{i=1}^{M} n_i(\omega)e^{-j\omega t_i} = 0 \ . \tag{3.5}$$

One can also use principle component analysis to better separate incident and scattered wavefields, see *Rondenay et al.* [2005] for more details.

3.2.2 Multi-Channel Cross-Correlation

Multi-channel cross-correlation technique is widely used to measure the differential travel time [*VanDecar and Crosson*, 1990]. Here we carefully tailor it to estimate the relative delay time t_i in Eq. (3.1), identify and correct for opposite polarities, and to conduct data quality control.

For a given teleseismic event, a preliminary travel time t_i^P is calculated for receiver i based on a reference model, e.g. PREM [*Dziewoński and Anderson*, 1981]. The cross-correlation function between the ith and jth receiver can be defined as

$$\phi_{ij}(\tau)=\int_{t_0}^{t_0+T} d_i(t+t_i^P)d_j(t+t_j^P+\tau)\,dt \ , \qquad (3.6)$$

where τ is the lag time between two traces; t_0 is the time between the preliminary arrival time and the starting time of correlation window; T is the length of correlation window, which can be chosen 2~3 cycles of the dominant frequency.

The relative delay time between two traces is

$$\Delta t_{ij}=t_i-t_j=t_i^P-t_j^P-\tau_{ij}^{\max} \ , \qquad (3.7)$$

where $\tau_{ij}^{\max}$ is the lag time between two traces when the absolute value of ϕ_{ij} reaches the maximum. In practice, one can constrain the range of $\tau_{ij}^{\max}$ in order to avoid cycle skips.

For M stations, Eq. (3.7) generates $M(M\text{-}1)/2$ overdetermined equations, and the following constraint equation can be added to remove the systematic error [*VanDecar and Crosson*, 1990]:

$$\sum \Delta t_{ij}=0 \ . \qquad (3.8)$$

Combing Eq. (3.7) and (3.8), the relative time delay for ith receiver t_i can be calculated by a least-square fitting.

The cross-correlation coefficient for two traces can be expressed as

$$r_{ij} = \frac{\phi_{ij}(\tau_{ij}^{\max})}{\sigma_i \sigma_j} , \tag{3.9}$$

where σ_i^2 is the variance of ith trace in the correlation time window. Note the cross-correlation coefficient in Eq. (3.9) is normalized and $-1 \le r_{ij} \le 1$. So it is an important indicator for data quality control. For a dead or highly distorted trace, the cross-correlation coefficient is consistently low (with a small variance). For a good trace with opposite polarity, it would produce high value but negative correlation coefficients.

3.2.3 Polarity Corrections and Data Selection

For true amplitude seismic imaging, it is important to take the intrinsic radiation patterns of earthquakes into account, both the polarity (sign) and amplitude of traces. In practice, the trace can be normalized with respect to a reference phase to compensate the amplitude effect. The polarity of the trace needs to be corrected if necessary, otherwise the data would be destructively stacked and produce degenerated migration image.

In a double-couple fault model, the radiation pattern for *P*, *SV*, and *SH* modes can be expressed as

$$\begin{aligned} F^{\mathrm{P}} &= \cos\lambda \sin\delta \sin^2 i_\xi \sin 2(\phi - \phi_s) - \cos\lambda \cos\delta \sin 2i_\xi \cos(\phi - \phi_s) \\ &\quad + \sin\lambda \sin 2\delta (\cos^2 i_\xi - \sin^2 i_\xi \sin^2(\phi - \phi_s)) \\ &\quad + \sin\lambda \cos 2\delta \sin 2i_\xi \sin(\phi - \phi_s) \end{aligned} \tag{3.10}$$

$$\begin{aligned} F^{\mathrm{SV}} &= \sin\lambda \cos 2\delta \cos 2i_\xi \sin(\phi - \phi_s) - \cos\lambda \cos\delta \cos 2i_\xi \cos(\phi - \phi_s) \\ &\quad + \frac{1}{2}\cos\lambda \sin\delta \sin 2i_\xi \sin 2(\phi - \phi_s) \\ &\quad - \frac{1}{2}\sin\lambda \sin 2\delta \sin 2i_\xi (1 + \sin^2(\phi - \phi_s)) \end{aligned} \tag{3.11}$$

$$\begin{aligned} F^{\mathrm{SH}} &= \cos\lambda \cos\delta \cos i_\xi \sin(\phi - \phi_s) + \cos\lambda \sin\delta \sin i_\xi \cos 2(\phi - \phi_s) \\ &\quad + \sin\lambda \cos 2\delta \cos i_\xi \cos(\phi - \phi_s) \\ &\quad - \frac{1}{2}\sin\lambda \sin 2\delta \sin i_\xi \sin 2(\phi - \phi_s) \end{aligned} \tag{3.12}$$

where ϕ_s, δ, and λ are fault strike, dip, and rake angle respectively; i_ξ is takeoff angle of the outgoing ray and ϕ is source-receiver azimuth [*Aki and Richards*, 2002].

The Harvard CMT solution for the source mechanism can be used as a reference for trace polarity correction. However, it is estimated from long period seismograms recorded on globally distributed seismographic stations [see *Ekström et al.*, 2012 and references therein], and not very accurate for short period body waves. We show in Fig. 3-1 the discrepancies between the polarity prediction from the CMT solution and real data measurements. The trace polarity is calculated from the CMT solution for direct *P* arrivals (Fig. 3-1A) and *SH* arrivals (Fig. 3-1B) with Eq. (3.10) and Eq. (3.12) respectively. The beach balls depict the CMT solution and triangles are seismic stations. The green triangles are stations that agree with the real data measurements. The red ones are those with opposite polarity to the CMT prediction, and the black indicate unclear polarity picks. This comparison suggests that there are quite a few outliners in both *P* and *S* polarity prediction from CMT solutions.

Here we use a data quality indicator to correct the trace polarity based on multi-channel cross-correlation. As mentioned above, the cross-correlation coefficient in Eq. (3.9) is a useful indicator for the data quality. An average cross-correlation coefficient $\overline{r_i}$ for ith trace can be introduced as

$$\overline{r_i} = \frac{1}{M-1} \sum_{j=1, j \neq i}^{M} \left| r_{ij} \right| . \tag{3.13}$$

Those traces with $\overline{r_i}$ below a given threshold are considered as noisy (or dead) data, and discarded for further processing. The trace with the maximum value of $\overline{r_i}$ is then used as a reference trace. The polarity of all other traces is corrected accordingly,

$$\mathrm{sign}(d_i) = \mathrm{sign}(r_{i,ref}) , \tag{3.14}$$

where $r_{i,ref}$ is the cross-correlation coefficient between ith trace and the reference trace. It is possible that after polarity correction the sign for one earthquake may differ from another event (e.g. all first arrivals are negative for one event, but positive for another). However, after the source deconvolution, all events are considered as isotropic explosive point sources consistently.

3.2.4 Synthetic and Real Data Examples

In this section, we present two examples of polarity correction and data selection, one of which is synthetic data and another is from real data. We firstly generate 20 traces with Ricker wavelets to mimic first arrivals, shown in Fig. 3-2A. The amplitudes of Ricker wavelets are modulated by a cosine function, while the polarity and time delay are chosen randomly. In addition, 25% Gaussian noise is superposed for each trace. All the traces are then aligned by multi-channel cross-correlation (Fig. 3-2B) with Eq. (3.6)-(3.8). The cross-correlation coefficient matrix (Fig. 3-3A) and average cross-correlation coefficients (Fig. 3-3B) are calculated as well. Note the 3rd trace and 19th trace are removed automatically due to low cross-correlation coefficients (below our threshold 0.5). After alignment, the 17th trace with maximum average cross-correlation coefficient is selected as the reference trace, and the polarity of all other traces is corrected accordingly with Eq. (3.14). The final result is demonstrated in Fig. 3-2C.

We apply this procedure to a real dataset. One event recorded by USArray is shown in Fig. 3-4. There are four clear phases in the data, direct *S* wave, *ScS*, and their depth phases. The raw data is aligned with a preliminary *S* wave arrival time based on ak135 model [*Kennett et al.*, 1995] (Fig. 3-4A). The traces are further aligned with MCCC (Fig. 3-4B), and some bad traces are removed at the same time (the threshold is 0.4, see Fig. 3-5). The polarity is then flipped if necessary. The surviving traces (Fig. 3-4C) are the input of next preprocessing step, source signature deconvolution.

3.3 Wiener Deconvolution

The observed signal $d(t)$ can be expressed as the convolution of a source wavelet (or signature) with the Earth's impulse response as well as noise,

$$d(t) = s(t) * g(t) + n(t) \ , \qquad (3.15)$$

where $s(t)$ is the source signature and $n(t)$ represents noise.

The deconvolution process, which tries to retrieve the Earth's impulse response $g(t)$ from the data given the source wavelet $s(t)$, is in general an ill-posed problem because of the existence of random noise $n(t)$, limited frequency bandwidth and inaccurate source signature estimation [see *Chen et al.*, 2010]. Such an inversion problem can be stabilized by a water level term, and an approximate solution of $g(t)$ in the frequency domain [e.g. *Wiggins and Ralph*, 1976] is

$$\hat{g}(\omega)=\frac{s^*(\omega)}{|s(\omega)|^2+\delta}d(\omega) \ , \tag{3.16}$$

where δ is the water level, representing the expected noise power. The asterisk superscript denotes the complex conjugate. ω is angular frequency. Normally the value of water level δ is independent of frequency, which implicitly assumes that the noise is white.

Haldorsen et al. [1994] obtained an optimal deconvolution filter by using the data redundancy in a seismic array, without white noise assumption. It is successfully applied in vertical seismic profiling (VSP) [e.g. *Haldorsen et al.*, 1994] and teleseismic data studies [e.g. *Chen et al.*, 2010]. Here we summarize the method for the purpose of self-containment.

Suppose the observed array data $d_m(t)$, with totally M traces, have a common source signature $s(t)$ with a variable noise $n_m(t)$:

$$d_m(t)=s(t)+n_m(t) \ . \tag{3.17}$$

In the frequency domain, Eq. (3.17) can be written as:

$$d_m(\omega)=s(\omega)+n_m(\omega) \ . \tag{3.18}$$

Here the source signature $s(\omega)$ consists of all aligned signals contributing to the source estimation, whereas the noise term $n_m(\omega)$ represents both uncorrelated random noise and coherent signal with the source but with a different moveout across the array.

The deconvolution filter $S(\omega)$ can be determined as the solution to Eq. (3.18) with the spectral whitening constraint:

$$S(\omega)d_m(\omega)=1 \ . \tag{3.19}$$

The least-square solution of Eq. (3.18) and (3.19) is expressed as

$$S(\omega)=\frac{s^*(\omega)}{E_T(\omega)} \ , \tag{3.20}$$

where $E_T(\omega)$ is the average total energy of the raw traces,

$$E_T(\omega)=\frac{1}{M}\sum_{m=1}^{M}\left|d_m(\omega)\right|^2 \ . \tag{3.21}$$

Plugging Eq. (3.18) and (3.21) into Eq. (3.20), we obtain

$$S(\omega)=\frac{s^*(\omega)}{\left|s(\omega)\right|^2+E_N(\omega)} \ , \tag{3.22}$$

where $E_N(\omega)$ is the average noise energy,

$$E_N(\omega)=\frac{1}{M}\sum_{m=1}^{M}\left|d_m(\omega)-s(\omega)\right|^2 \ . \tag{3.23}$$

Therefore, the estimated Green's function can be written as,

$$\hat{g}_m(\omega)=\frac{s^*(\omega)}{\left|s(\omega)\right|^2+E_N(\omega)}d_m(\omega) \tag{3.24}$$

or

$$\hat{g}_m(\omega) = \frac{s^*(\omega)}{E_T(\omega)} d_m(\omega) \ . \qquad (3.25)$$

Comparing Eq. (3.24) with Eq. (3.16), the constant water level is replaced with a noise adaptive regularization term. The optimal deconvolution filter (in the least-square sense) acts as a data-adaptive, band-limiting filter attenuating frequencies where the signal-to-noise ratio is small, and will preserve the part of noise that is spectrally coherent with the source [*Haldorsen et al.*, 1994].

In traditional water level deconvolution, it is crucial to find the most proper parameter δ, which is usually done by trail and error. This is subjective and tedious, however, and not suitable for massive data processing. On the other hand, Wiener deconvolution adaptively estimates the noise spectrum and is an optimal inverse filter. The overhead computational cost is neglectable compared with the water level deconvolution. Two examples of real data are given in Fig. 3-6, 3-7. The raw data are aligned with the direct *S* wave (Fig. 3-6A and 3-7A). The source signatures are estimated by PCA (Fig. 3-6B and 3-7B). We compare the Wiener deconvolution (Fig. 3-6C and 3-7C) with the water level method (Fig. 3-6D and 3-7D). The water level δ is selected as 1% of the maximum value of the source amplitude spectrum. The random (but may not white) noise is suppressed much better by Wiener inverse filter than by the water level method both in high SNR data (Fig. 3-6) and noisy data (Fig. 3-7).

It is straightforward to generalize Eq. (3.24)-(3.25) to multi-component data. In *P* wave receiver function studies, for instance, the source signature is estimated from the vertical component. The vertical and radial receiver functions are retrieved by source deconvolution from vertical and radial components respectively. We demonstrate this using two events recorded by Hi-Climb array in Tibet (Fig. 3-8, 3-10). For comparison, different values for the water level δ are chosen from 0.01% to 1%. The (radial) receiver functions vary with the parameter δ in water level deconvolution (Fig. 3-9, 3-11). For the first earthquake, a comparable result is achieved with 0.01% as the water level (Fig. 3-8D and Fig. 3-9B). However, there is no acceptable result in the second case, even though a grid search for water level is performed (Fig. 3-11), which suggests that white noise assumption is not always proper in the real data.

3.4 Summary

In this chapter we propose a semi-automatic workflow for the preprocessing of large data sets. We mainly focus on the source normalization. Through multi-channel cross-correlation technique the polarity and relative delay time of traces are corrected, and noisy traces are discarded as well. The source signature is then estimated by principle component analysis and deconvolved by means of a Wiener inverse filter. Examples show that such a workflow is less subjective and more efficient than alternative schemes. Moreover, it is designed to be suitable for massive data processing, especially for dense array data.

References

Aki, K., and P. G. Richards (2002), *Quantitative seismology*, University Science Books.

Cao, Q., P. Wang, R. van der Hilst, M. de Hoop, and S.-H. Shim (2010), Imaging the upper mantle transition zone with a generalized Radon transform of SS precursors, *Phys. Earth Planet. Inter.*, *180*(1), 80-91.

Chen, C. W., D. Miller, H. Djikpesse, J. Haldorsen, and S. Rondenay (2010), Array-conditioned deconvolution of multiple-component teleseismic recordings, *Geophys. J. Inter.*, *182*(2), 967-976.

Dziewoński, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*(4), 297-356.

Dziewoński, A. M., T. A. Chou, and J. H. Woodhouse (1981), Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, *86*(B4), 2825-2852.

Ekström, G., M. Nettles, and A. M. Dziewoński (2012), The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. Inter.*, *200-201*(0), 1-9.

Haldorsen, J. B. U., D. E. Miller, and J. J. Walsh (1994), Multichannel Wiener deconvolution of vertical seismic profiles, *Geophysics*, *59*(10), 1500-1511.

Kennett, B., E. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Inter.*, *122*(1), 108-124.

Nábělek, J. L. (1984), Determination of earthquake source parameters from inversion of body waves, phD thesis, MIT.

Rondenay, S., M. G. Bostock, and K. M. Fischer (2005), Multichannel inversion of scattered teleseismic body waves: Practical considerations and applicability, *Geophys. Monograph-American Geophys. Union, 157.*

Rondenay, S. (2009), Upper mantle imaging with array recordings of converted and scattered teleseismic waves, *Surv. Geophys.*, *30*(4), 377-405.

Rost, S., and C. Thomas (2002), Array seismology: Methods and applications, *Rev. Geophys.*, *40*(3), 1008.

Sipkin, S. A. (1982), Estimation of earthquake source parameters by the inversion of waveform data: synthetic waveforms, *Phys. Earth Planet. Inter.*, *30*(2), 242-259.

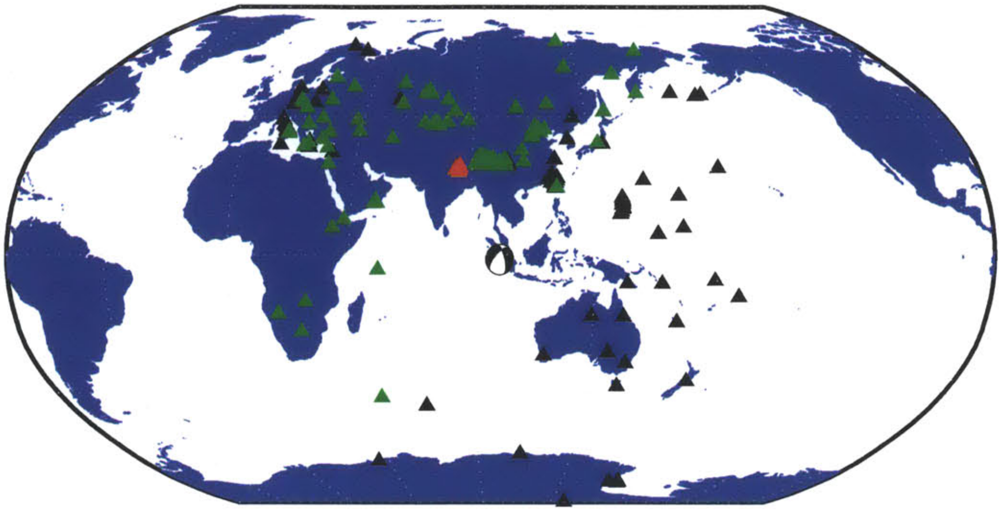
VanDecar, J., and R. Crosson (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seismol. Soc. Am.*, *80*(1), 150-159.

Wang, P., M. de Hoop, R. van der Hilst, P. Ma, and L. Tenorio (2006), Imaging of structure at and near the core mantle boundary using a generalized radon transform: 1. Construction of image gathers, *J. Geophys. Res.*, *111*.

Wiggins, R. W. C., and A. Ralph (1976), Source shape estimation and deconvolution of teleseismic bodywaves, *Geophys. J. R. Astr. Soc.*, *47*(1), 151-177.

Zheng, Y. (2007), Imaging upper mantle discontinuities and Earth's small scale heterogeneities, phD thesis, Univerisity of California, Santa Cruz.

A



B

Figure 3-1 Comparison between the polarity predicted from the CMT soltuion and real data measurements for (A) direct *P* arrivals and (B) *SH* arrivals. The beach balls depict the CMT solutions and the triangles are seismic stations. The green triangles are stations whose polarity picks agree with the CMT predictions. The red ones are those with the opposite polarity to the CMT predictions, and the black indicate unclear polarity picks from the data.

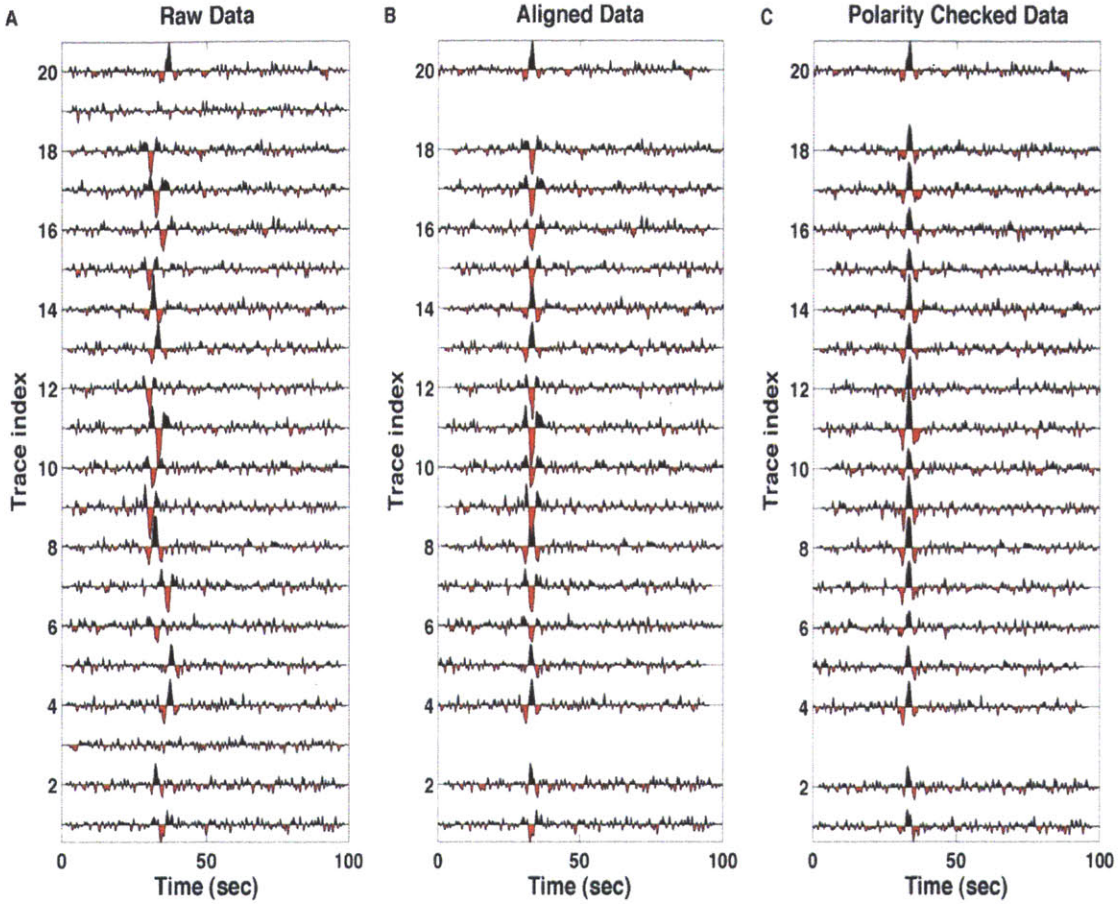


Figure 3-2 A synthetic example of data selection and polarity correction. (A) Raw data. The amplitudes of Ricker wavelets are modulated by a cosine function. The polarity and relative delay time are generated randomly. About 25% Gaussian noise is added to the traces. (B) Alignment of the traces by multi-channel cross-correlation. Trace 3 and 19 are removed due to low average correlation coefficients (threshold value is 0.5). (C) Aligned traces after polarity correction.

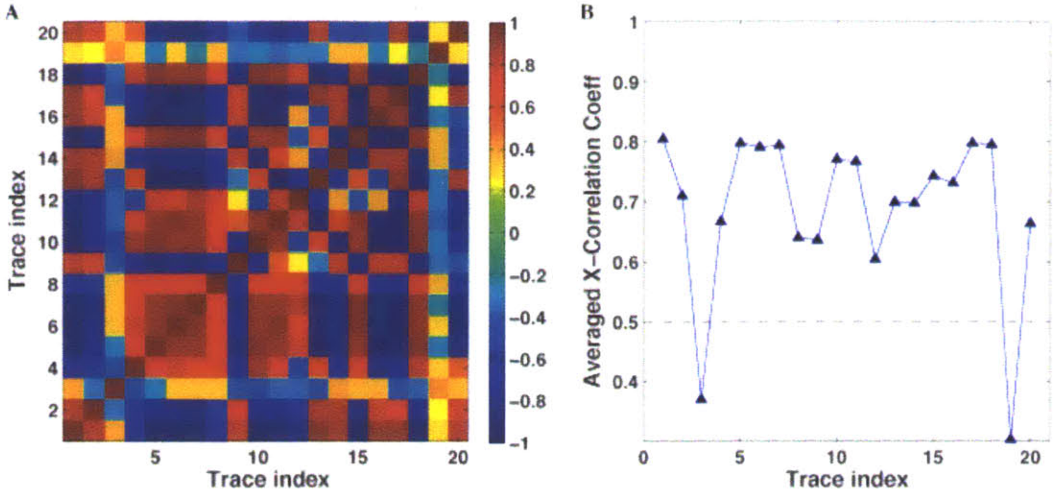


Figure 3-3 (A) The cross-correlation coefficient matrix of the synthetic data in Fig. 3-2A. (B) Average cross-correlation coefficients for the data in Fig. 3-2A. The threshold value for data selection is 0.5.

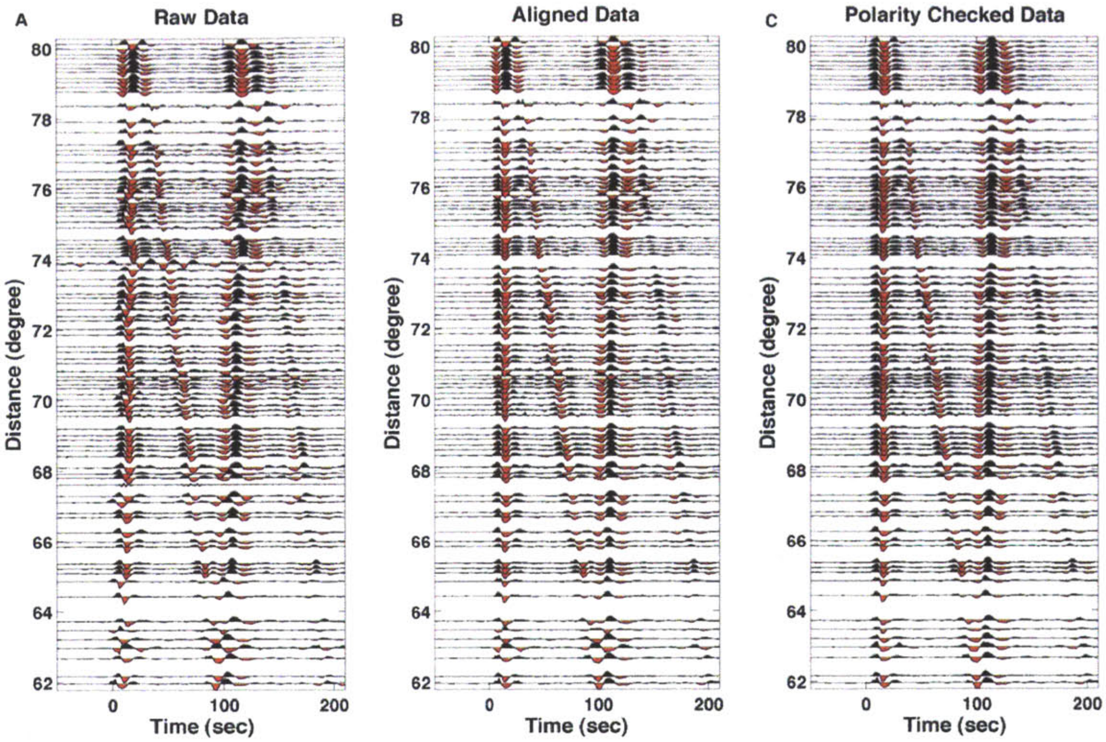


Figure 3-4 A real data example of data selection and polarity correction. (A) The raw data aligned with the direct *S* wave (predicted on ak135 model). The following phases are *ScS* and the depth phases of *S* and *ScS*. (B) Alignment of traces by multi-channel cross-correlation. The threshold value is chosen as 0.4 for the data selection. (C) Aligned traces after polarity correction.

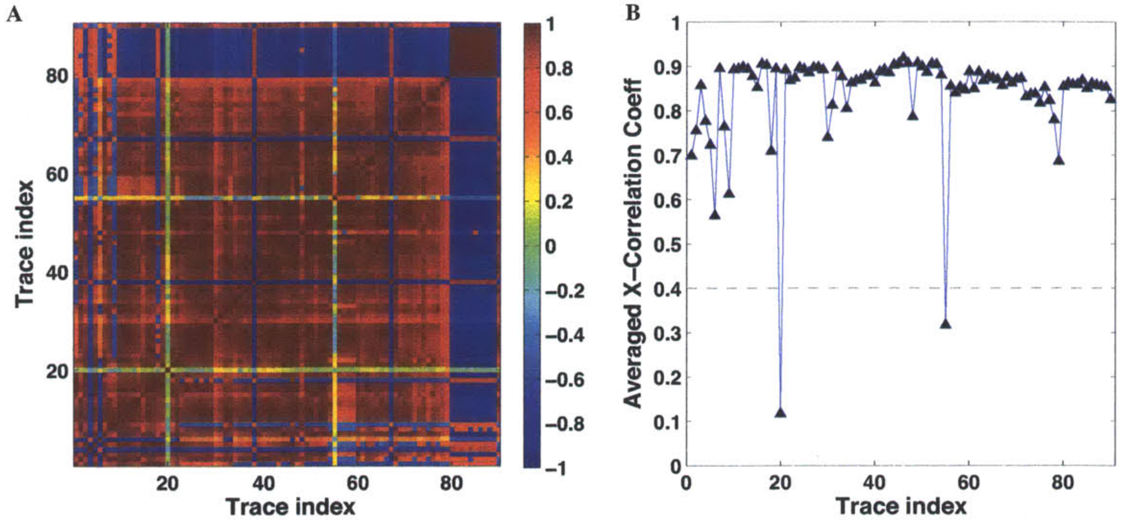


Figure 3-5 (A) The cross-correlation coefficient matrix for the data in Fig. 3-4A. (B) The corresponding average cross-correlation coefficients. The threshold value for the data selection is 0.4.

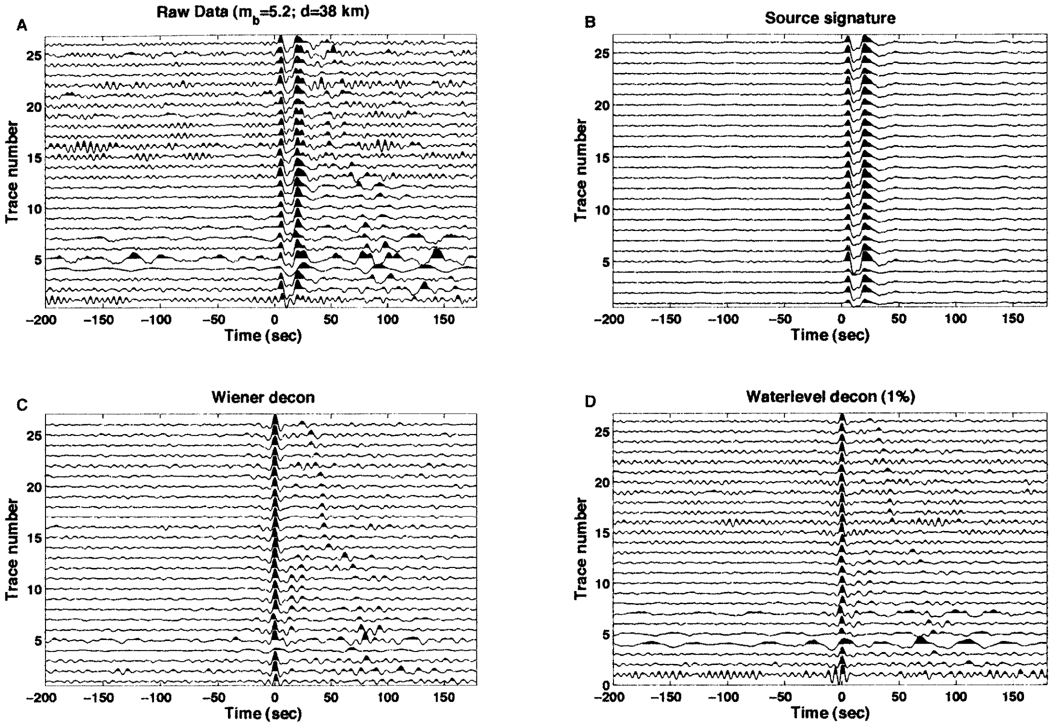


Figure 3-6 An example of source signature estimation and deconvolution. (A) The raw data aligned with the direct *S* wave. The source magnitude is (m_b) 5.2 and the focal depth is 38 km. (B) Source signature calculated by principle component analysis. Only the largest eigenvalue is included for the source estimation. (C) Traces after Wiener deconvolution. (D) Deconvolved traces by water level deconvolution. The water level is selected as 1% of the maximum value of the source amplitude spectrum.

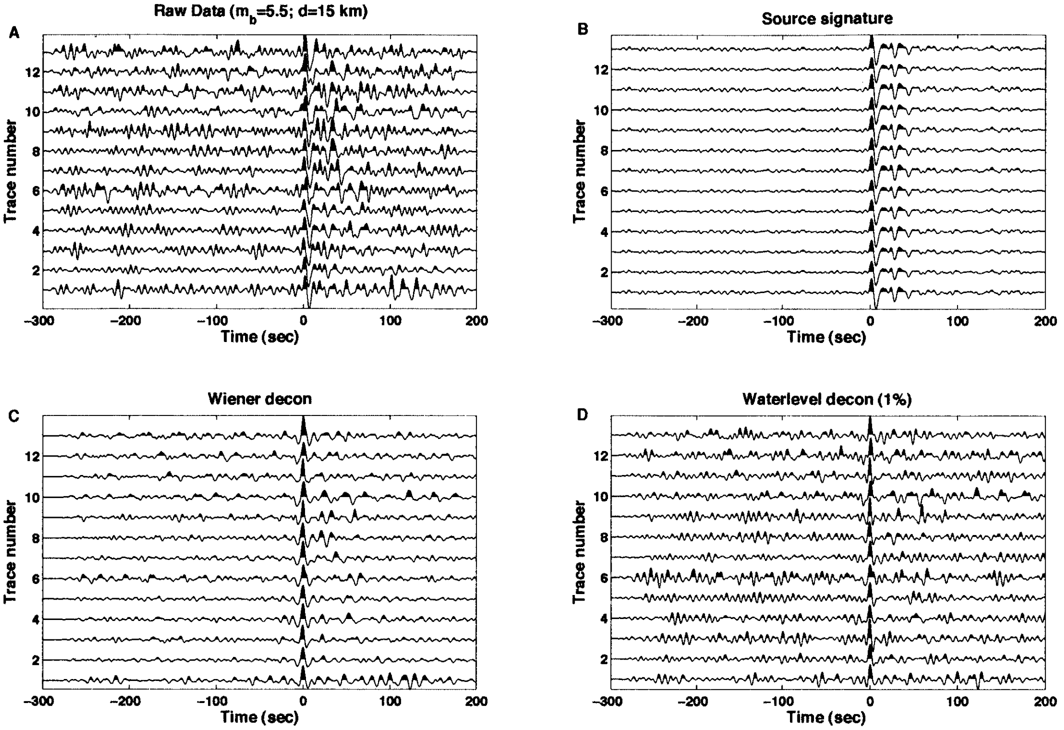


Figure 3-7 (A) The raw data aligned with the direct *S* wave. The source magnitude is (m_b) 5.5 and the focal depth is 15 km. (B) Source signature calculated by principle component analysis (only the largest eigenvalue is used). (C) Traces after Wiener deconvolution. (D) Traces after water level deconvolution. The water level is used as 1% of the maximum value of the source amplitude spectrum.

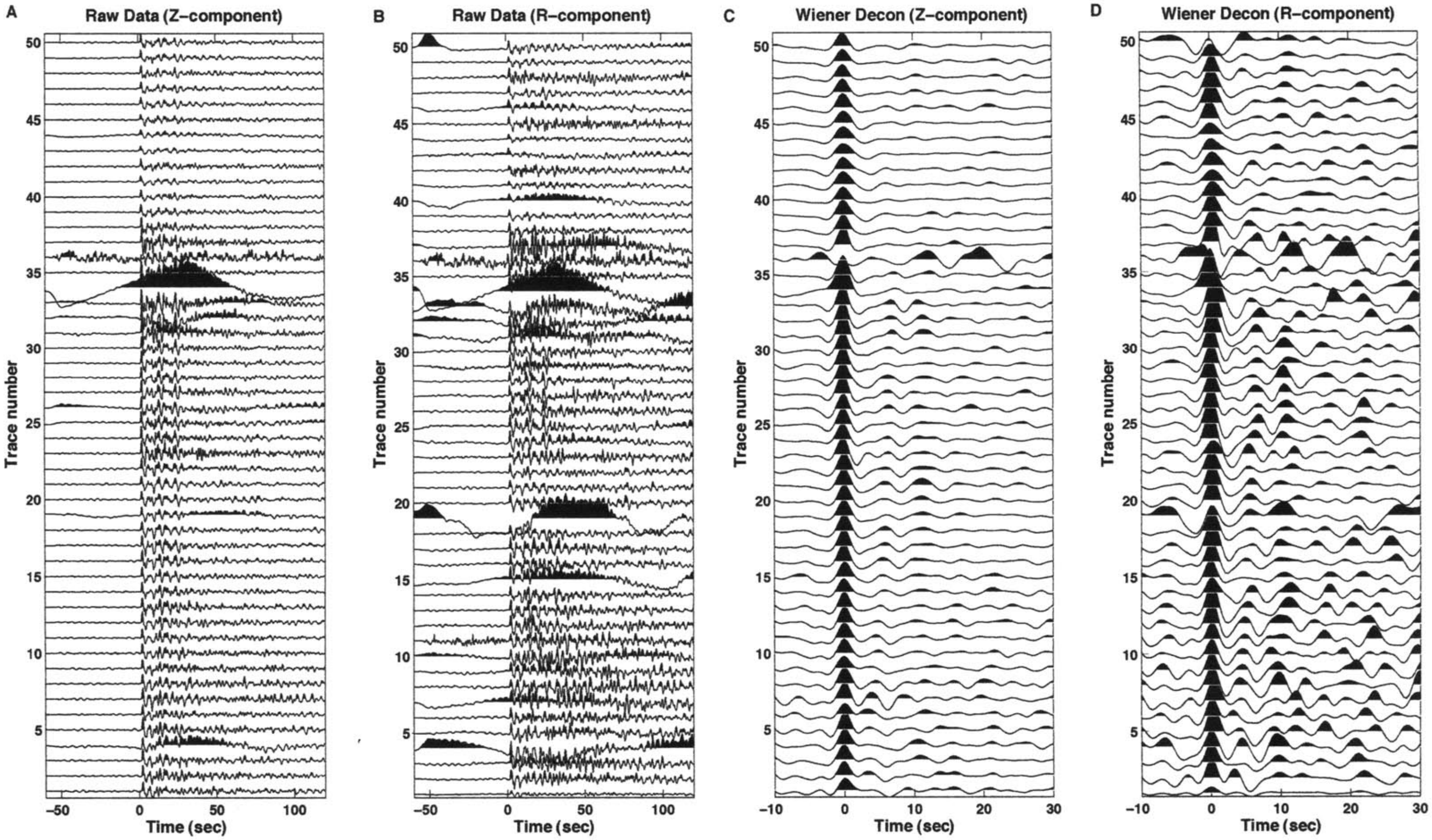


Figure 3-8 Demonstration of Wiener deconvolution for multi-component data. The raw data ((A) vertical and (B) radial component) are aligned with the direct *P* wave. The vertical and radial receiver functions (after Wiener deconvolution) are shown in (C) and (D), respectively.

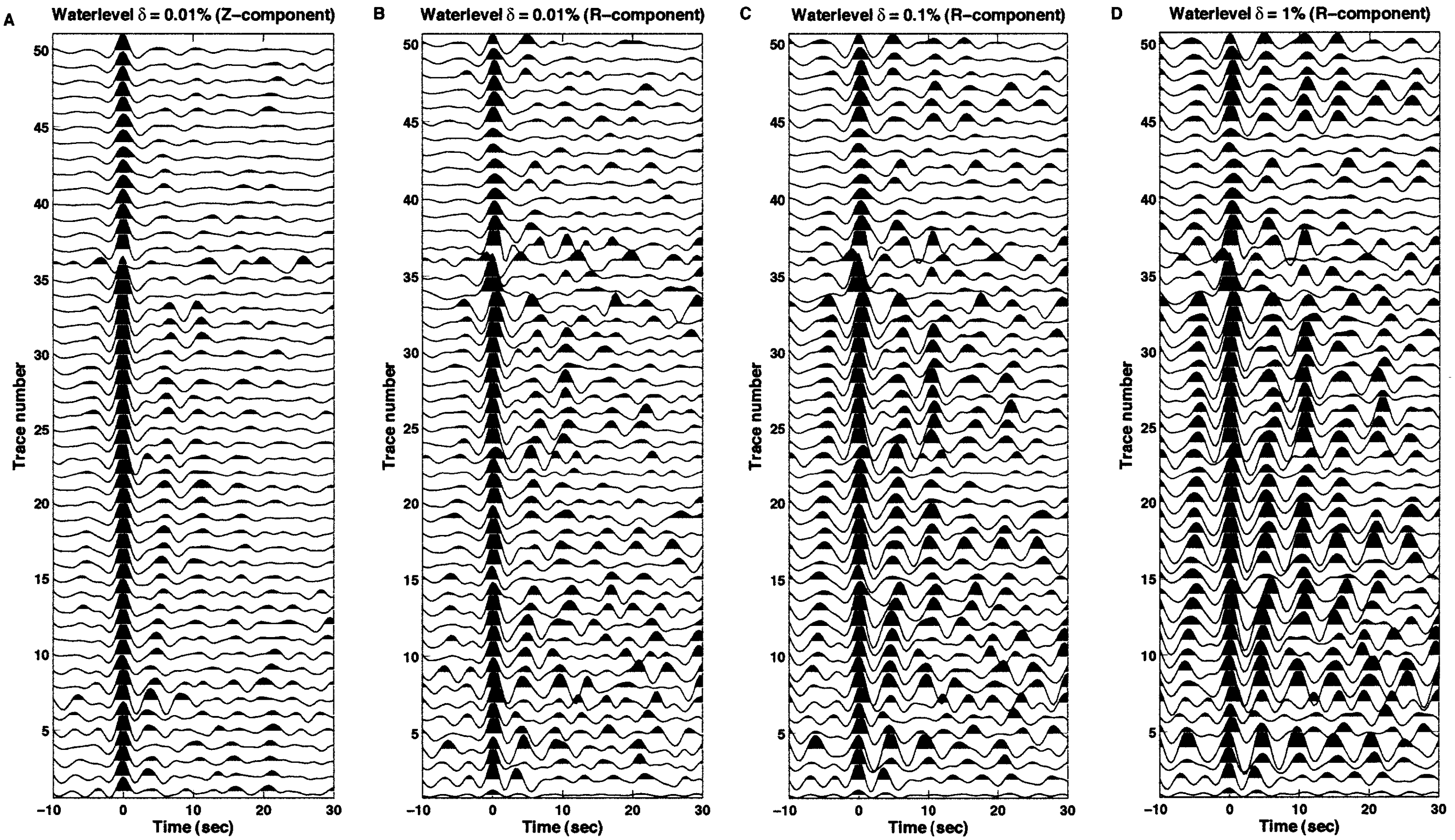


Figure 3-9 Influence of parameter δ on the water level deconvolution. The raw data are the same as in Fig. 3-8A and 3-8B. The water level changes from (A, B) 0.01% to (C) 0.1% and (D) 1% . Only the variation of the radial receiver functions (B-D) is shown here, since the vertical receiver functions (A) do not change dramatically.

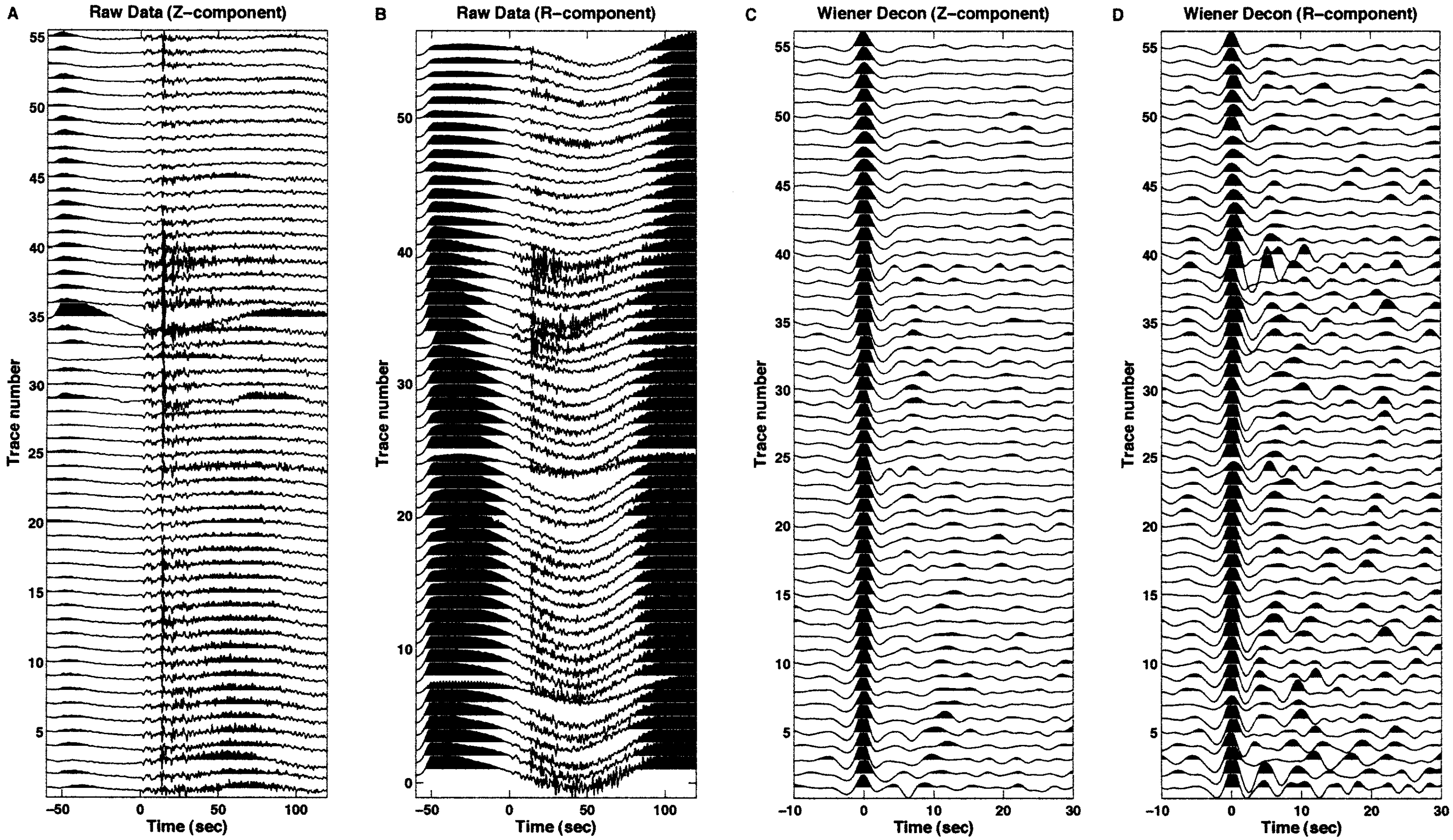


Figure 3-10 Another example for Wiener deconvolution in multi-component data. The raw data ((A) vertical and (B) radial component) are aligned with the direct *P* wave. The vertical and radial receiver functions (after Wiener deconvolution) are shown in (C) and (D), respectively.

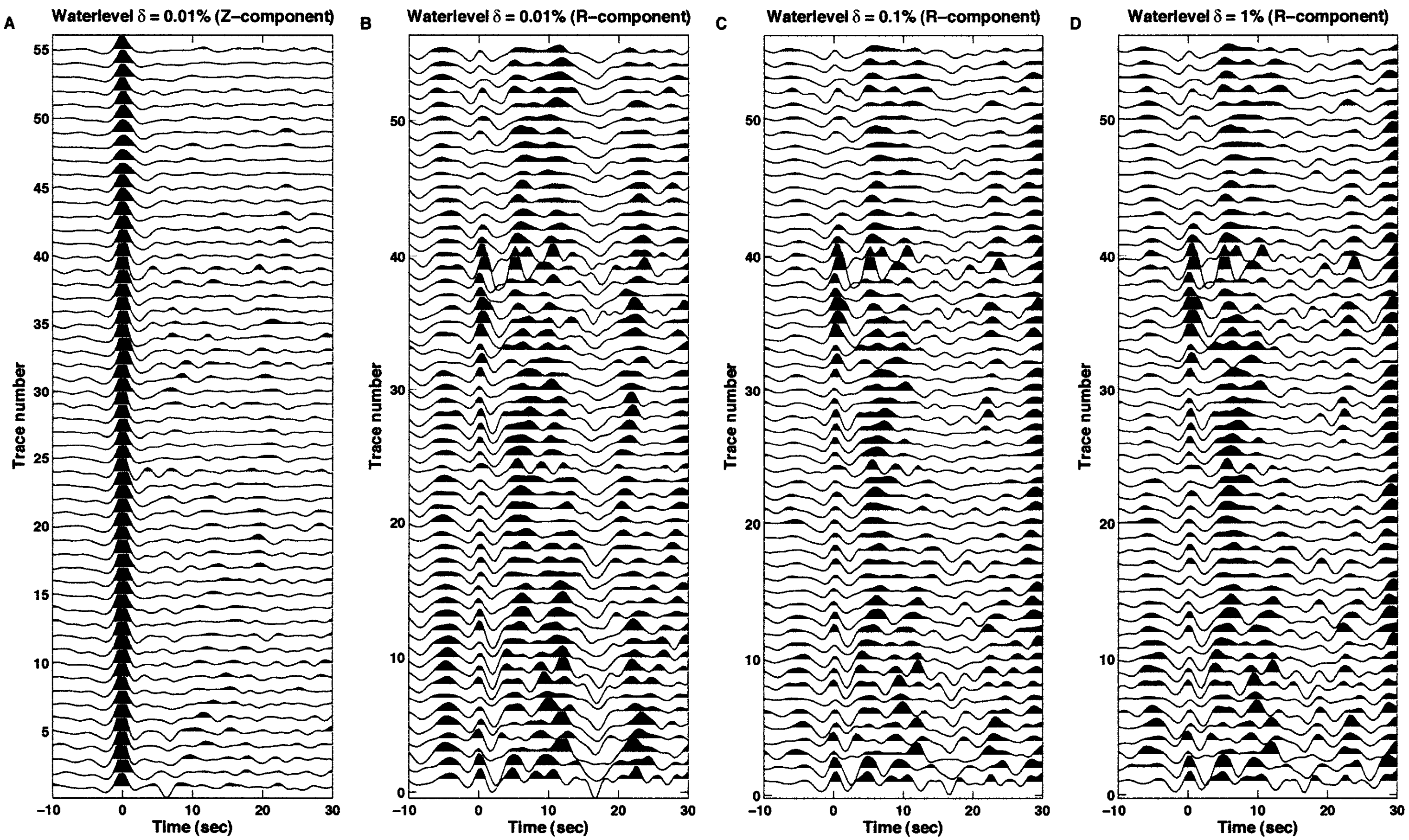


Figure 3-11 Influence of parameter δ on the water level deconvolution. The raw data are the same as shown in Fig. 3-10A and 3-10B. The water level changes from (A, B) 0.01% to (C) 0.1% and (D) 1%. Only the variation of the radial receiver functions (B-D) is shown here.

Chapter 4

Wavefield Interpolation in Curvelet Domain with Sparsity Promotion

In this chapter we apply curvelet-based wavefield recovery algorithm to teleseismic datasets. Irregularly and sparsely sampled data are interpolated to regular acquisition grids, which is crucial for wave equation based migration. The underdetermined problem with sparsity-promoting ℓ_1-norm regularization is solved by a spectral projected-gradient method, which is more efficient than the often used "cooling" method [*Figueiredo et al.*, 2007]. We investigate the effect of the undersampling ratio, from 0.5 to 0.15, on the teleseismic wavefield recovery. A simple mask function is applied to reduce the model space in severely undersampling scenarios. In the presence of noise, the Pareto curve can be explored by solving a series of so-called Lasso problems to find the optimal trade-off between the data fit and sparsity of the solution. Both synthetic and real data examples demonstrate that curvelet-based interpolation with sparsity promotion works remarkably well on teleseismic data.

4.1 Introduction

In seismic acquisition, the continuous wavefield is recorded by a finite number of receivers. To avoid aliasing and loss of information, seismic data should be sampled according to the Shannon-Nyquist criterion, that is, at least two sample points in each period or wavelength in the sense of Fourier. Due to the limitation of logistics and cost, however, this cannot always be achieved in practice, and the wavefield is often sparsely and insufficiently sampled along the spatial coordinates. On the other hand, many important and advanced techniques, for instance, denoising and wave equation migration, have stringent sampling requirement [*Liu and Sacchi*, 2004; *Xu et al.*, 2005]. So the recovery of missing traces is an important preprocessing step and the quality of reconstruction will impact the subsequent seismic data analysis.

There are two main classes of data interpolation algorithms: wave-equation methods and signal-processing methods. Wave-equation based methods simulate the wave propagation and can reconstruct the missing traces accurately, provided the information about wave speeds is available [e.g. *Ronen*, 1987; *Stolt*, 2002]. Signal processing methods usually exploit sparse transforms to map the recorded data into other domains to compress the redundant data and require no information about the medium property. Popular transforms include the Fourier transform [e.g. *Spitz*, 1991; *Sacchi et al.*, 1998; *Xu et al.*, 2005; *Zwartjes and Sacchi*, 2007], the Radon transform [e.g. *Kabir and Verschuur*, 1995; *Trad et al.*, 2002] and the curvelet transform [e.g. *Herrmann and Hennenfent*, 2008; *Naghizadeh and Sacchi*, 2010a].

Curvelets are optimally sparse representations of the wave propagator [*Candès and Demanet*, 2005]. The wave front (or the seismic data) can be sparsely represented by those curvelets aligned with the wave front (or the moveout curves). *Herrmann and Hennenfent* [2008] applied curvelet reconstruction with sparsity-promoting inversion (CRSI) for seismic data interpolation. The missing data can be efficiently reconstructed even for sub-Nyquist sampling. Such is especially attractive for teleseismic observations, in which the wavefield is severely undersampled. Dense seismographic arrays are commonly deployed with station spacing of 5-10 km, but high-resolution seismic imaging methods, for instance wave-equation migration, require even denser sampling (~ 2 km).

In this study, we focus on the recovery quality of curvelet interpolation for typical teleseismic scenarios. We first investigate the effect of sampling ratio on the quality of data reconstruction. A simple mask matrix in the curvelet domain is introduced for severely under-sampled scenarios. The Pareto curve (L-curve) is then exploited by solving so called Lasso problems in the presence of noise. The new method is demonstrated both on synthetic and real data examples.

4.2 Curvelet Transform

In the past decades, intense research has focused on the multi-scale (multi-resolution) analysis, e.g. wavelets, curvelets, and beamlets [e.g. *Daubechies*, 1990; *Candès and Donoho*, 2004; *Wu and Chen*, 2006]. Compared with wavelets, curvelets are anisotropic elements with support obeying the parabolic principle (width $\approx$ length^2), and are optimally sparse representations of the wave propagator [*Candès and Demanet*, 2005]. Analogous to the wavelets, the curvelet family can be constructed by scaling, rotating and translating of a mother function. Fig. 4-1 shows some examples of curvelets in both space domain (left) and frequency domain (right). In the space domain, the curvelet is smooth along the wave front direction but oscillatory and rapidly decay in the other direction. An intuitive explanation is that it behaves like a localized wave front with different orientation and scale. In the frequency domain, the curvelets tile the frequency plane into multi-scale and multi-angular wedges. The angular sampling doubles every other scale, so the curvelets become more anisotropic on finer scales, thus providing higher directional resolution. The mathematical basics and numerical implementation of the curvelet transform are fully developed and discussed in literatures [e.g. *Candès and Donoho*, 2004; *Candès et al.*, 2006; *Hennenfent et al.*, 2010]. Here we use the discrete curvelet transform developed by *Candès et al.* [2006], which is available at http://www.curvelet.org .

Due to the anisotropic, dual-domain localized and oscillatory nature, curvelets provide an optimally sparse representation of seismic data [e.g. *Candès and Demanet*, 2005; *Douma and de Hoop*, 2007]. Intuitively speaking, only a few curvelets that align with the wave front are required to represent one seismic event. We compare sparse reconstruction in different transform domains, shown in Fig. 4-2. The original shot gather is transformed to the Fourier, wavelet, and curvelet domains, and only the 1% coefficients with the largest amplitude are kept before inverse transformation back. It is clear that the curvelet transform gives the best result (Fig. 4-2D), while the Fourier and wavelet methods suffer from significant artifacts. The Fourier transform consists of an orthogonal but global basis, so it is not competent to approximate localized wave fronts by superposition of monochromatic plane waves. The "blob-like" wavelet frame is localized but isotropic, and is less efficient to fit "wavefront-like" seismic events. The performance in Fig. 4-2 is also supported by the decay rate of the Fourier, wavelet, and curvelet coefficients, plotted in Fig. 4-3. It is proved that for 2-D functions that are twice-differentiable and contain singularities along piecewise twice differentiable curves, the Fourier transform attains an asymptotic decay of the k-term approximation error of $O(k^{-1/2})$, while wavelet is $O(k^{-1})$ and curvelet is $O(k^{-2})$ [*Herrmann and Hennenfent*, 2008].

4. 3 Wavefield Reconstruction with Sparsity Promotion

In the seismic data reconstruction, we consider the observed data **d** as a subset of the desired interpolated data **m**. It can be represented as

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n} \ , \tag{4.1}$$

where **G** is a sampling operator and **n** is additive noise [*Naghizadeh and Sacchi*, 2010b]. Obviously Eq. (4.1) is an underdetermined problem, and some prior information must be introduced as a regularization term to constrain the solution. We follow the strategy proposed by *Herrmann and Hennenfent* [2008] to reconstruct the wavefield with sparsity promotion in curvelet domain. In this study, a spectral projected-gradient algorithm (SPG ℓ_1) [*Figueiredo et al.*, 2007; *van den Berg and Friedlander*, 2008] is adopted as the one-norm solver instead of iterative soft thresholding with cooling (ISTc) method [*Daubechies et al.*, 2004; *Herrmann and Hennenfent*, 2008].

The curvelet analysis operator **C** decomposes desired data **m** to curvelet coefficients **x** as:

$$\mathbf{x} = \mathbf{C}\mathbf{m} \ , \tag{4.2}$$

and **m** can be synthesized as

$$\mathbf{m} = \mathbf{C}^T \mathbf{x} \ , \tag{4.3}$$

where $\mathbf{C}^\mathrm{T}$ is the adjoint operator of curvelet transform, which is equivalent to the inverse of **C** since the discrete curvelet frame is a tight frame [*Candès and Donoho*, 2004]. Plugging Eq. (4.3) into Eq. (4.1), the curvelet coefficients **x** of the desired data can be found by minimizing the following cost function:

$$J = \frac{1}{2}\|\mathbf{d} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1 \tag{4.4}$$

where $\mathbf{A}=\mathbf{G}\mathbf{C}^\mathrm{T}$, and the positive parameter λ is the Lagrange multiplier indicating the trade-off between the data misfit and one-norm of the solution. The desired interpolated data **m** then can be estimated by Eq. (4.3). The best damping parameter λ in Eq. (4.4) is generally unknown. The Pareto curve (or L-curve) is usually exploited to find the optimal damping parameter. *Van den Berg and Friedlander* [2008] proved that this curve is convex and continuously differentiable, and gave an explicit relationship to two other closely related optimization problems stated in the following, basis pursuit denoise (BPDN) and Lasso. Basis pursuit denoise problem is described as

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{x}\|_1, \quad \text{subject to } \|\mathbf{Ax}-\mathbf{d}\|_2 \leq \sigma \ . \tag{4.5}$$

Lasso problem is

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{Ax}-\mathbf{d}\|_2, \quad \text{subject to } \|\mathbf{x}\|_1 \leq \tau \ . \tag{4.6}$$

These three problems are equivalent in some sense, and for some appropriate parameters of λ, τ, and σ, the solutions of Eq. (4.4), (4.5) and (4.6) coincide. In practice, Eq. (4.5) is preferred if the noise level σ in the data is known, and Eq. (4.6) can be used if the upper bound of one-norm of the curvelets can be estimated.

Since curvelets are direction selective, one can choose a weighting function, e.g. using a vector of Lagrange multipliers rather than a scalar in Eq. (4.4), to penalize nearly vertical curvelets that are not likely true in the seismic data gather in *t-x* domain. However, subjective selection of multiple parameters would complicate the inversion process. A simple way to mute the vertical curvelets is to add a mask matrix **M** as $\mathbf{A}=\mathbf{G}\mathbf{C}^{\mathrm{T}}\mathbf{M}$ in the curvelet domain [*Herrmann and Hennenfent*, 2008]. One can taper out nearly vertical curvelets only on fine scales, as illustrated in Fig. 5-4. The dip (or slope) of the mask function is analogous to that of a dip filter in the *f-k* domain. For simplicity but without loss of generality, we consider for a band-limited shot gather in 2D case the horizontal wavenumber can be expressed as

$$k_x = k\cos\theta = \frac{2\pi f}{v}\cos\theta \ , \tag{4.7}$$

where v is the medium velocity; f is the frequency and θ is the horizontal direction angle of the wavenumber vector. Assuming no evanescent waves recorded in the data, which means $|\cos\theta| \leq 1$, the maximum slope in the *f-k* domain is

$$s_{\max} = \left(\frac{k_x}{f}\right)_{\max} = \frac{2\pi}{v_{\min}}, \tag{4.8}$$

where $v_{\min}$ is the minimum medium velocity or apparent velocity. The mask matrix **M** can be designed according to Eq. (4.8). In the following section, a mask matrix is applied in the wavefield interpolation for severely undersampled datasets, which reduces the number of unknowns and improves the data recovery quality.

4.4 Numerical Examples

Since the compressive sampling recovery depends on factors such as acquisition footprint and signal characteristics [*Herrmann and Hennenfent*, 2008], it is difficult to predict the curvelet interpolation quality accurately. In this section, we investigate the effects of undersampling ratio and random noise on typical synthetic teleseismic data, which can give guidelines for the real data applications.

4.4.1 Noise-free Synthetic Examples

We first create a synthetic teleseismic section, shown in Fig. 4-5A. The first arrival is approximately a plane wave, and later arrivals are converted waves and multiples. Notice that polarity changes, conflicting dips and caustics can be observed in the later events. The spectrum is illustrated in Fig. 4-6A, which is band-limited in the wavenumber-frequency domain. This original dataset $\mathbf{m}_0$ is used as a ground-truth solution to evaluate the quality of curvelet interpolation. The reconstruction quality is evaluated in decibels (dB) by the measure [*Hennenfent and Herrmann*, 2008]

$$Q = -20\log_{10}\left(\frac{\|\mathbf{m}_0 - \hat{\mathbf{m}}\|_2}{\|\mathbf{m}_0\|_2}\right), \tag{4.9}$$

where $\hat{\mathbf{m}}$ is the reconstructed data.

In the first experiment, 50% of original traces are randomly removed, as in Fig. 4-5B (the missing traces are replaced by zero traces). The *f-k* amplitude spectrum after the trace removal is plotted in Fig. 4-6B. In contrast to regularly undersampled data, there is no clear wraparound (replica) alias but random noise in the *f-k* spectrum, which is characterized by the random sampling function [*Oppenheim et al.*, 1999]. Since it is noise free, we solve the basis pursuit denoise problem in the curvelet domain with $\sigma \to 0$ (here $\sigma = 0.001\|d\|_2$ is used) in Eq. (4.5), and the interpolated data $\mathbf{m}$ is then obtained by applying the adjoint curvelet transform in Eq. (4.3). The recovered data is shown in Fig. 4-5C, and the corresponding *f-k* amplitude spectrum is depicted in Fig. 4-6C. The difference between the recovered data and the ground-truth solution is illustrated in Fig. 4-5D. The recovery quality Q measured by Eq. (4.9) is 34.60 dB. It is clear that wavefield interpolation works excellently and that the missing traces are recovered almost perfectly.

In the next experiment, we remove randomly 85% traces in the original dataset, plotted in Fig. 4-7A (amplitude spectrum in Fig. 4-8A). This scenario is typical in the teleseismic studies. The wavefield recorded by an array with 10 km spacing, for instance, needs to be interpolated into 2 km spacing grids to meet the requirement of the wave-equation migration. In this example, we examine the performance of mask matrix $\mathbf{M}$ in operator $\mathbf{A}$ in Eq. (4.5). The basis pursuit denoise problem is first solved with the same parameters as above. The interpolated data is plotted in Fig. 4-7B with the amplitude spectrum in Fig. 4-8B. Major features including caustics are recovered and the recovery quality is 11.0 dB. Some artificial localized events and oscillations (e.g. around the first arrival) can be observed in Fig. 4-7B, especially in the large gap fillings. Spectral leakage is noticeable after the curvelet interpolation (Fig. 4-8B). For a comparison, a mask matrix with an apparent velocity 4 km/s (roughly the minimum S wave velocity here) is applied in the data interpolation, and the recovery quality increases to 18.37 dB, shown in Fig. 4-7C (Fig. 4-8C for the spectrum). Compared with Fig. 4-7B, the artifacts in the interpolation are efficiently suppressed. Spectral leakage (Fig. 4-8C) is mitigated as well. The recovery with a simple mask function is remarkable given the sparsity of spatial samplings. One can choose a larger apparent velocity to squeeze the model space if more prior information about the medium and wavefield is known.

4.4.2 Synthetic Examples with Noise

In the reality, there is always noise in the recorded data. In this section, the effect of random noise on the curvelet interpolation is investigated. For simplicity only white Gaussian noise is used in the synthetic experiments. The Pareto curve is explored to find the best trade-off parameters σ and τ in Eq. (4.5) and (4.6) respectively. Here we sample the Pareto curve by a series of Lasso problems in Eq. (4.6). The upper bound τ of one-norm of curvelets is estimated in the following way. For the recorded data in *t-x* domain, in each row (fix the time) we replace each missing sample with the mean value of its horizontal neighbors. Such a process is conducted iteratively to fill the gaps from both sides. This simple method gives a low quality interpolation, but provides a good reference for the upper bound of τ. The norm of curvelets from the neighborhood interpolation above is similar to that of the ground true solution, and the difference is marginal.

Similar to the noise-free examples above, we provide two scenarios with 50% and 85% missing data, respectively. In each case, 30% white Gaussian noise is added to the recorded data. We first remove 50% traces and add random noise, depicted in Fig. 4-9A (spectrum in Fig. 4-10A). A series of parameter τ are used to sample the Pareto curve, plotted in Fig. 4-11A. The recovery quality varies with respect to the parameter τ as well (Fig. 4-11B). To understand better the effect of parameter τ, three points on the Pareto curve are selected (colored circles in Fig. 4-11), and the corresponding interpolated results are illustrated in Fig. 4-9(B-D), as well as the amplitude spectra in Fig. 4-10(B-D).

It is noticed that with a small τ, the first arrival and some of the following converted phases are recovered (Fig. 4-9B). More subtle features as multiples can be observed as τ increases (Fig. 4-9C). Above some certain point, however, there is no more improvement in the data recovery. Fine and nearly vertical curvelets are introduced to over-fit the noisy data (Fig. 4-9D). The recovery quality even decreases though the residual of data fitting decreases as well (Fig. 4-11). The best recovery quality is 16.55 dB in this example (the red circle in Fig. 4-11). After interpolation, one can further denoise the interpolated data (remove the artificial curvelets) by a threshold algorithm [e.g. *Starck et al.*, 2002; *Ali et al.*, 2010], which is beyond the scope of this paper.

Next, we remove 85% of traces and add 30% noise to the data, shown in Fig. 4-12A with the amplitude spectrum in Fig. 4-13A. Similar to the noise-free case, the influence of the mask matrix **M** is probed. The Pareto curves with and without the mask matrix are plotted as red and blue lines in Fig. 4-14A respectively. The recovery quality curves are in Fig. 4-14B. Here the apparent velocity is 4 km/s. The best parameter τ (marked as circles in Fig. 4-14) yields 9.09 dB recovery quality in the absence of the mask matrix, and it increases to 12.54 dB with the mask matrix. The corresponding interpolated data are shown in Fig. 4-12B and 4-12C respectively (for spectra, see Fig. 4-13(B-C)). After applying a simple mask in the curvelet domain, some fine scale artifacts are notably suppressed in the interpolated result. Compared with the noise-free case (Fig. 4-7C), some subtle events in Fig. 4-12C, e.g. multiples after 50 seconds, are not observable. The recovery for the direct wave and primary converted waves is acceptable, though there are artificial oscillations around the first arrival.

4.4.3 A Real Data Example

In this section, we apply the algorithm in a real dataset. The vertical component of the raw data for one earthquake, recorded by a broadband seismographic array in Tibet, is aligned with the first *P*-arrival, shown in Fig. 4-15A. There are 58 stations with spacing varying from 2 km to 26 km. The temporal frequency is 0.05 Hz to 1 Hz. We try to interpolate the randomly distributed data into regularly sampled traces with spatial sampling of 2 km, which means to recover 80% missing traces. As in the synthetic examples, a Pareto curve is explored by solving a series of Lasso problems, depicted in Fig. 4-15B. The apparent velocity for the mask matrix is chosen as 6 km/s. The best parameter τ is selected along the Pareto curve as a red circle (Fig. 4-15B). The corresponding interpolated result is shown in Fig. 4-15C. The amplitudes along the horizontal and slanting (e.g. around 120 second) events vary smoothly. There are no significant artifacts in the interpolated data. A zoomed-in comparison with the raw traces is plotted in Fig. 4-15D. Major events agree well, while some subtle wiggles in the raw data are considered as uncorrelated noise and dimmed out in the recovery.

4.5 Summary

In this chapter, we focus on the teleseismic data interpolation via curvelet domain with sparsity promotion. We solve the underdetermined problem with a spectral projected-gradient method. A simple mask function is applied for interpolations in severely undersampled scenarios. For real data applications with noise, we explore the Pareto curve by solving Lasso problems to find the optimal trade-off between the data fit and sparsity of the solution. For typical teleseismic acquisition geometry with 50% to 85% missing traces, both synthetic and real data demonstrate that the curvelet-based interpolation works remarkably well.

References

Ali, A. D., P. D. Swami, and J. Singhai (2010), Modified curvelet thresholding algorithm for image denoising, *J. Comput. Sci.*, *6*(1), 18.

Candès, E., and D. Donoho (2004), New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities, *Comm. Pure Appl. Math.*, *57*(2), 219-266.

Candès, E., and L. Demanet (2005), The curvelet representation of wave propagators is optimally sparse, *Comm. Pure Appl. Math.*, *58*(11), 1472-1528.

Candès, E., L. Demanet, D. Donoho, and L. Ying (2006), Fast discrete curvelet transforms, *Multiscale Model. Sim.*, *5*(3), 861-899.

Daubechies, I. (1990), The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inf. Theory*, *36*(5), 961-1005.

Daubechies, I., M. Defrise, and C. De Mol (2004), An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Comm. Pure Appl. Math.*, *57*(11), 1413-1457.

Douma, H., and M. V. de Hoop (2007), Leading-order seismic imaging using curvelets, *Geophysics*, *72*(6), S231.

Figueiredo, M. r. A., R. D. Nowak, and S. J. Wright (2007), Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems, *IEEE J. Sel. Topics Signal Process.*, *1*(4), 586-597.

Hennenfent, G., and F. J. Herrmann (2008), Simply denoise: Wavefield reconstruction via jittered undersampling, *Geophysics*, *73*(3), V19-V28.

Hennenfent, G., L. Fenelon, and F. J. Herrmann (2010), Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach, *Geophysics*, *75*(6), Wb203-Wb210.

Herrmann, F. J., and G. Hennenfent (2008), Non-parametric seismic data recovery with curvelet frames, *Geophys. J. Inter.*, *173*(1), 233-248.

Kabir, M. N., and D. Verschuur (1995), Restoration of missing offsets by parabolic Radon transform, *Geophys. Prospect.*, *43*(3), 347-368.

Liu, B., and M. D. Sacchi (2004), Minimum weighted norm interpolation of seismic records, *Geophysics*, *69*(6), 1560-1568.

Naghizadeh, M., and M. D. Sacchi (2010a), Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data, *Geophysics*, *75*(6), WB189-WB202.

Naghizadeh, M., and M. D. Sacchi (2010b), On sampling functions and Fourier reconstruction methods, *Geophysics*, *75*(6), Wb137-Wb151.

Oppenheim, A. V., R. W. Schafer, and J. R. Buck (1999), *Discrete-time signal processing*, Prentice Hall Upper Saddle River.

Ronen, J. (1987), Wave-equation trace interpolation, *Geophysics*, *52*(7), 973-984.

Sacchi, M. D., T. J. Ulrych, and C. J. Walker (1998), Interpolation and extrapolation using a high-resolution discrete Fourier transform, *IEEE Trans. Signal Process.*, *46*(1), 31-38.

Spitz, S. (1991), Seismic trace interpolation in the *F-X* domain, *Geophysics*, *56*(6), 785-794.

Starck, J.-L., E. Candès, and D. L. Donoho (2002), The curvelet transform for image denoising, *IEEE Trans. Image Process.*, *11*(6), 670-684.

Stolt, R. H. (2002), Seismic data mapping and reconstruction, *Geophysics*, *67*(3), 890-908.

Trad, D. O., T. J. Ulrych, and M. D. Sacchi (2002), Accurate interpolation with high-resolution time-variant Radon transforms, *Geophysics*, *67*(2), 644-656.

van den Berg, E., and M. P. Friedlander (2008), Probing the Pareto frontier for basis pursuit solutions, *SIAM J. Sci. Comput.*, *31*(2), 890-912.

Wu, R. S., and L. Chen (2006), Directional illumination analysis using beamlet decomposition and propagation, *Geophysics*, *71*(4), S147.

Xu, S., Y. Zhang, D. Pham, and G. Lambare (2005), Antileakage Fourier transform for seismic data regularization, *Geophysics*, *70*(4), V87.

Zwartjes, P. M., and M. D. Sacchi (2007), Fourier reconstruction of nonuniformly sampled, aliased seismic data, *Geophysics*, *72*(1), V21-V32.

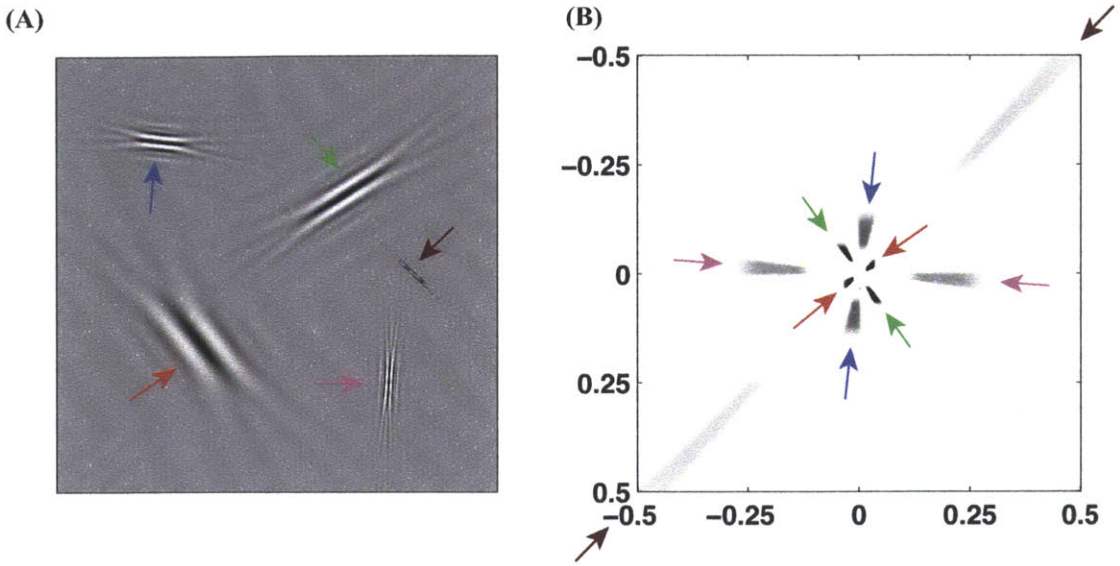


Figure 4-1 Examples of curvelets in the spatial domain and frequency domain. (A) Five curvelets in the spatial domain with different scales and orientations. (B) The corresponding wedges in the frequency domain. Each pair of the opposing wedges, denoted by arrows in different colors, represents a curvelet in the spatial domain. Note the finer scale curvelets occupy outer tiles in the frequency plane, and more anisotropic. This figure is modified from *Herrmann and Hennenfent* [2008].

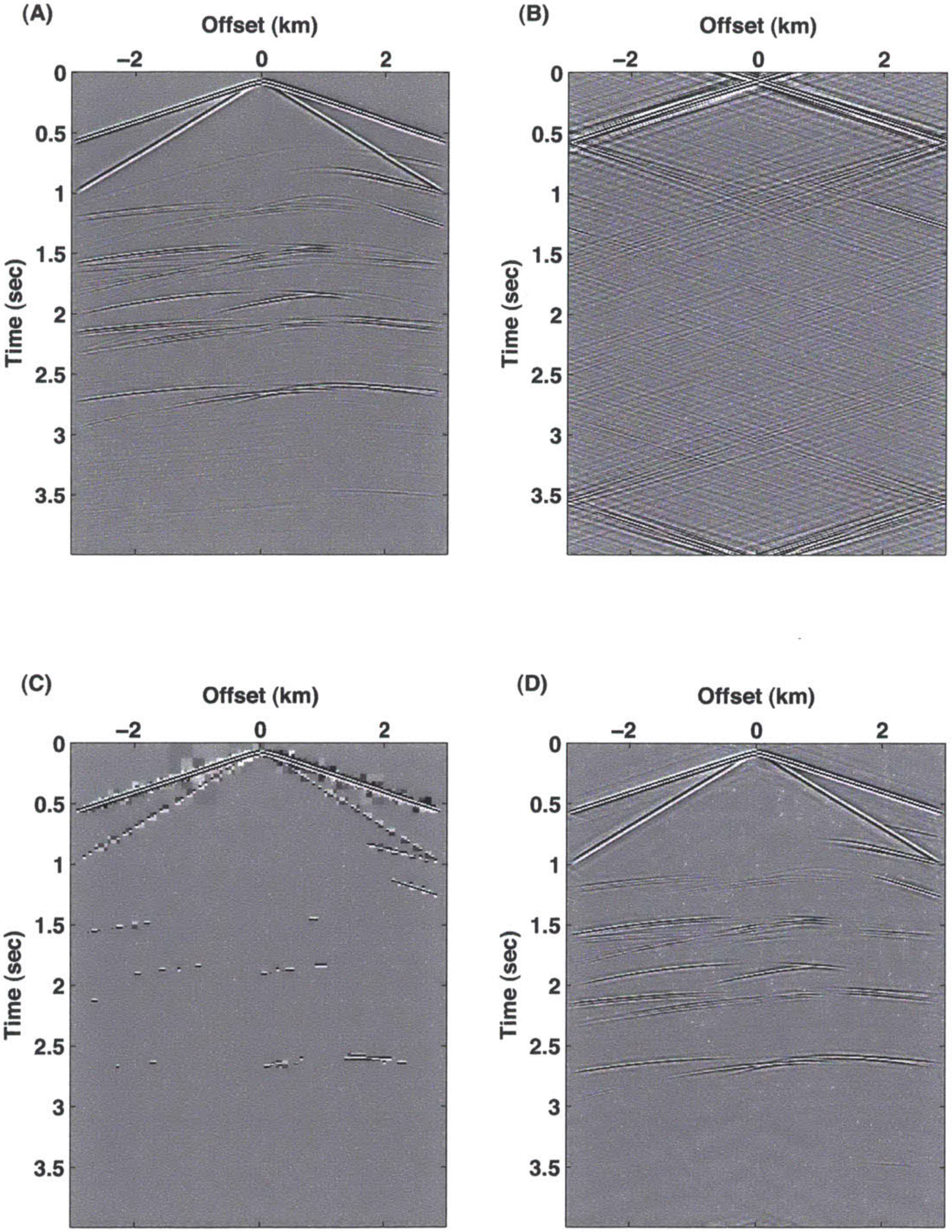


Figure 4-2 Sparse reconstruction in different domains. (A) The original shot gather generated by a finite difference method. We reconstruct the data by keeping only the 1% of the largest amplitude coefficients in the (B) Fourier domain, (C) wavelet domain with Daubechies frames, and (D) curvelet domain.

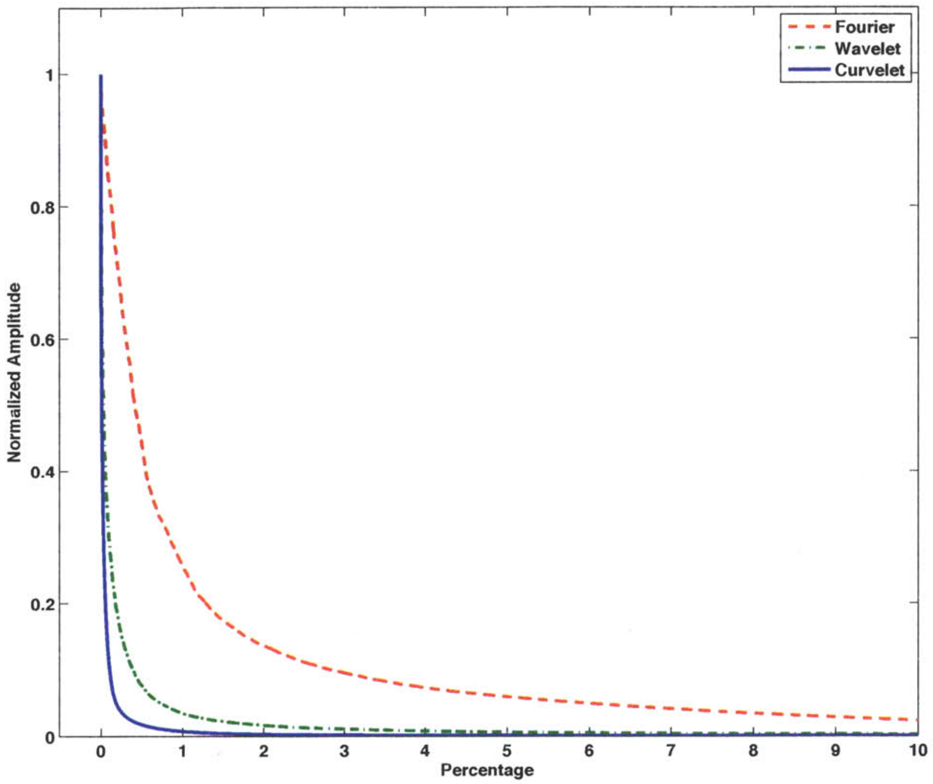


Figure 4-3 The decay rate of the coefficients in the Fourier domain (red dashed line), the wavelet domain (green dashed line), and the curvelet domain (blue solid line). The coefficients are calculated based on the synthetic data in Fig. 4-2A.

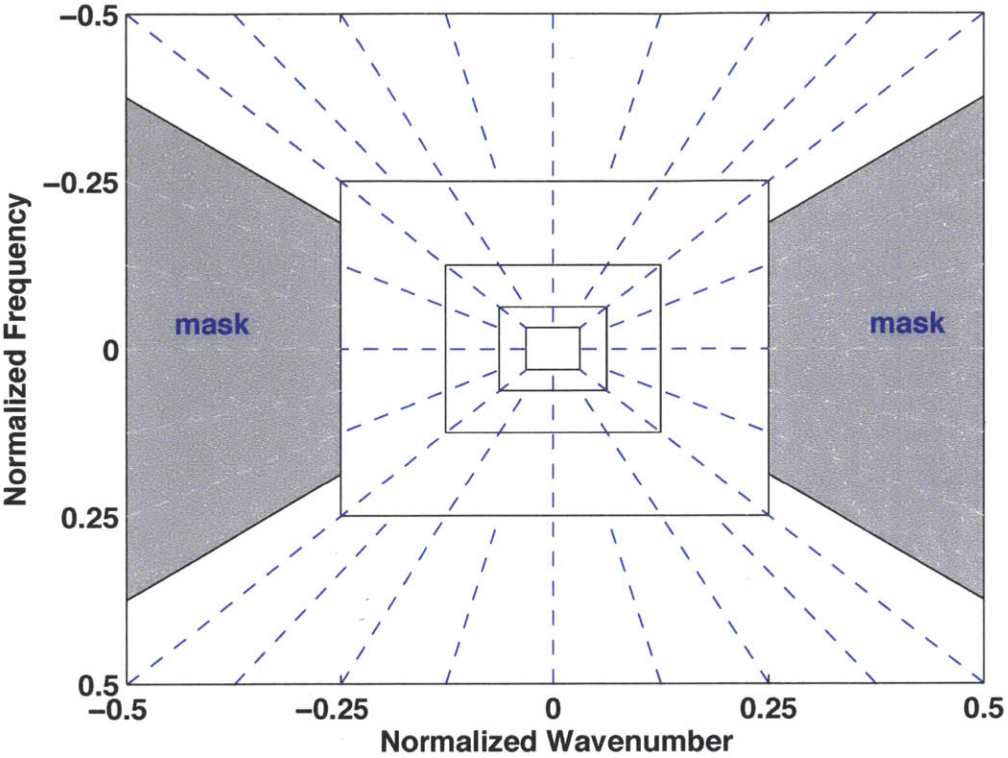


Figure 4-4 Illustration of angular discrimination in the curvelet domain. Due to the mutual coherence between the nearly vertical-oriented curvelets and the sampling gaps, the curvelets within the mask (gray regions) can be tapered out in the wavefield interpolation.

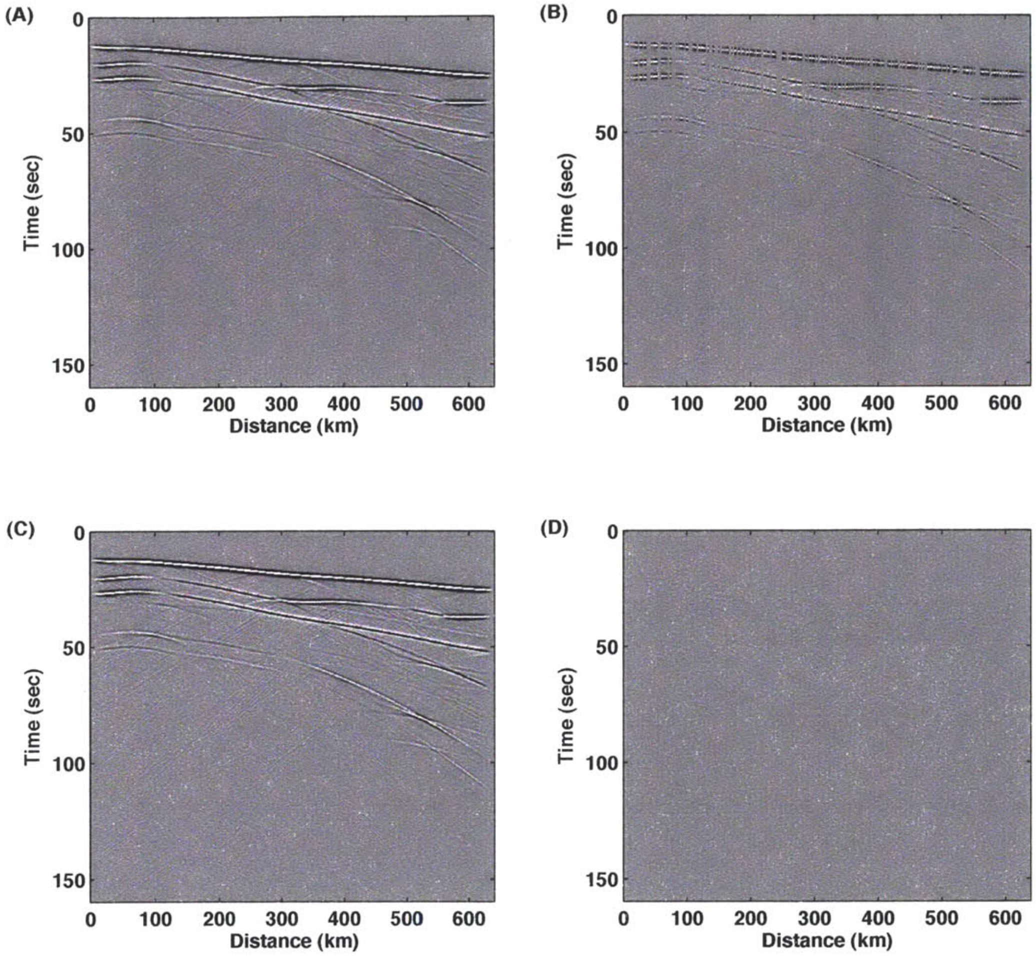


Figure 4-5 A synthetic example of curvelet interpolation. (A) The original data. (B) Data after removing 50% of the original traces randomly. (C) Recovered data by curvelet interpolation with sparsity promotion. (D) The difference between the interpolated data and the original data. All figures are displayed on the same color scale, clipped by ±5% of the maximum value of the original data.

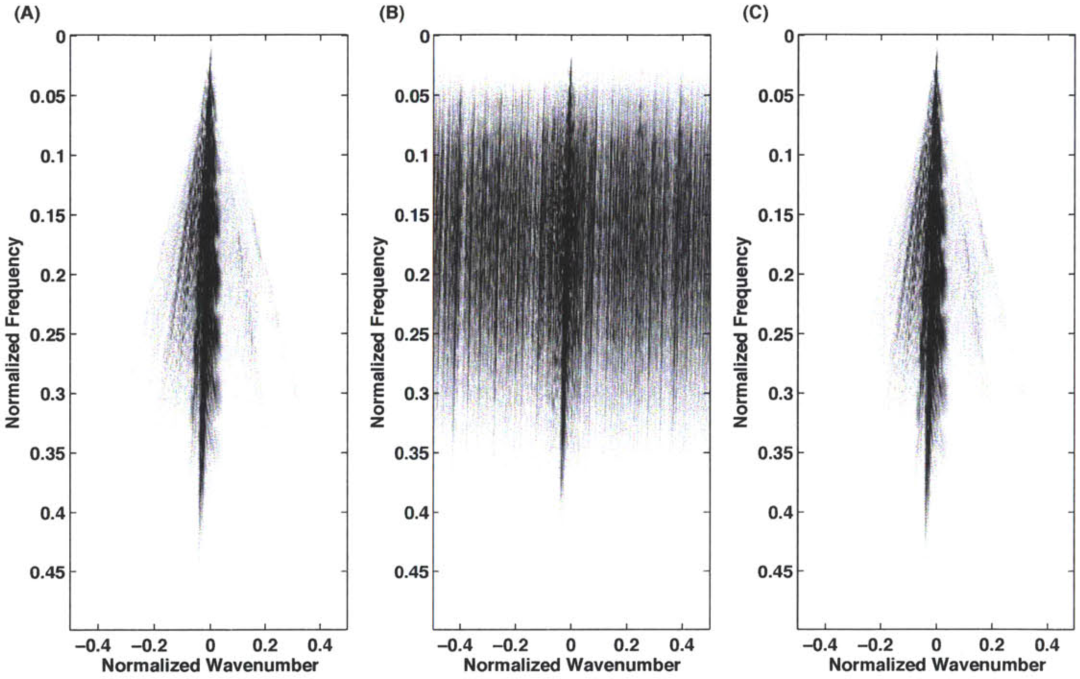


Figure 4-6 The *f-k* spectra of the data in Fig. 4-5 (A-C), respectively.

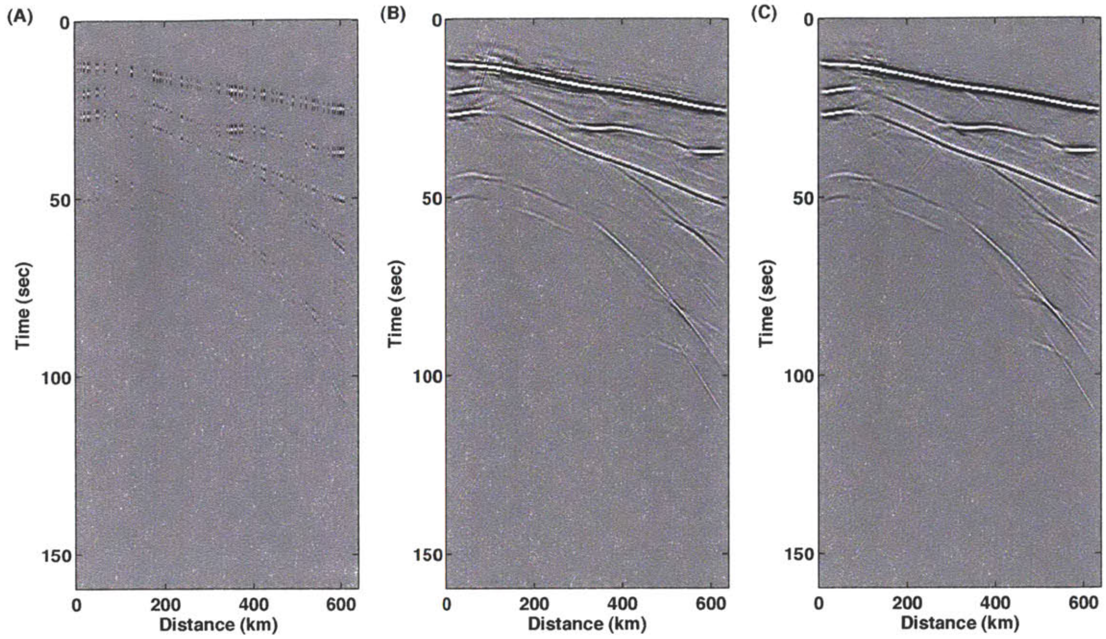


Figure 4-7 The curvelet interpolation in a severely undersampled case. (A) Data with 85% missing traces. The original data is the same as in Fig. 4-5A. (B) Interpolated data without a mask matrix. (C) Interpolated data with an angular mask matrix, and the apparent velocity is chosen as 4 km/s.

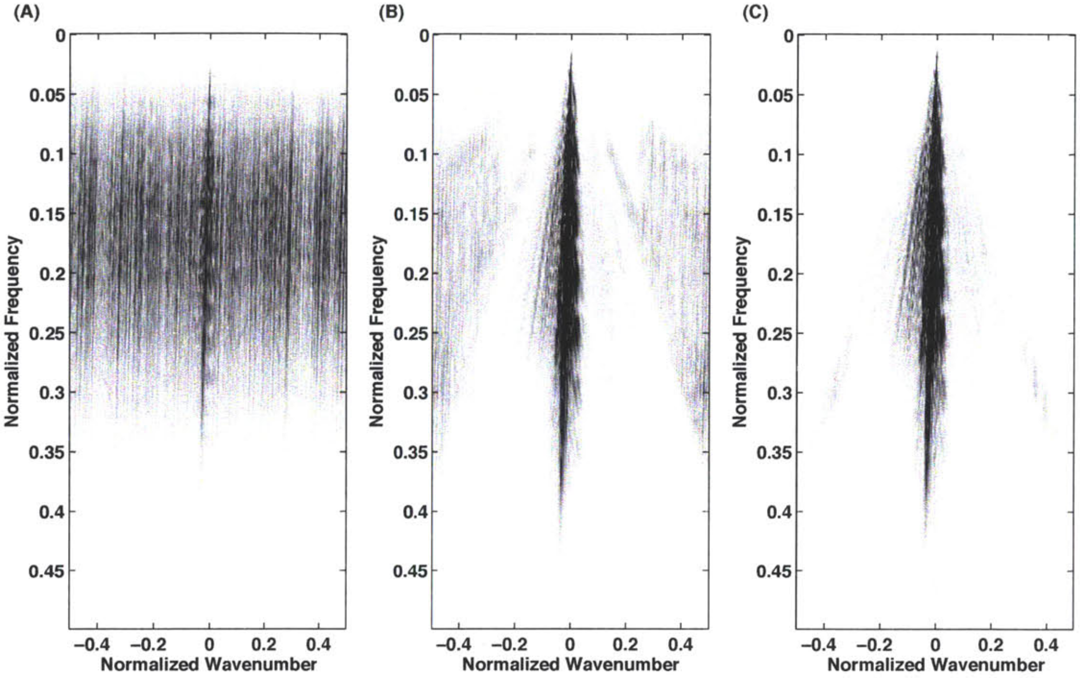


Figure 4-8 The *f-k* spectra of the data in Fig. 4-7 (A-C), respectively.

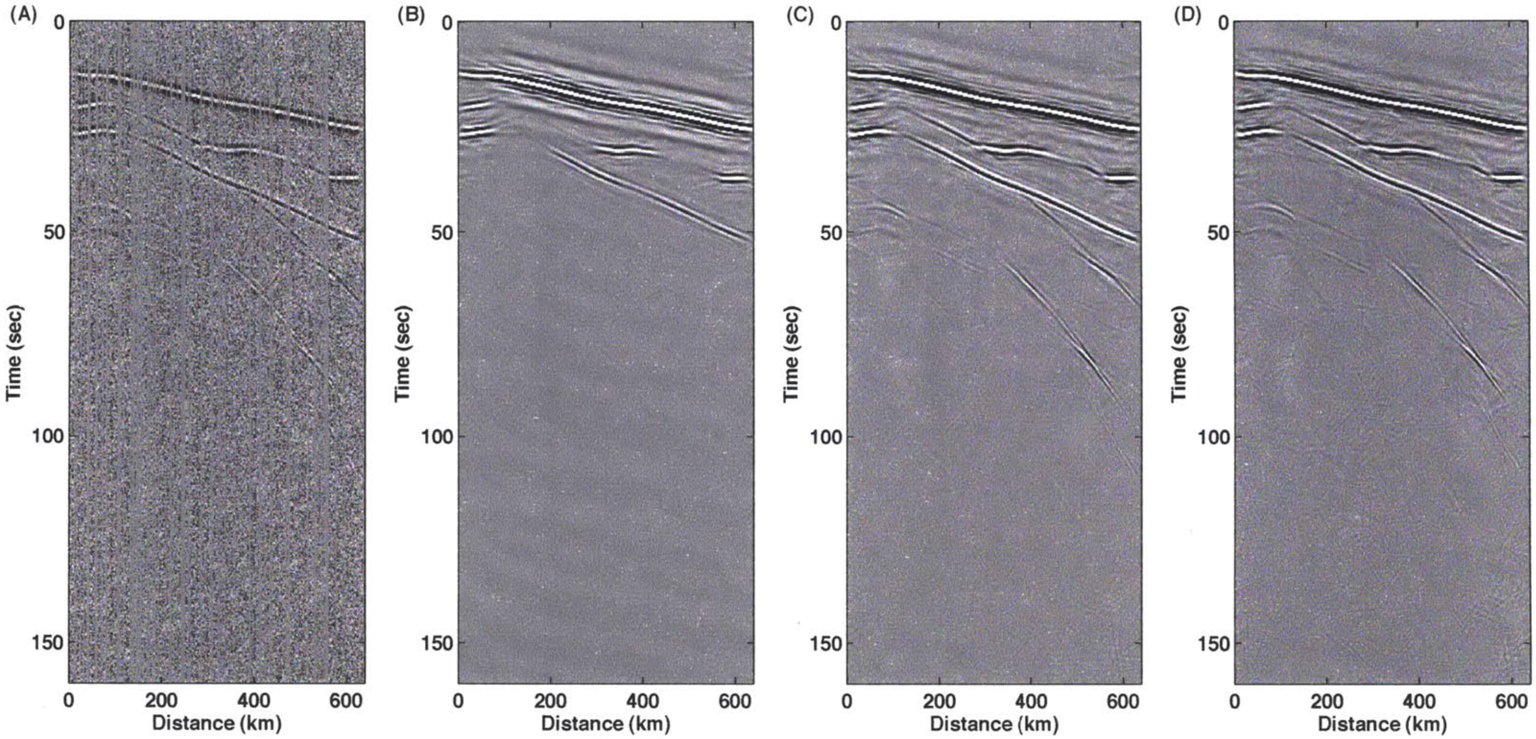


Figure 4-9 (A) Noisy data constructed by removing 50% of the original traces randomly and then adding 30% white Gaussian noise. (B-D) Interpolated data for different parameter τ along the Pareto curve indicated as black, red and green circles in Fig. 4-11A, respectively.

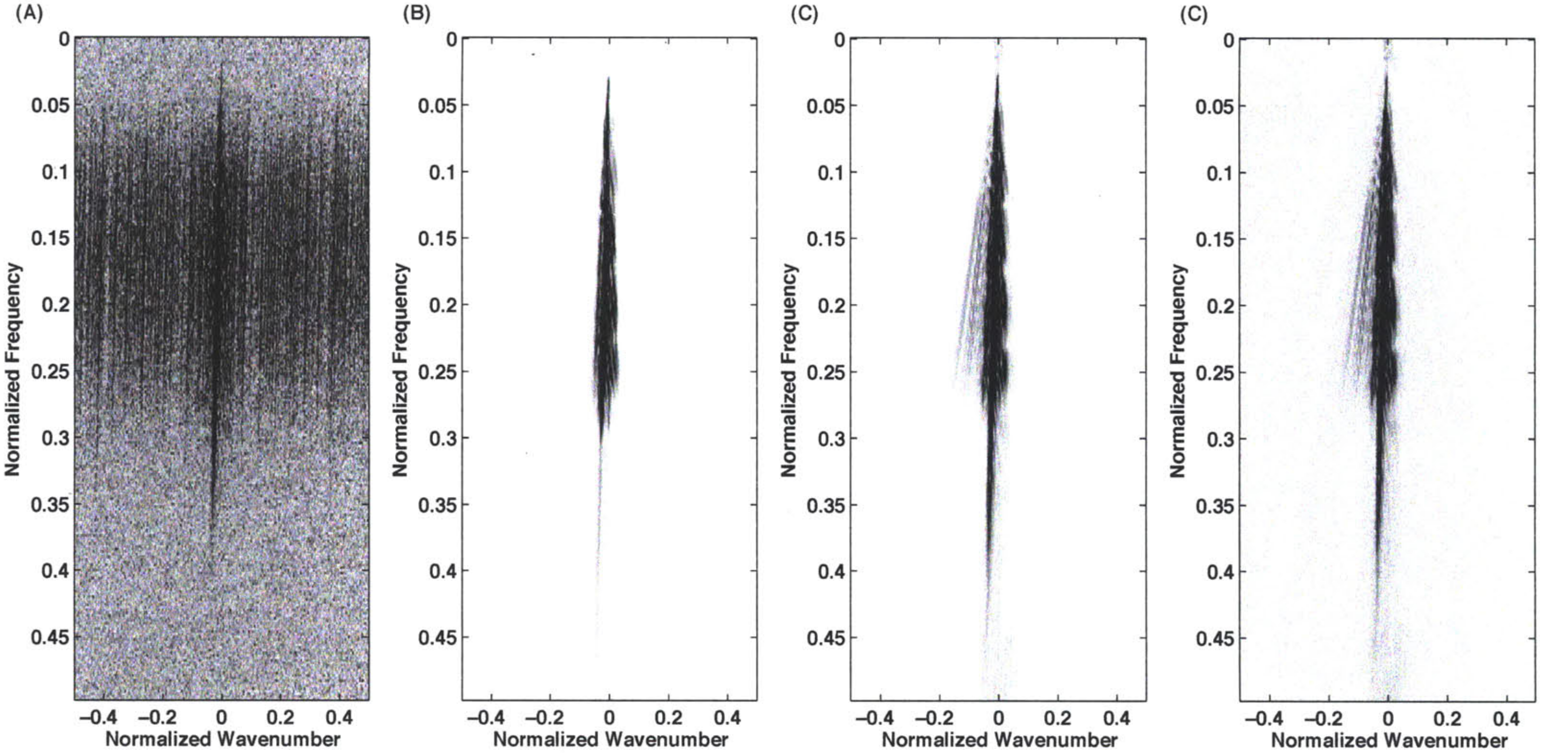


Figure 4-10 The *f-k* spectra of the data in Fig. 4-9 (A-D), respectively

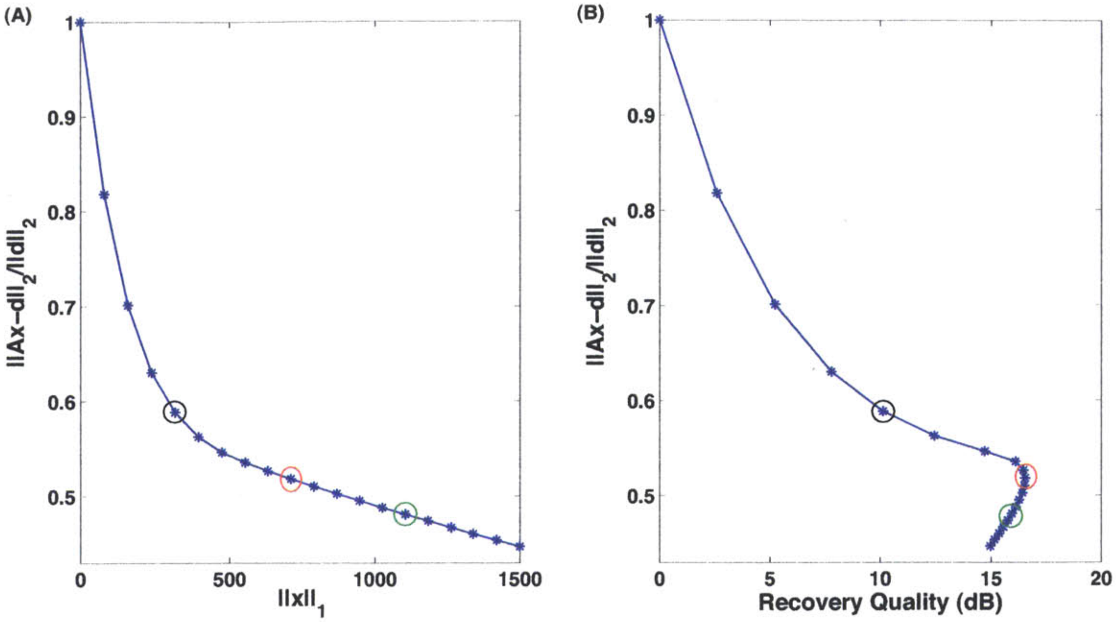


Figure 4-11 (A) The Pareto curve for the data recovery with noise. The original data is shown in Fig. 4-9A. The horizontal axis is one-norm of the curvelets, and the vertical axis is two-norm of the residual of data fitting (normalized by the norm of the data). The star symbols are those numerical sampling points along the Pareto curve by solving a series of Lasso problems. Three points depicted as colored circles are selected to investigate the influence of parameter τ on the curvelet interpolation. The corresponding interpolated results are shown in Fig. 4-9(B-D). (B) The recovery quality (measured in dB) varies with parameter τ sampled along the Pareto curve in the left panel.

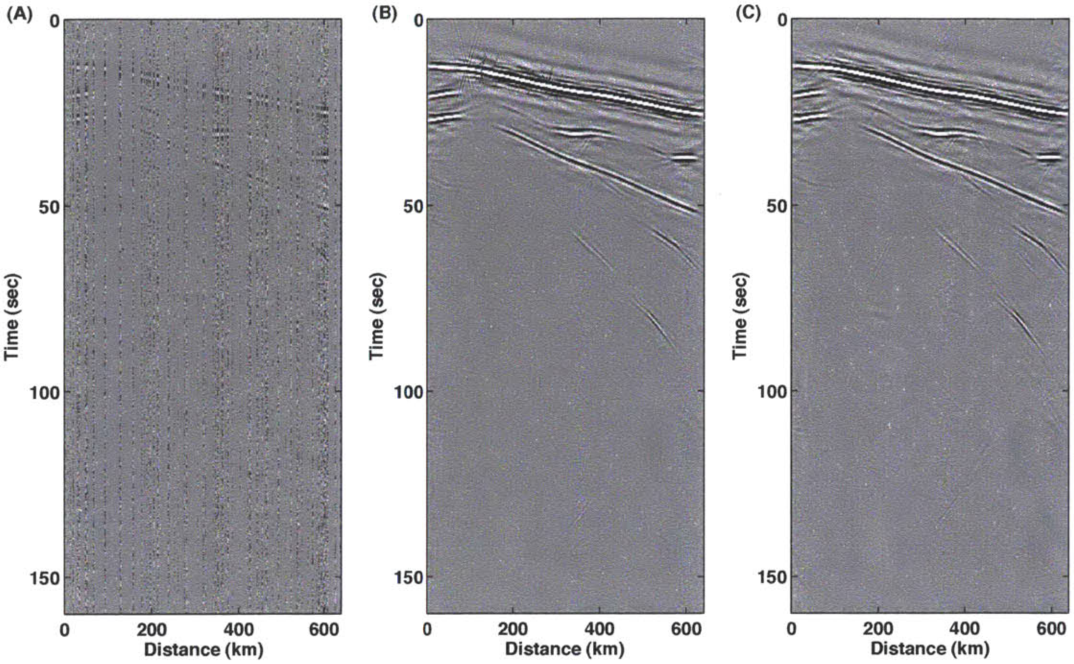


Figure 4-12 (A) Noisy data constructed by randomly removing 85% of the original traces in Fig. 4-5A, and then adding 30% white Gaussian noise. (B) Interpolated data without a mask matrix. The parameter τ used in the Lasso problem is depicted as a black circle along the blue Pareto curve in Fig. 4-14A. (C) Interpolated data with a mask matrix. The apparent velocity is 4km/s and the parameter τ is shown as a black circle on the red curve in Fig. 4-14A.

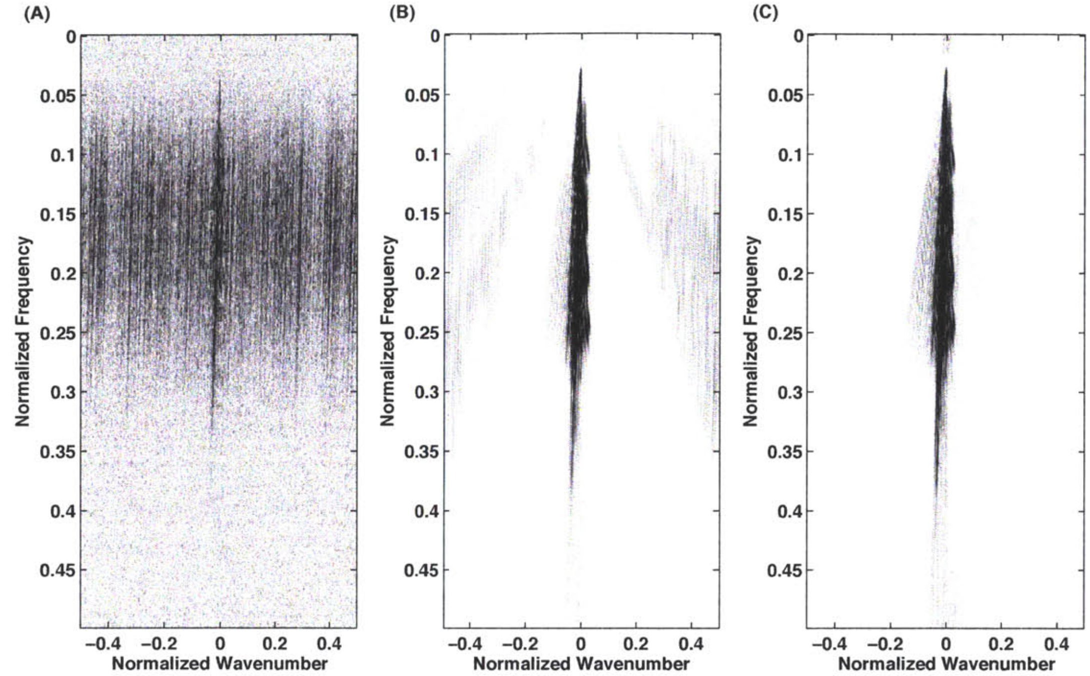


Figure 4-13 The *f-k* spectra of the data in Fig. 4-12 (A-C), respectively.

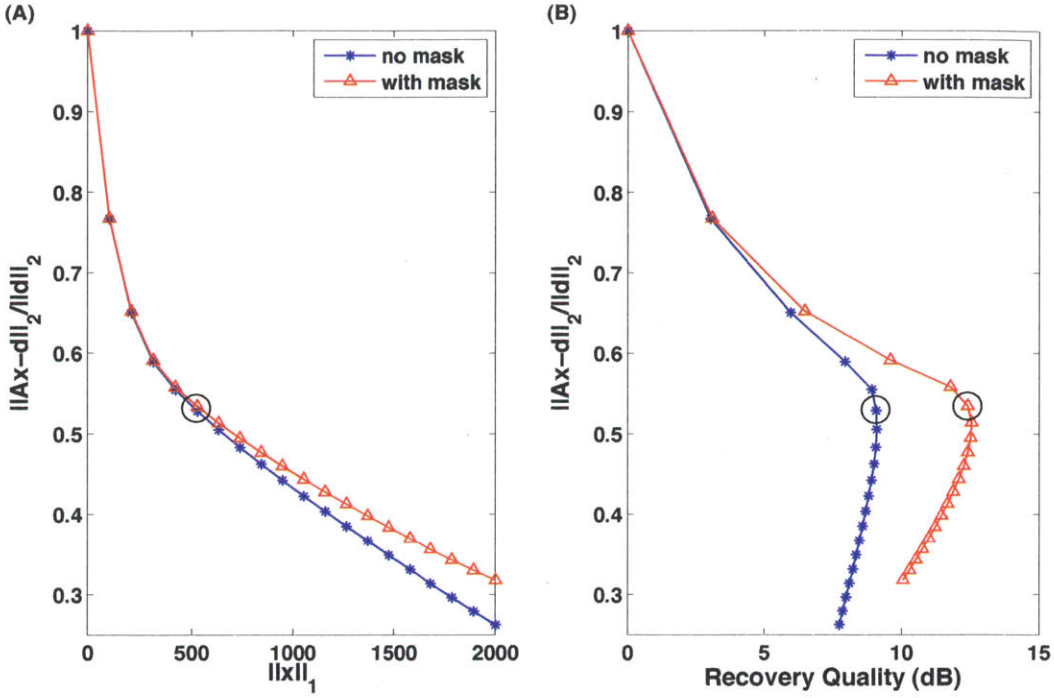


Figure 4-14 (A) The Pareto curves for data recovery with noise. The original data is shown in Fig. 4-12A. The horizontal axis is one-norm of the curvelets, and vertical axis is two-norm of the residual of data fitting (normalized by the norm of the data). The blue curve is calculated without an angular mask matrix in the interpolation, while a mask matrix with the apparent velocity 4 km/s is applied for the red curve. The star and triangle symbols are those numerical sampling points along the Pareto curves. (B) The corresponding recovery quality (measured in dB) along the Pareto curves in the left panel. The circles indicate the optimal trade-off between two-norm of data fitting and one-norm of curvelets. The corresponding interpolated data are shown in Fig. 4-12B and 4-12C, respectively.

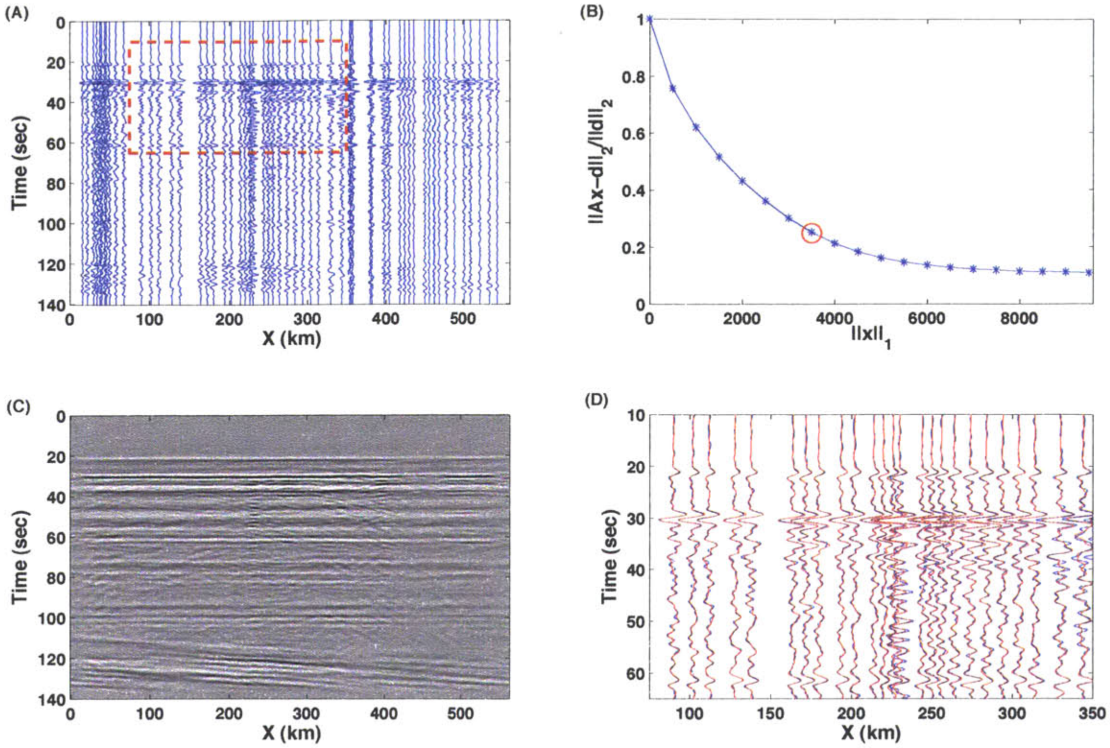


Figure 4-15 A real data example of wavefield interpolation. (A) The recorded data aligned with the first *P*-wave arrival. (B) The Pareto curve of the curvelet interpolation with a mask matrix, and the apparent velocity is chosen as 6 km/s. (C) Interpolated data with the parameter τ marked as a red circle in panel (B). (D) Zoomed-in comparison (clipped from panel (A), depicted as a red dashed box) between the interpolated and recorded traces. The recorded traces are shown in blue wiggles, and the interpolated data are plotted in red.

Chapter 5

Application of Passive Source Reverse Time Migration to Sparsely Sampled Teleseismic Data

In this chapter, we apply passive source reverse time migration (RTM) to sparsely sampled teleseismic data. RTM requires densely and regularly sampled data, which with present-day deployments can rarely be met in teleseismic observations. Therefore, wavefield interpolation in curvelet domain is introduced as an important preprocessing step for RTM. We first review the theory of passive source RTM and curvelet interpolation. Synthetic examples then demonstrate that RTM image is superior to traditional CCP stacking if dense array data is available. With the aid of wavefield interpolation, about 6 km station spacing can satisfy the stringent sampling requirement of RTM for teleseismic imaging. Application to Hi-CLIMB data in Tibet reveals a clear and continuous Moho as well as fine structures in the crust.

5.1 Introduction

Increasingly dense seismographic arrays are being deployed recently to explore finer structures in the crust and upper mantle. For teleseismic imaging, common conversion point (CCP) stacking is routinely and successfully applied in the receiver function studies [e.g. *Dueker and Sheehan*, 1997; *Zhu*, 2000; *Gilbert et al.*, 2003]. The simple assumption in CCP stacking (horizontal interface), however, hinders accurate imaging of geologically complex structures, such as steep faults and laterally discontinuous interfaces, and will diminish the scientific return on the investment in increasingly dense array deployments. Novel imaging methods, such as seismic migration, are required, which can exploit full complexity of recorded wavefields and rely on less prior information about the Earth's structures.

As in reflection seismology, different types of migration methods can be applied to teleseismic studies, such as ray-based Kirchhoff migration [e.g. *Ryberg and Weber*, 2000; *Rondenay et al.*, 2001; *Poppeliers and Pavlis*, 2003] and wave equation based migration [e.g. *Chen et al.*, 2005; *Shang et al.*, 2012]. Ray-based methods rely on the asymptotic solution (the high-frequency limit) of the wave equation, and fail in the presence of caustics, multiple arrivals, and other complex wave phenomena. In contrast, wave equation migration solves the wave equation directly and takes the wavefield complexity into account automatically. In particular, passive source reverse time migration (RTM) proposed by *Shang et al.* [2012] has several distinct advantages over previous methods. Firstly, multi-component raw data (subject to source signature removal) from individual earthquakes are used as the boundary condition for wavefield time reverse continuation, rather than stacked receiver functions that would degrade spatial resolution due to the implicit 1D velocity assumption. Furthermore, the two-way elastic wave equation is explicitly solved and can easily handle strong medium complexity.

RTM requires regularly and densely sampled data, that is, in principle at least two sample points in each temporal and spatial cycle according to Shannon-Nyquist sampling theorem [e.g. *Liu and Sacchi*, 2004; *Xu et al.*, 2005]. In seismic acquisitions, however, and especially in case of teleseismic observations, the continuous wavefield is often sparsely (aliased) and irregularly sampled along spatial coordinates. Such a limited acquisition footprint can introduce significant "noise" in the RTM images. Therefore seismic data regularization is one of the key preprocessing steps in RTM applications. The seismic traces need to be interpolated (or extrapolated) from acquired data on sparse and irregular locations to dense and regular grids.

Wavefield interpolation algorithms usually exploit sparse transforms to compress the data redundancy in other domains, such as the Fourier transform [e.g. *Spitz*, 1991; *Sacchi et al.*, 1998; *Zwartjes and Sacchi*, 2007], the Radon transform [e.g. *Kabir and Verschuur*, 1995; *Trad et al.*, 2002], and the curvelet transform [e.g. *Herrmann and Hennenfent*, 2008; *Naghizadeh and Sacchi*, 2010a]. Among them, the curvelet transform is especially favorable in teleseismic study, because curvelets are optimally sparse representations of the wave propagator and can efficiently represent seismic data [*Candès and Demanet*, 2005]. *Herrmann and Hennenfent* [2008] interpolated seismic data by curvelet reconstruction with sparsity-promoting inversion (CRSI). The missing traces can be efficiently reconstructed even the sampling ratio is below Nyquist frequency. In other words, with the aid of compressive sampling, an optimally sparse acquisition can be designed without losses of wavefield information.

As an example of our passive source RTM application, we use data recorded by Hi-CLIMB (Himalayan-Tibetan Continental Lithosphere during Mountain Building) broadband seismographic array in Tibet. With station spacing of 5-15 km and an aperture of over 500 km, Hi-CLIMB array provides an excellent opportunity to delineate the geometry of tectonic plates beneath the Tibet plateau, especially the north limit of underthrusting Indian crust and lithosphere, which is vital to understand the uplift mechanism in the continental collision zone [*Nábělek et al.*, 2005; *Nábělek et al.*, 2009; *Nowack et al.*, 2010].

In this paper, we first review the theory of passive source RTM and curvelet interpolation. Synthetic experiments are then performed to demonstrate the workflow on teleseismic imaging. Finally passive source RTM is applied to Hi-CLIMB data for lithosphere imaging.

5.2 Passive Source Reverse Time Migration

The principle of passive source RTM is simple, and mainly utilizes the wave mode conversion in the elastic wave propagation. Teleseismic *P* (or *S*) waves can convert to other modes (*P*-to-*S* or *S*-to-*P*) when impinging on an interface or scatterer from below. The transmitted and converted waves then propagate with different wave speeds to receivers on the Earth's surface. By time reversal, due to the wavespeed difference for *P* and *S* waves, the relevant parts of the transmitted wave front coincide with the converted wave only at the time of mode conversion. The conversion time (and conversion location given a reference velocity model) can be determined by applying a cross-correlation like imaging condition.

In teleseismic applications there are three main steps: First, the snapshots of the elastic wavefield are reconstructed from the recorded multicomponent array data, using reverse-time continuation. Second, *P* and *S* components for each snapshot are separated by polarization decomposition. Finally, a cross-correlation imaging condition between *P* and *S* waves is applied. The final image is then obtained by summation of partial images from individual events. Note that in active source RTM both forward simulation from source side and backward propagation from receiver side are required, while only one back-propagation from receiver side is necessary for passive source RTM. This eliminates the uncertainty of source location and source time function. We present here the main equations of passive source RTM for the purpose of self-containment. For detailed derivation and implementation, we refer to *Shang et al.* [2012], that is Chapter 2 of this thesis, and references therein.

In 3-D isotropic elastic media, the back-propagated displacement wavefield from receiver side $\mathbf{u}_r$ can be decoupled into *P*, *SV* and *SH* components by projection operators Q^* [*Brytik et al.*, 2011]:

$$u_{r,\mathrm{P}} = Q_{\mathrm{P}}^* \, \mathbf{u}_r \quad u_{r,\mathrm{SV}} = Q_{\mathrm{SV}}^* \, \mathbf{u}_r \quad \text{and} \quad u_{r,\mathrm{SH}} = Q_{\mathrm{SH}}^* \, \mathbf{u}_r \;, \tag{5.1}$$

which are defined as

$$\begin{aligned} Q_{\mathrm{P}}^* &= (-\Delta)^{-1/2} \frac{1}{i} \begin{pmatrix} \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \end{pmatrix} \\ Q_{\mathrm{SV}}^* &= (-\Delta)^{-1/2} (-\Delta)^{-1/2} \begin{pmatrix} \dfrac{\partial^2}{\partial x_1 \partial x_3} & \dfrac{\partial^2}{\partial x_2 \partial x_3} & -\dfrac{\partial^2}{\partial x_1^2} - \dfrac{\partial^2}{\partial x_2^2} \end{pmatrix} \\ Q_{\mathrm{SH}}^* &= (-\Delta')^{-1/2} \frac{1}{i} \begin{pmatrix} -\dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_1} & 0 \end{pmatrix} \end{aligned} \tag{5.2}$$

Here, Δ is the Laplacian operator, Δ' is $\partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$, and i is the imaginary unit. In 2-D case the projection operators degenerate to:

$$\begin{aligned} Q_{\mathrm{P}}^* &= (-\Delta)^{-1/2} \frac{1}{i} \begin{pmatrix} \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} \end{pmatrix} \\ Q_{\mathrm{S}}^* &= (-\Delta')^{-1/2} \frac{1}{i} \begin{pmatrix} -\dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_1} \end{pmatrix} \end{aligned} \tag{5.3}$$

The imaging condition for *P* wave incidence from a single source is [*Brytik et al.*, 2012]:

$$\begin{aligned} I(\mathbf{x}) = \int_0^T \Bigg\{ & \left(\frac{1}{i} \partial_t u_{r,\mathrm{p}}(\mathbf{x},t) \right) \left(\frac{1}{i} \partial_t u_{r,\mathrm{sv}}(\mathbf{x},t) \right) \\ & + \left(\frac{1}{i} \nabla \right) u_{r,\mathrm{p}}(\mathbf{x},t) \cdot \left(\frac{1}{i} \beta(\mathbf{x}) (-\Delta)^{-1/2} \partial_t \nabla \right) u_{r,\mathrm{sv}}(\mathbf{x},t) \Bigg\} dt \end{aligned} \tag{5.4}$$

where T is the total recorded time, and $\beta(\mathbf{x})$ is *S* wave velocity. For *S* wave incidence, one can exchange $u_{r,\mathrm{p}}$ with $u_{r,\mathrm{sv}}$ and replace $\beta(\mathbf{x})$ by $\alpha(\mathbf{x})$ (*P* wave velocity). At sufficiently high frequencies the operator $(-\Delta)^{-1/2}\partial_t$ can be approximated (in the asymptotic limit) as $\beta(\mathbf{x})$ – or $\alpha(\mathbf{x})$ for *S* wave incidence.

5.3 Curvelet Interpolation with Sparsity Promotion

Following the analysis in Chapter 4, we consider the observed data **d** as a subset of the desired interpolated data **m**. This relationship can be represented as

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n} \ , \tag{5.5}$$

where **G** is a sampling operator and **n** is additive noise [*Naghizadeh and Sacchi*, 2010b]. We can solve this underdetermined problem by finding curvelets **x** that minimizes

$$\frac{1}{2}\|\mathbf{d}-\mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1 \ , \tag{5.6}$$

where $\mathbf{A}=\mathbf{G}\mathbf{C}^{\mathrm{T}}$, $\mathbf{C}^{\mathrm{T}}$ the adjoint of the curvelet transform operator, and λ a Lagrange multiplier that can be used to control the trade-off between the data misfit and the sparsity (one-norm) of the solution. The desired interpolated data **m** then can be calculated from $\mathbf{m} = \mathbf{C}^{\mathrm{T}}\mathbf{x}$.

The optimal damping parameter λ can be determined from L- (or Pareto) curves by solving either a series of basis pursuit denoising (BPDN) or Lasso problems. The basis pursuit denoising problem is described as:

$$\underset{\mathbf{x}}{\text{minimize}} \ \ \|\mathbf{x}\|_1, \quad \text{subject to} \ \ \|\mathbf{A}\mathbf{x}-\mathbf{d}\| \le \sigma \ , \tag{5.7}$$

whereas the Lasso problem is defined as:

$$\underset{\mathbf{x}}{\text{minimize}} \ \ \|\mathbf{A}\mathbf{x}-\mathbf{d}\|_2, \quad \text{subject to} \ \ \|\mathbf{x}\|_1 \le \tau \ . \tag{5.8}$$

As described in Chapter 4, we can choose a weighting function, that is, a mask matrix **M** in operator **A**, so that $\mathbf{A}=\mathbf{G}\mathbf{C}^{\mathbf{T}}\mathbf{M}$ in the curvelet domain [*Herrmann and Hennenfent*, 2008], to penalize (or mute) nearly vertical curvelets (horizontally propagated wave packets) that are not likely true in *t*-**x** domain.

5.4 Synthetic Examples

In this section, we apply passive source RTM with curvelet interpolation to data simulated (by a finite difference code) from a known velocity model. RTM is first applied without wavefield interpolation, and compared with the results from traditional CCP stacking method. Curvelet interpolation is then performed as a preprocessing step of RTM, which would improve the image quality remarkably.

5.4.1 Synthetic Models and Data

The test model, depicted in Fig. 5.1, is designed with in mind the application to data from the Hi-CLIMB project in southern Tibet. For simplicity, we include two main layers in the model, which could represent crust and mantle, separated from each other by an interface, say the "Moho". In the middle part of this section the synthetic Moho is disrupted and the velocity increases gradually from the shallow to deep layer. Furthermore, the shallow layer consists of two blocks (with boundary at 450 km horizontal distance) in order to represent different geological units in Tibet, and depth offset (of 20 km) of the synthetic Moho is introduced at 150 km horizontal distance.

The synthetic data are generated by a finite difference method. Ten plane *P* waves are injected in the mantle as incident waves, five of which are from the left side and others are from the right side (Fig. 5-1A). The incident angle with vertical axis is from 20° to 40° with 5° increment. The corresponding epicentral-distance varies from 30° to 90°, which is typical for teleseismic studies. A Ricker wavelet is used as the source time function, and the central frequency is 0.5 Hz (~1 Hz for the maximum frequency). For simplicity but without loss of generality, no free surface condition is applied in the forward simulation. There are only direct *P* wave and converted waves in the data, shown in Fig. 5-2 as examples. The free surface condition would introduce surface related multiples and affect the final images, but that is beyond the scope of this study. In order to investigate the effect of receiver spacing, the recorded wavefield, e.g. in Fig. 5-2, is randomly sampled along the spatial coordinate with an average interval of 2 km, 6 km, 10 km and 20 km.

5.4.2 Comparison with CCP Stacking

CCP stacks of data from ten events with different station spacing (from 2 km to 20 km) are shown in Fig. 5-3. Average 1D wavespeed models are used as the background models for ray tracing. The relatively flat part of the synthetic Moho (that is, at 0-100 km and 500-600 km horizontal distance) is well imaged even with 20 km spacing samplings (Fig. 5-3D). The image quality of the dipping Moho (100-300 km and 400-500 km) improves as the receiver spacing decreases; see, for instance, the diminishing staircases from Fig. 5-3D to 5-3C to 5-3B. However, the improvement becomes trivial at certain point, and

Fig. 5-3A and Fig. 5-3B are almost identical. The vertical or steeply dipping structures are not resolved at all. The details of the disrupted Moho are not well recovered.

We then conduct passive source RTM without wavefield interpolation on the same datasets. For a fair comparison, the background velocity models are the same as above. The images from differently spaced data are shown in Fig. 5-4. For coarse sampling, e.g. 20 km in Fig. 5-4D, there are many circle-like artifacts due to the spatial aliasing, and the Moho is barely seen in the image. With increasing number of stations, such artifacts are gradually suppressed. The kink at 150 km horizontal distance starts to show up after the station interval is less than 10 km (Fig. 5-4C). The topography of the gradual Moho in the middle is observable in the case of 6 km (Fig. 5-4B). For an array with 2 km interval (Fig. 5-4A), the Moho is delineated remarkably well with the absence of spatial alias. The Moho kink, weak disrupted Moho, and even the vertical suture are imaged with high clarity.

5.4.3 Passive Source RTM with Wavefield Interpolation

In next experiment, curvelet interpolation is employed as a preprocessing step of RTM. The randomly sampled data with average interval of 6 km, 10 km, and 20 km (as used in Fig. 5-4(B-D)) are interpolated to regular grids with 2 km increment before applying RTM. Examples are shown in Fig. 5-5 (6 km interval), Fig 5-6 (10 km) and Fig. 5-7 (20 km) for one event. In each scenario, two components of the sparse data are plotted in the upper panels and interpolated data are in the bottom respectively. Since it is noise free, we use the basis pursuit problem to recover the missing traces for each component. A mask matrix **M** with minimum apparent velocity 4 km/s is applied in the inversion. The complete dense datasets (2 km interval) are depicted in Fig. 5-2(A-B), and are considered as the ground-truth solution. The recovery quality can be evaluated in decibels (dB) by the measure $Q = -20\log_{10}(\|\mathbf{m}_0 - \hat{\mathbf{m}}\|_2 / \|\mathbf{m}_0\|_2)$, where $\mathbf{m}_0$ is the ground-truth solution and $\hat{\mathbf{m}}$ is the interpolated data [*Hennenfent and Herrmann*, 2008].

For the data with 6 km spacing, the recovery quality is 27.5 dB in the horizontal component and 21.9 dB in the vertical component (Fig. 5-5(C-D)). For more sparse case (with 10km interval), the recovery factor Q is 17.6 dB and 16.0 dB for horizontal and vertical component respectively (Fig. 5-6(C-D)). The curvelet interpolation works acceptable in the extreme case (with 20 km interval). The horizontal component recovery quality is 7.5 dB and the vertical one is 8.6 dB (Fig. 5-7(C-D)). Even in the severely undersampled case (Fig. 5-7(C-D)), major features in the data, such as converted waves from the flat part of the Moho and diffraction from the Moho kink, are recovered considerably well, though some artifacts are notable, e.g. oscillation around the first arrival and fine scale curvelets.

After wavefield interpolation, passive source RTM is applied on the regular and dense sampled data. The RTM images are shown in Fig. 5-8 (from top to bottom) for the original data with 6 km, 10 km and 20 km spacing respectively. The spatial alias is efficiently suppressed in the images, compared with Fig. 5-4(B-D). For the data interpolated from 6 km grids (Fig. 5-8A), the final RTM image is almost identical to the one from dense sampled data with 2 km interval (Fig. 5-4A). The disrupted Moho is clearly imaged in the case of 10 km (Fig. 5-8B), while the Moho kink is observable in the extreme case (Fig. 5-8C). The effect of artifacts introduced in the curvelet interpolation is not apparent for large-scale structures in the RTM images, partially due to their localization nature.

5.5 Application to Hi-CLIMB Array Data

In this section, we apply our 2D RTM (with trace interpolation) to data from the northern segment of Hi-CLIMB in order to image the lithosphere beneath Tibet. First, a great circle is estimated by a least-square fitting of the locations of a total of 71 stations, shown in Fig. 5-9. The azimuth angle is 23.1° west-of-north. The horizontal coordinate is set along the linear profile, starting from the position of 29° N and 86° E. Note the linear profile (x-axis hereafter) extends northwards to the center of the Qiangtang block, across Indus-Yarlung suture (IYS) and Bangong-Nujiang suture (BNS), which are the main collision frontiers between different tectonic blocks [*Yin and Harrison*, 2000].

In order to reduce the effect of 3D geometry, only earthquakes (m_b>5.0) approximately aligned with the x-axis are selected. The back azimuth is confined to within a 30° cone with respect to the x-axis. The epicentral-distance range is restricted from 30° to 90°, and for each selected earthquake we require it is recorded by at least 40 stations. After careful selections of data, two groups of seismic sources are used for RTM imaging. The southeastern (SE) group consists of 70 earthquakes, most of which cluster around the Java trench, while only 5 scattered events are in the northwestern (NW) group due to the low seismicity (Fig. 5-10A). The approximate *Ps* conversion points from a total of 75 events are plotted in Fig. 5-10B, assuming a Moho depth of 70 km. The half width of the swath is less than 15 km; this is comparable to the scale of the Fresnel zone, which suggests that small scale heterogeneities away from the *x*-axis form at most a limited contribution to the final 2D image [*Nowack et al.*, 2010].

For each earthquake, the recorded stations are projected along the epicentral-distance contours onto the *x*-axis. Two horizontal components (E-W and N-S components) of the seismic data are rotated to the *x*- and its perpendicular directions. In 2D RTM, only *x*- and *z*- (vertical) components are used for wavefield time reverse continuation. After removal of bad data, there are a total of about 4,000 traces for each component. The gap between adjacent traces varies from 1 km to 60 km with an average of 10 km. The histogram is illustrated in Fig. 5-11. The distance along the *x*-axis is then discretized into grids with 2 km spacing, and the trace locations are rounded up to the nearest grids. The average round-up error is about 0.5 km. The missing traces are interlaced with zeros.

Trace interpolation (into 2 km grids) is then performed component by component for each event. Due to the presence of noise in real data, we consider the Lasso problem for the analysis of the Pareto curve. One example is demonstrated in Fig. 5-12 to Fig. 5-14. Two components of raw data are plotted in Fig. 5-12. The traces are aligned with the first *P* arrival. Note the horizontal events are dominant features in the teleseismic data. Then, Pareto curves are calculated for each component, and the best regularization parameter is selected around the turning point of the Pareto curve (depicted as black circles in Fig. 5-

13). The corresponding interpolated data for two components are shown in Fig. 5-14. The amplitude along horizontal events varies smoothly, e.g. around 20 and 40 second. The slanting event between 70 to 80 second is also recovered. After interpolation, the source time function is estimated from the vertical component by principle component analysis [e.g. *Rondenay et al.*, 2005] and then deconvolved from both components by Wiener deconvolution [e.g. *Chen et al.*, 2010a]. The deconvolved two-component data are shown in Fig. 5-15.

Passive source RTM is then applied to the interpolated datasets. For the background velocity models, *P* and *S* wave tomographic results from *Hung et al.* [2010] are incorporated with the reference models used in *Nowack et al.* [2010]. The frequency band is from 0.05 Hz to 0.8 Hz. We first stack partial images from the southeastern group (Fig. 5-16C) and the northwestern events (Fig. 5-16D) separately, and then stack these two to enhance coherent features (Fig. 5-16B). In general the quality of the SE image (Fig. 5-16C) is better than the NW image (Fig. 5-16D). Southeastern events mainly cluster at the Java trench with epicentral distances between 30° and 40°, whereas the northwestern events range from 50° to 80°. Due to their distinct and complementary nature, the two images are stacked with equal weights. Otherwise, the final image would be dominated by the SE image component. The final image clearly shows a continuous Moho at a depth of about 70 km, marked as a black dashed line in Fig. 5-16B. We note that for this result we did not correct for elevation (Fig. 5-16A). The northward thinning of the crust in the Qiangtang block is consistent with previous studies [e.g. *Tseng et al.*, 2009]. More than one strong interface are present in the crust between 31°N and 32°N, suggesting complex crustal structures in this zone.

5.6 Discussion

The current RTM imaging workflow includes several steps that could introduce uncertainties (or artifacts) in the final image. The curvelet interpolation rounds locations to the nearest regular grids, so that the curvelet coefficients are not accurately calculated. This could be remedied by applying a non-equispaced curvelet transform in the sparsity promoting inversion [e.g. *Hennenfent et al.*, 2010]. In our Hi-CLIMB application, however, a typical round-up error is 0.5 km, while the dominant wavelength of S wave is about 5 km. Moreover, horizontal events dominate in the teleseismic dataset after alignment with the first P arrival, for instance, in Fig. 5-12. So the effect of inaccurate trace location on curvelet coefficient estimation is negligible for our application, especially for coarse scale curvelets.

As all other imaging methods the accuracy of RTM images depends, to some extent, on the background models. Small fluctuations of wavespeed in the model will affect the amplitude of RTM image, but the locations of prominent scatterers are not very sensitive to the details of background models. Here we assume that the background models are isotropic, so that P and S waves can be separated by polarization decomposition in Eq. (5.3). In the presence of anisotropy qP and qS waves can be separated by solving the Christoffel equation [e.g. *Dellinger and Etgen*, 1990; *Yan and Sava*, 2009]. Along the northern segment of the Hi-CLIMB array, significant shear wave splitting (~0.8 sec and mainly east-west as the fast direction) suggests strong anisotropy in the crust and upper mantle [*Chen et al.*, 2010b]. Because of the nearly north-south linear geometry of sources and receivers, however, RTM imaging with Hi-CLIMB data is not severely affected by the azimuthal anisotropy.

In Fig. 5-17 we compare our RTM results with those obtained with Gaussian beam migration [*Nowack et al.*, 2010]. The RTM image reveals similar but much cleaner lithosphere structures. In particular, the continuity of the Moho across the BNS suture zone (between 31° to 32°) is clearly delineated, while this is not well resolved by Gaussian beam migration.

In the SE image, double layers (at 50-70 km depth) are observed between 29. 7°N and 31°N. *Nábělek et al.* [2009] observed a similar feature in their CCP stacks and interpreted the deeper one as eclogitic Indian Moho, which might subduct northward up to 31°N. However, in view of the discrepancies between our SE and NW RTM images, such a far-reaching conclusion may be premature. *Nábělek et al.* [2009] also used surface multiples, such as *PpPs* and *PpSs*, and in future studies we will explore if their use in RTM can improve the resolution.

5.7 Summary

In this chapter, we first review the theory of passive source RTM and curvelet interpolation with sparsity promotion. A series of synthetic experiments demonstrates that with increasing dense seismic array, passive source RTM can exploit more about the wavefield complexity, and is superior to CCP stacking. In practice, teleseismic array sampling rarely meets stringent requirement of RTM. Synthetic examples show that with curvelet interpolation as a preprocessing step, ~6 km station spacing is enough for teleseismic imaging of the crust and lithosphere. Finally application of wavefield interpolation and RTM to Hi-CLIMB data in Tibet clearly reveals a continuous Moho near 70 km depth as well as fine structures in the crust.

References

Brytik, V., M. V. de Hoop, H. F. Smith, and G. Uhlmann (2011), Decoupling of modes for the elastic wave equation in media of limited smoothness, *Commun. Part. Diff. Eq.*, *36*(10), 1683-1693.

Brytik, V., M. V. De Hoop, and R. D. van der Hilst (2012), Elastic-wave inverse scattering based on reverse time migration with active and passive source reflection data, in *Inside Out - Inverse Problems, MSRI Publications*, edited, Cambridge University Press.

Candès, E., and L. Demanet (2005), The curvelet representation of wave propagators is optimally sparse, *Comm. Pure Appl. Math.*, *58*(11), 1472-1528.

Chen, C. W., D. Miller, H. Djikpesse, J. Haldorsen, and S. Rondenay (2010a), Array-conditioned deconvolution of multiple-component teleseismic recordings, *Geophys. J. Inter.*, *182*(2), 967-976.

Chen, L., L. Wen, and T. Zheng (2005), A wave equation migration method for receiver function imaging: 1. Theory, *J. Geophys. Res.*, *110*(B11), B11309.

Chen, W.-P., M. Martin, T.-L. Tseng, R. L. Nowack, S.-H. Hung, and B.-S. Huang (2010b), Shear-wave birefringence and current configuration of converging lithosphere under Tibet, *Earth Planet. Sci. Lett.*, *295*(1), 297-304.

Dellinger, J., and J. Etgen (1990), Wave-field separation in two-dimensional anisotropic media, *Geophysics*, *55*, 914-919.

Dueker, K. G., and A. F. Sheehan (1997), Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track, *J. Geophys. Res.*, *102*, 8313-8327.

Gilbert, H. J., A. F. Sheehan, K. G. Dueker, and P. Molnar (2003), Receiver functions in the western United States, with implications for upper mantle structure and dynamics, *J. Geophys. Res*, *108*, 2229.

Hennenfent, G., and F. J. Herrmann (2008), Simply denoise: Wavefield reconstruction via jittered undersampling, *Geophysics*, *73*(3), V19-V28.

Hennenfent, G., L. Fenelon, and F. J. Herrmann (2010), Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach, *Geophysics*, *75*(6), Wb203-Wb210.

Herrmann, F. J., and G. Hennenfent (2008), Non-parametric seismic data recovery with curvelet frames, *Geophys. J. Inter.*, *173*(1), 233-248.

Hung, S.-H., W.-P. Chen, L.-Y. Chiao, and T.-L. Tseng (2010), First multi-scale, finite-frequency tomography illuminates 3-D anatomy of the Tibetan plateau, *Geophys. Res. Lett.*, *37*(6), L06304.

Kabir, M. N., and D. Verschuur (1995), Restoration of missing offsets by parabolic Radon transform, *Geophys. Prospect.*, *43*(3), 347-368.

Kind, R., X. Yuan, J. Saul, D. Nelson, S. Sobolev, J. Mechie, W. Zhao, G. Kosarev, J. Ni, and U. Achauer (2002), Seismic images of crust and upper mantle beneath Tibet: Evidence for Eurasian plate subduction, *Science*, *298*(5596), 1219-1221.

Liu, B., and M. D. Sacchi (2004), Minimum weighted norm interpolation of seismic records, *Geophysics*, *69*(6), 1560-1568.

Nábělek, J., J. Vergne, and G. Hetényi (2005), Project Hi-CLIMB: a synoptic view of the Himalayan collision zone and Southern Tibet, paper presented at AGU Fall Meeting.

NábělekNábělek, J., G. Hetényi, J. Vergne, S. Sapkota, B. Kafle, M. Jiang, H. Su, J. Chen, and B. S. Huang (2009), Underplating in the Himalaya-Tibet collision zone revealed by the Hi-CLIMB experiment, *Science*, *325*(5946), 1371-1374.

Naghizadeh, M., and M. D. Sacchi (2010a), Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data, *Geophysics*, *75*(6), WB189-WB202.

Naghizadeh, M., and M. D. Sacchi (2010b), On sampling functions and Fourier reconstruction methods, *Geophysics*, *75*(6), Wb137-Wb151.

Nowack, R. L., W. P. Chen, and T. L. Tseng (2010), Application of gaussian-beam migration to multiscale imaging of the lithosphere beneath the Hi-CLIMB array in Tibet, *Bull. Seism. Soc. Am.*, *100*(4), 1743-1754.

Poppeliers, C., and G. L. Pavlis (2003), Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases: 1. Theory, *J. Geophys. Res.*, *108*(B2), 2112.

Rondenay, S., M. G. Bostock, and J. Shragge (2001), Multiparameter two-dimensional inversion of scattered teleseismic body waves 3. Application to the Cascadia 1993 data set, *J. Geophys. Res.*, *106*(B12), 30795-30807.

Rondenay, S., M. G. Bostock, and K. M. Fischer (2005), Multichannel inversion of scattered teleseismic body waves: Practical considerations and applicability, *Geophys. Monograph-American Geophys. Union*, *157*.

Ryberg, T., and M. Weber (2000), Receiver function arrays: a reflection seismic approach, *Geophys. J. Inter.*, *141*(1), 1-11.

Sacchi, M. D., T. J. Ulrych, and C. J. Walker (1998), Interpolation and extrapolation using a high-resolution discrete Fourier transform, *IEEE Trans. Signal Process.*, *46*(1), 31-38.

Shang, X., M. V. de Hoop, and R. D. van der Hilst (2012), Beyond receiver functions: Passive source reverse time migration and inverse scattering of converted waves, *Geophys. Res. Lett.*, *39*, L15308.

Spitz, S. (1991), Seismic trace interpolation in the *F-X* domain, *Geophysics*, *56*(6), 785-794.

Trad, D. O., T. J. Ulrych, and M. D. Sacchi (2002), Accurate interpolation with high-resolution time-variant Radon transforms, *Geophysics*, *67*(2), 644-656.

Tseng, T.-L., W.-P. Chen, and R. L. Nowack (2009), Northward thinning of Tibetan crust revealed by virtual seismic profiles, *Geophys. Res. Lett.*, *36*(24), L24304.

Xu, S., Y. Zhang, D. Pham, and G. Lambare (2005), Antileakage Fourier transform for seismic data regularization, *Geophysics*, *70*(4), V87.

Yan, J., and P. Sava (2009), Elastic wave-mode separation for VTI media, *Geophysics*, *74*(5), Wb19-Wb32.

Yin, A., and T. M. Harrison (2000), Geologic evolution of the Himalayan-Tibetan orogen, *Annu. Rev. Earth Planet. Sci.* , *28*(1), 211-280.

Zhu, L. (2000), Crustal structure across the San Andreas Fault, southern California from teleseismic converted waves, *Earth Planet. Sci. Lett.*, *179*(1), 183-190.

Zwartjes, P. M., and M. D. Sacchi (2007), Fourier reconstruction of nonuniformly sampled, aliased seismic data, *Geophysics*, *72*(1), V21-V32.

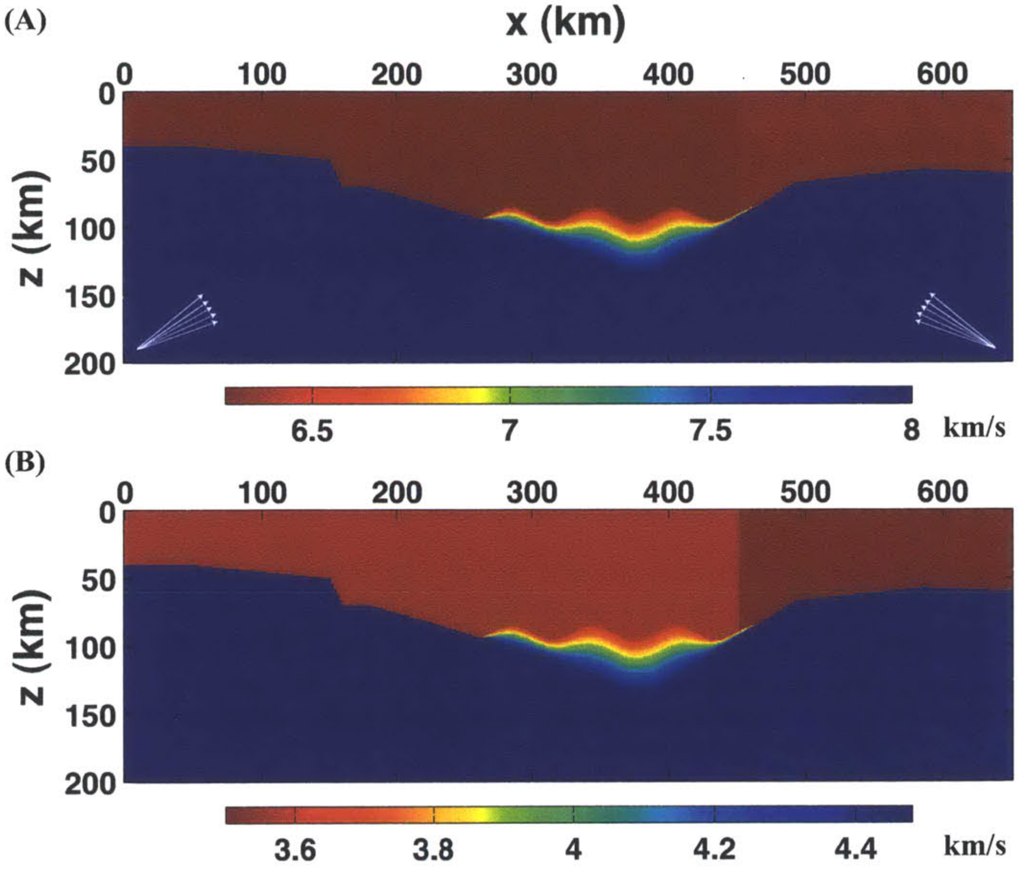


Figure 5-1 Synthetic models for (A) *P* and (B) *S* wave speeds. There are generally two layers in the models, crust (red) and mantle (blue). Two blocks are presented in the crust, separated at 450 km horizontal distance. In the middle part (from 250 km to 450 km), a transitional zone is added between the crust and mantle, in which the velocity increases gradually with depth. Ten events with plane wave incidence are simulated, and the incident angle (with respect to the vertical axis) is from 20 to 40 degree with an increment of 5. Five events are injected from the left side, and others are from the right side (white arrows in (A)).

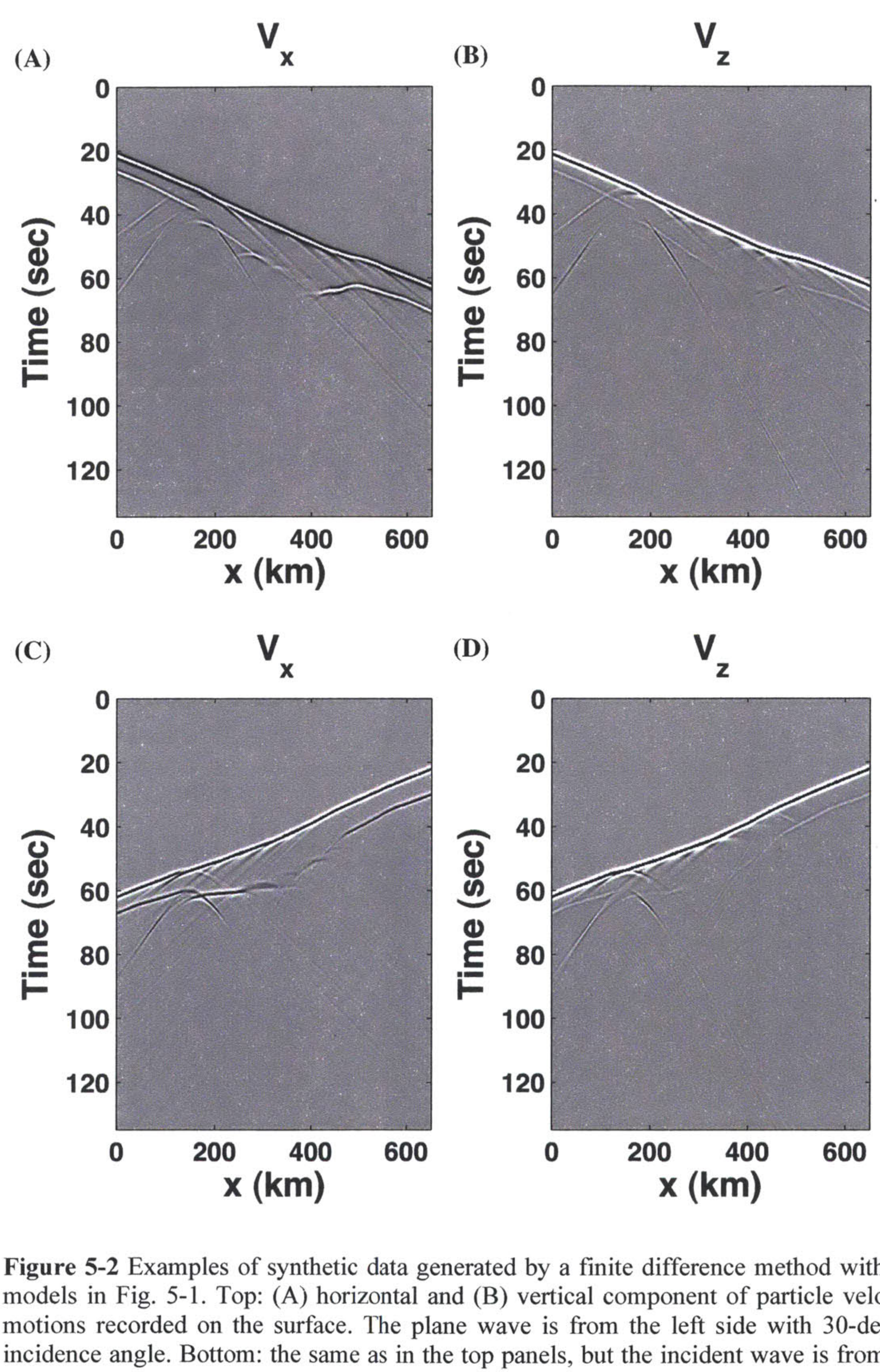


Figure 5-2 Examples of synthetic data generated by a finite difference method with the models in Fig. 5-1. Top: (A) horizontal and (B) vertical component of particle velocity motions recorded on the surface. The plane wave is from the left side with 30-degree incidence angle. Bottom: the same as in the top panels, but the incident wave is from the right side. A Ricker wavelet is used as the source time function, and the central frequency is 0.5 Hz.

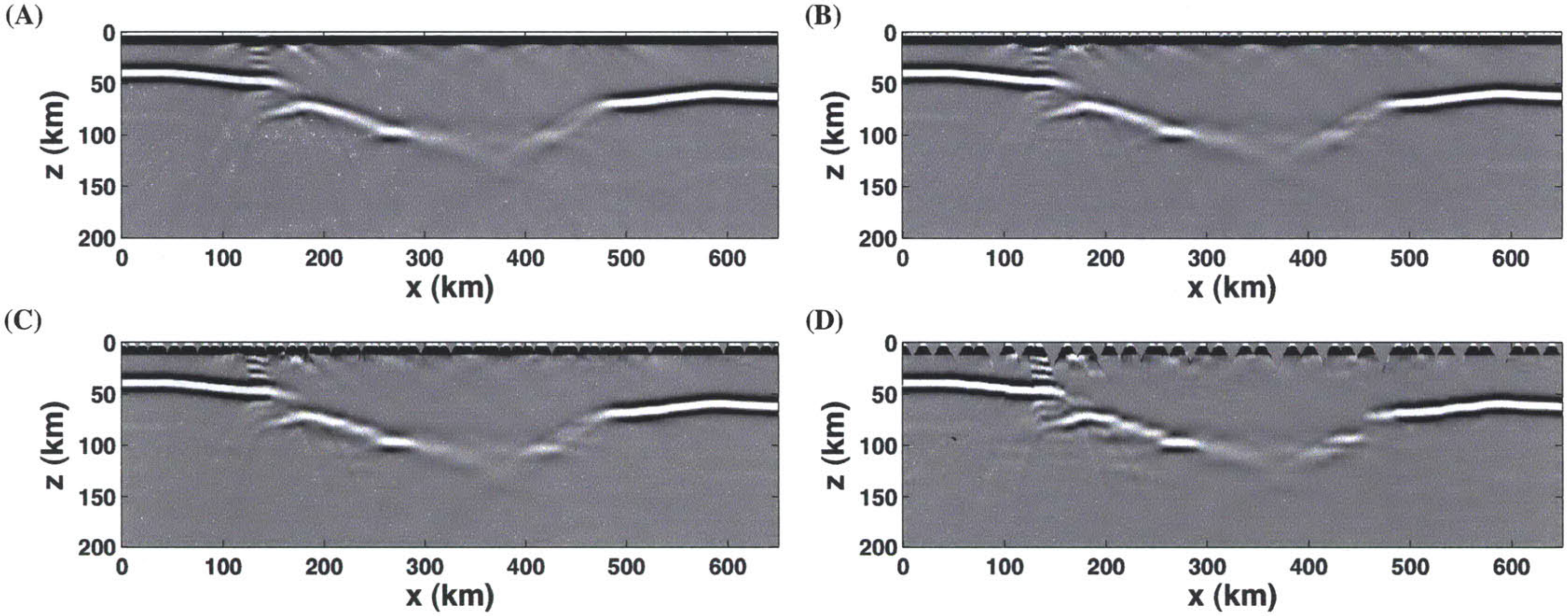


Figure 5-3 CCP stacking results with various station intervals. The stations are randomly distributed on the surface with an average spacing of (A) 2 km, (B) 6 km, (C) 10 km and (D) 20 km. For the ray tracing, 1D linear velocity models are used.

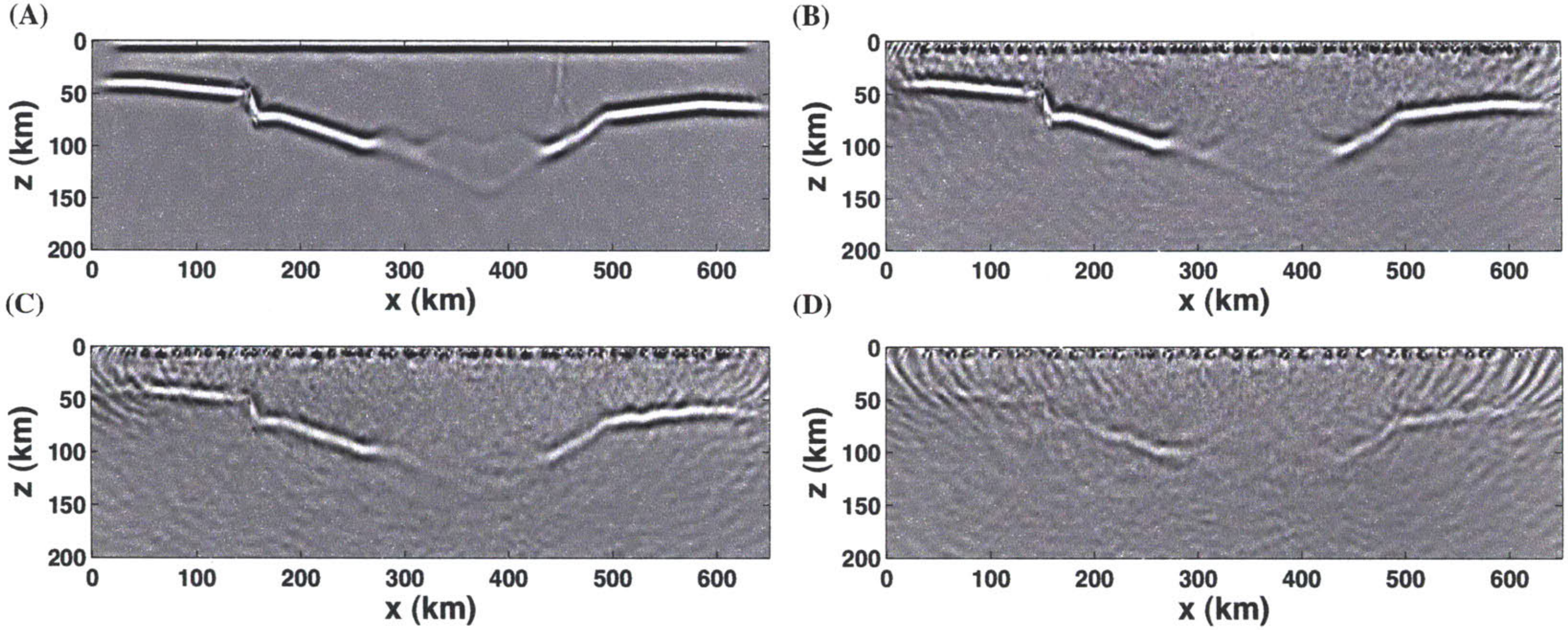


Figure 5-4 Passive source RTM results without wavefield interpolation. The input data and background velocity models are the same as in CCP stacking. The average station spacing is (A) 2 km, (B) 6 km, (C) 10 km and (D) 20 km.

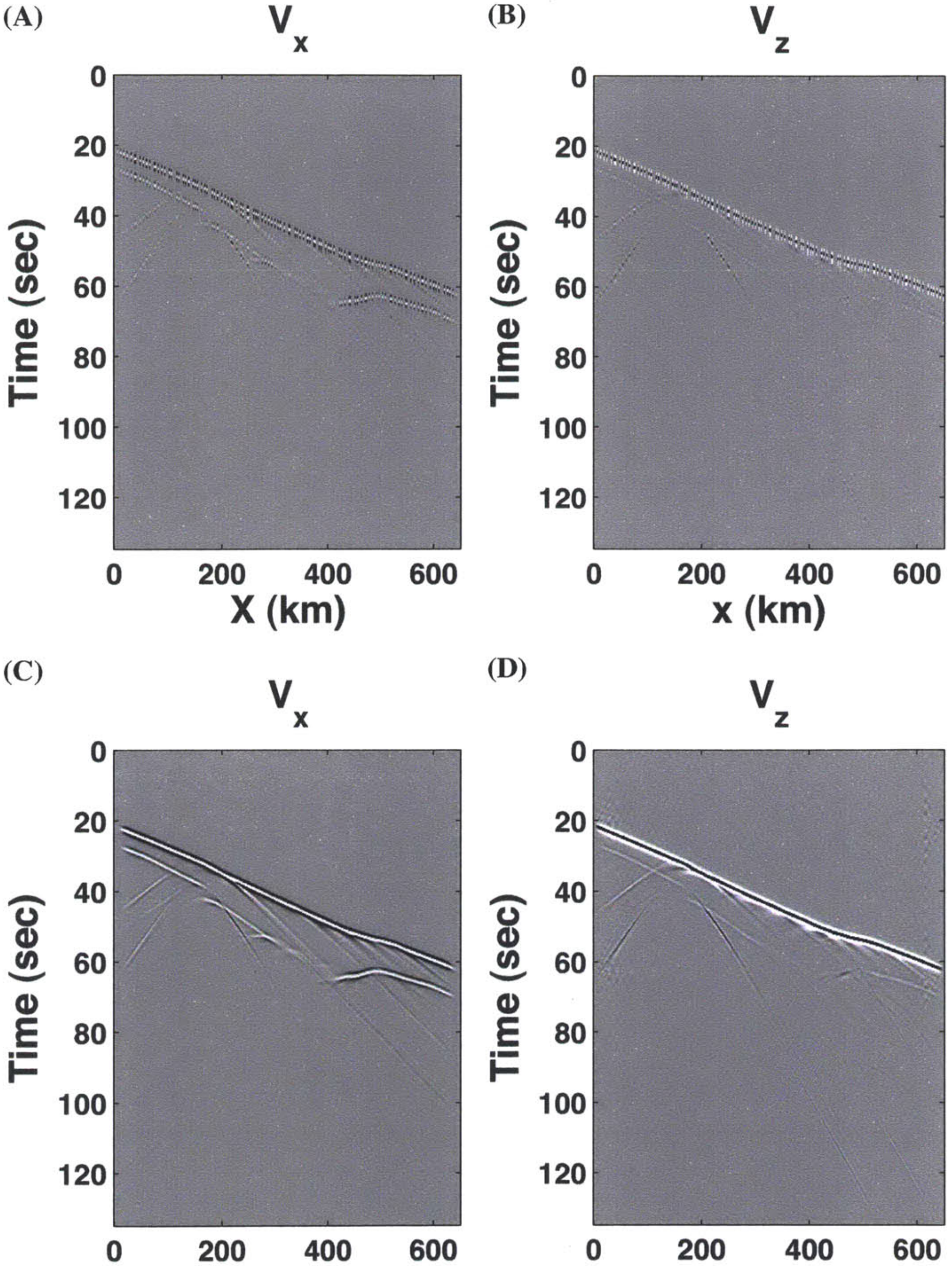


Figure 5-5 Top: (A) horizontal and (B) vertical component of randomly sampled data with an average interval of 6 km. Bottom: two components ((C) horizontal and (D) vertical) of the data interpolated into 2 km grids. The recovery quality is 27.5 dB for the horizontal component and 21.9 dB for the vertical one.

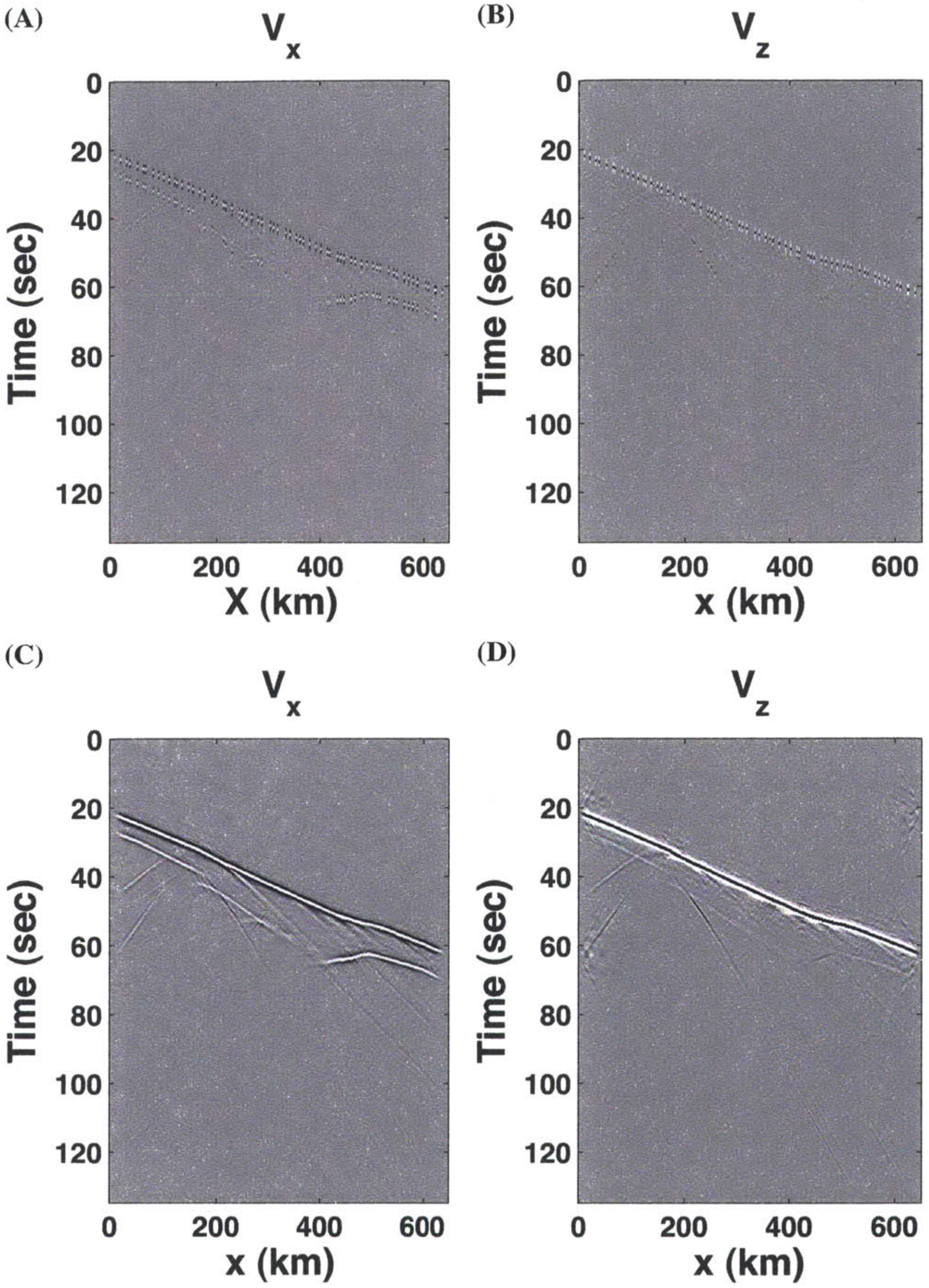


Figure 5-6 The same as in Fig. 5-5, but for the 10 km case. The recovery quality is 17.6 dB and 16.0 dB for the horizontal and vertical component respectively.

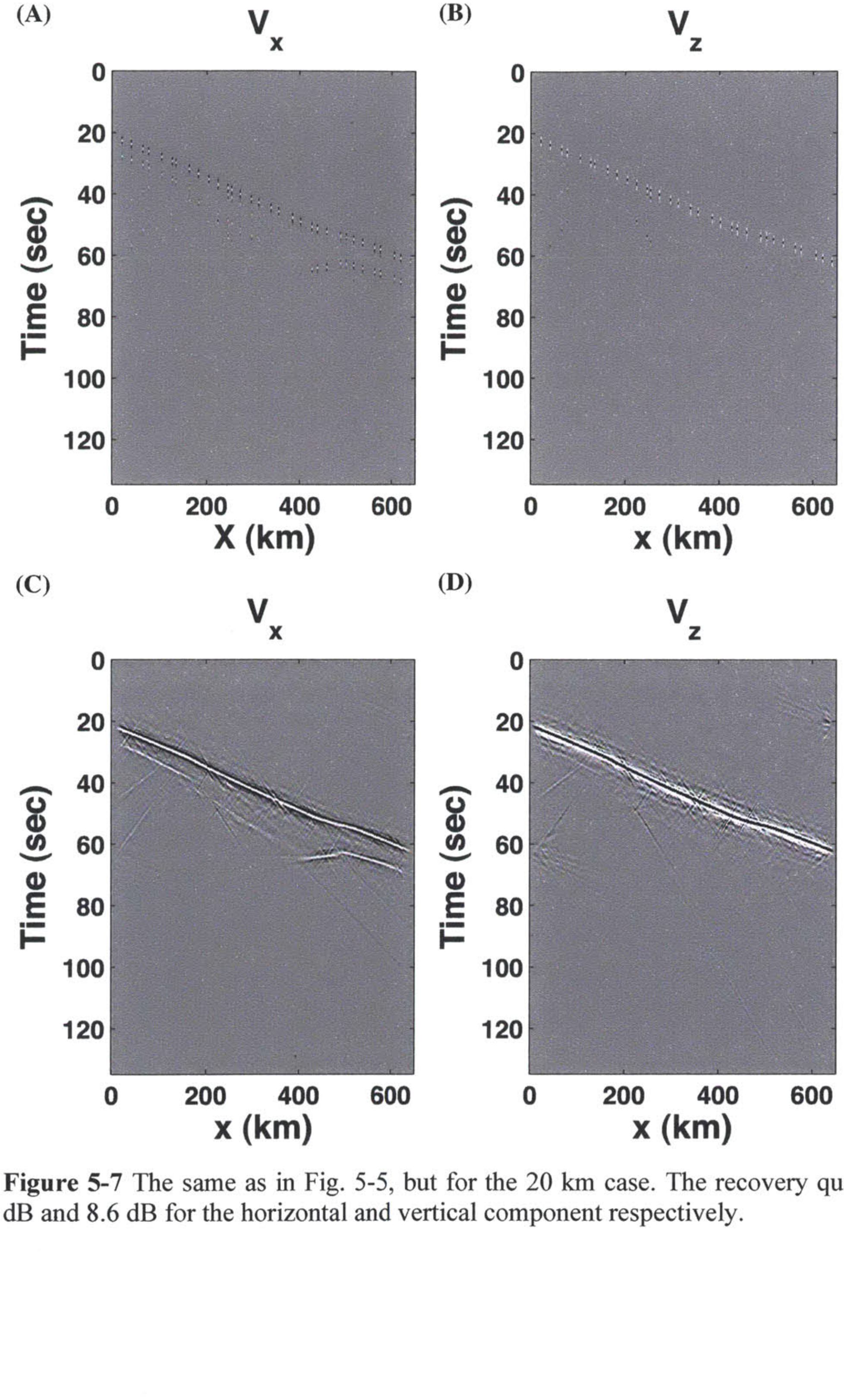


Figure 5-7 The same as in Fig. 5-5, but for the 20 km case. The recovery quality is 7.5 dB and 8.6 dB for the horizontal and vertical component respectively.

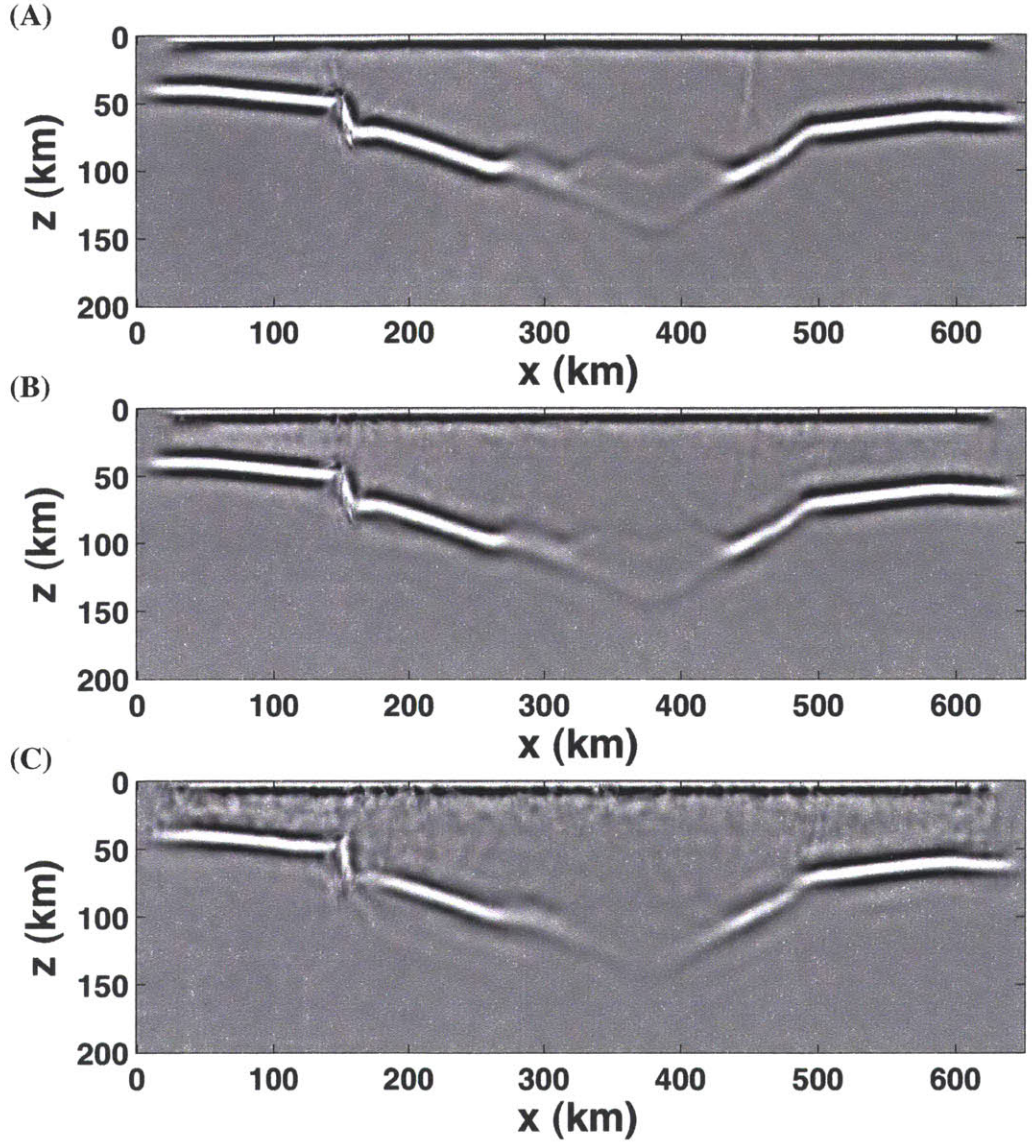


Figure 5-8 Passive source RTM results with curvelet interpolations. The input data are regularly sampled data with 2 km spacing interpolated from randomly distributed receivers with an average interval of (A) 6 km, (B) 10 km and (C) 20 km.

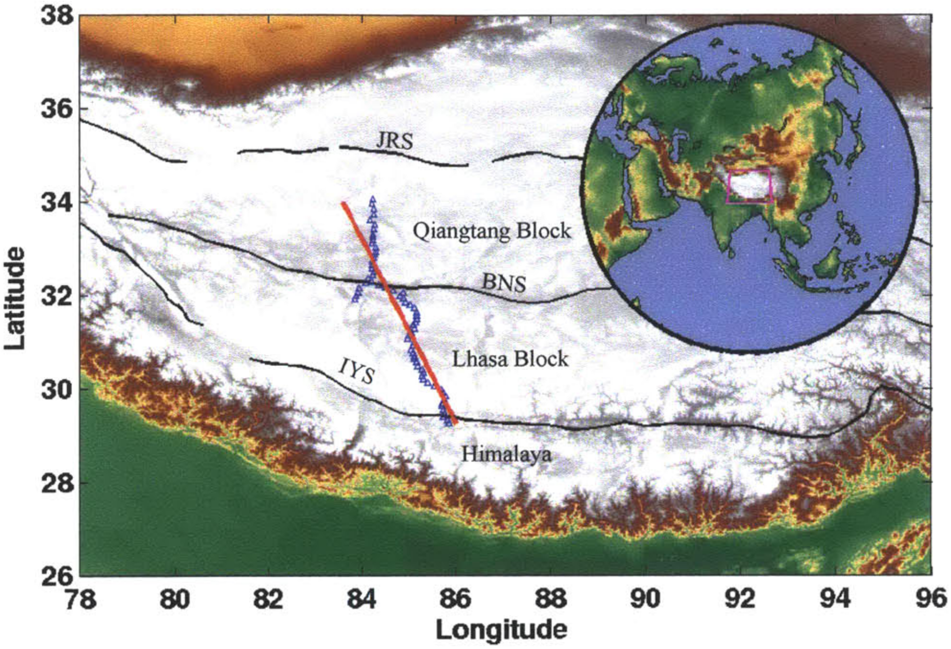


Figure 5-9 Geometry of the north segment of Hi-CLIMB array. The seismographs are depicted as blue triangles. A great circle (red line) is estimated by a least-square fitting of a total of 71 stations. The great circle starts at the point of 29° N and 86° E, with an angle of 23.1° counter-clockwise from the north. IYS stands for Indus-Yarlung suture; BNS is for Bangong-Nujiang suture and JRS is Jinsha River suture.

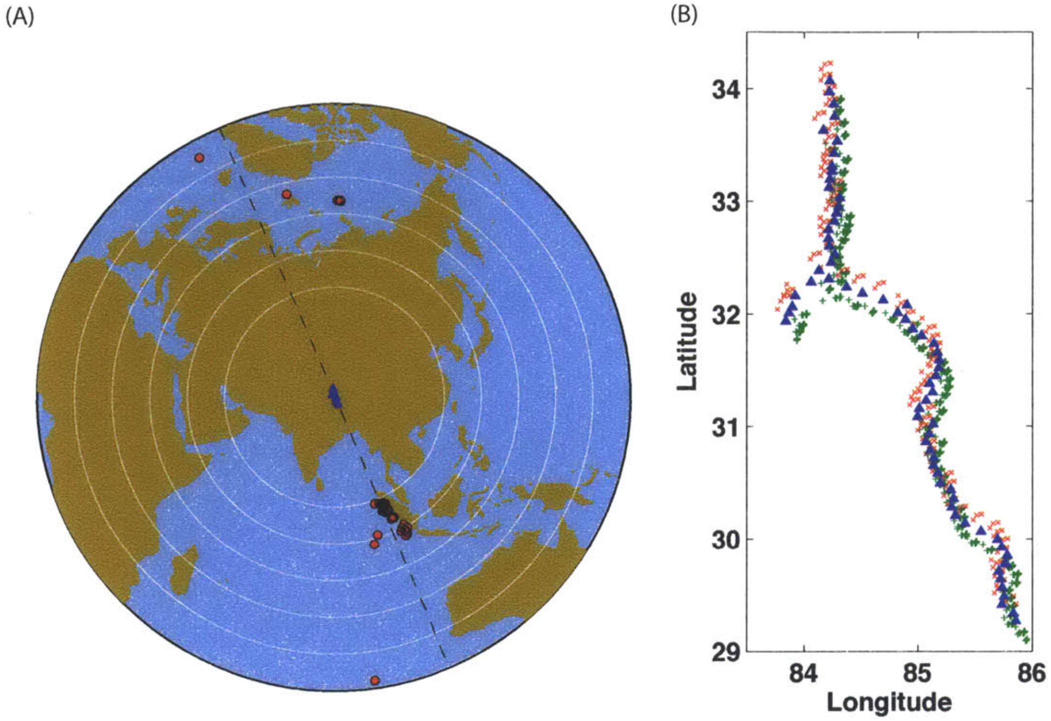


Figure 5-10 (A) Source distribution (red circles) used for the RTM imaging. Blue triangles represent stations in the north segment of Hi-CLIMB. Black dashed line is the great circle from a least-square fitting of station locations. (B) *Ps* conversion point distribution assuming the Moho depth is 70 km. The blue triangles are Hi-CLIMB stations. The green symbols are pierce points from southeastern events and the red are from northwestern events.

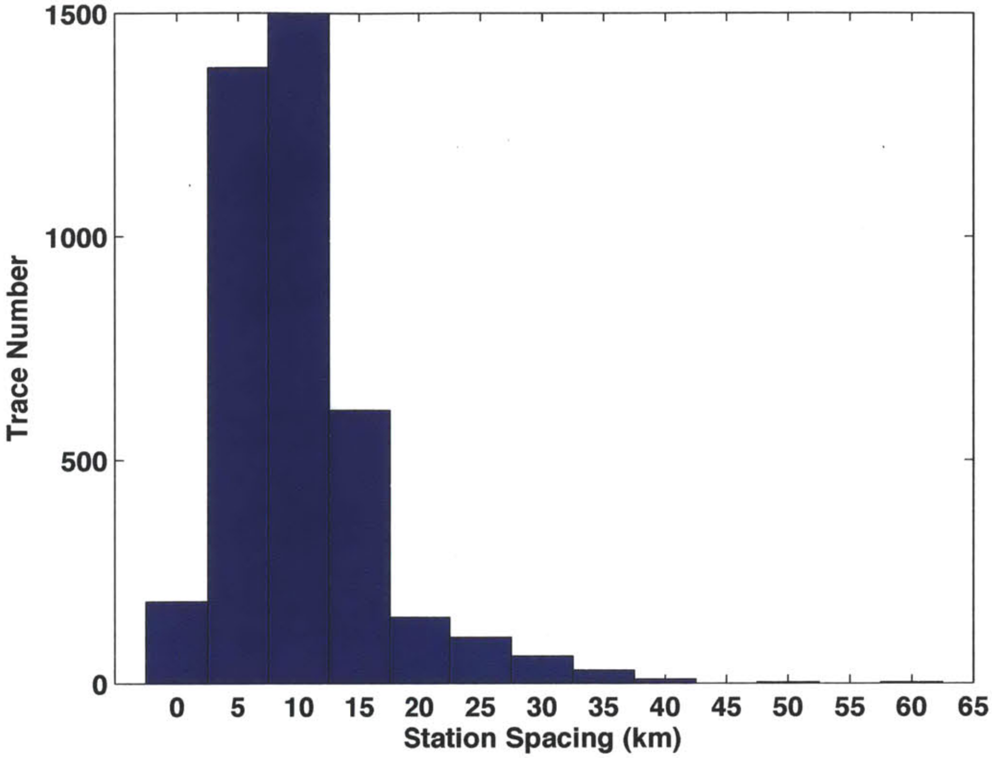


Figure 5-11 Histogram of trace intervals for 75 events used in the Hi-CLIMB RTM imaging.

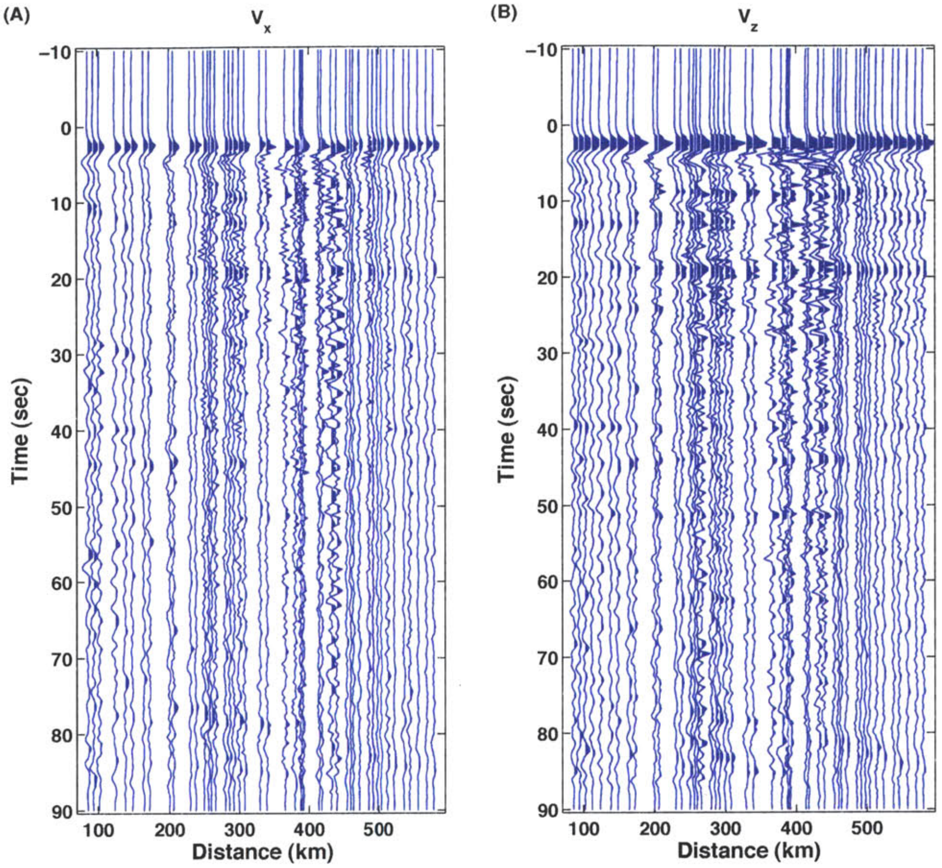


Figure 5-12 A real dataset recorded by the Hi-CLIMB array, with (A) horizontal and (B) vertical component. The data are aligned with the first *P* arrival, and bandpass filtered between 0.05 Hz and 0.8 Hz.

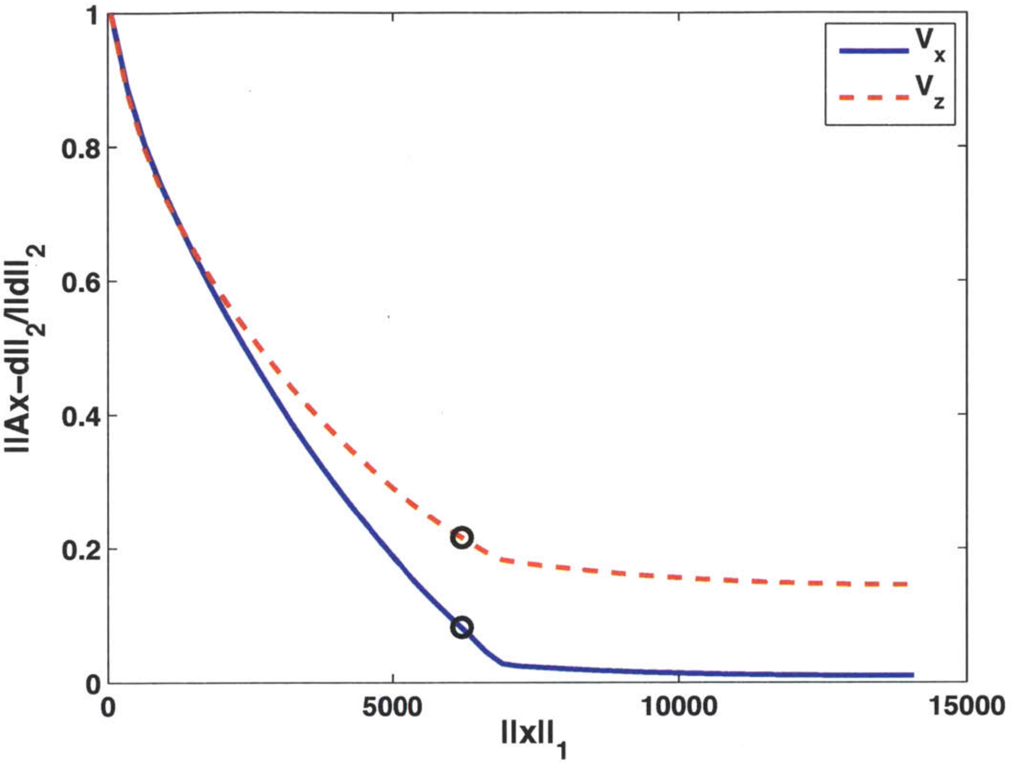


Figure 5-13 Pareto curves for the curvelet interpolation. The blue line is for the horizontal component and the red dashed line is for the vertical component. The raw data are depicted in Fig. 5-12. Black circles indicate the optimal trade-off between the data fitting and the sparsity of the solution.

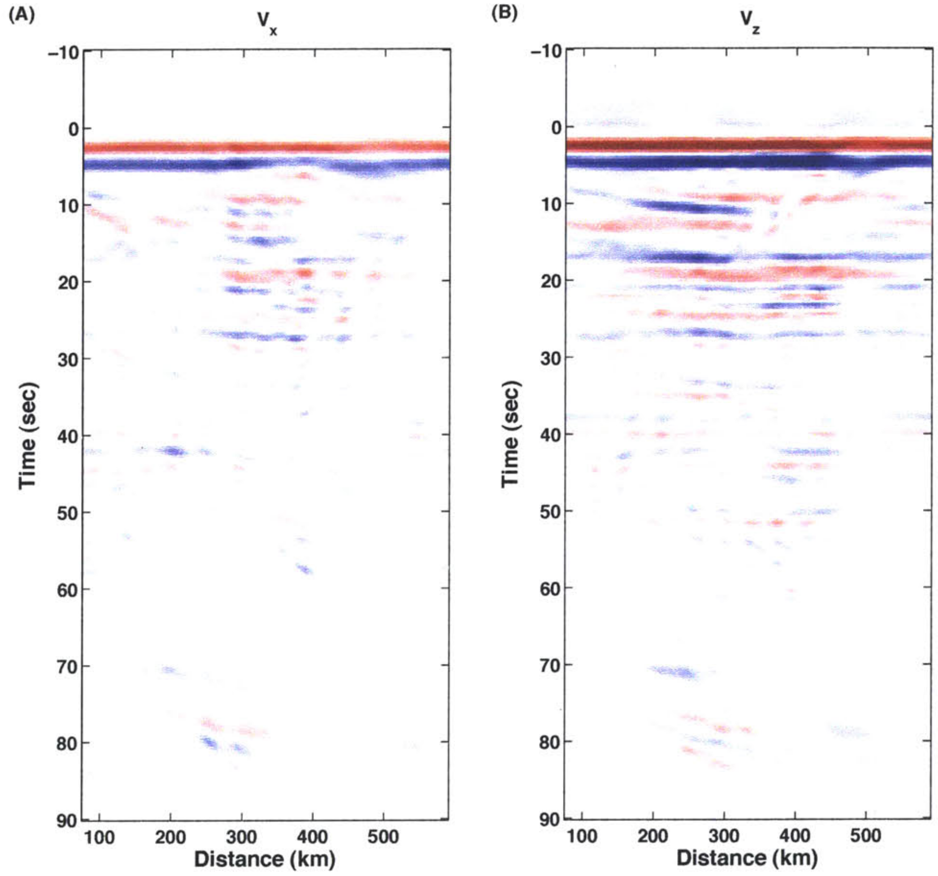


Figure 5-14 (A) Horizontal and (B) vertical component of the data after curvelet interpolation with the best trade-off parameters (black circles in Fig. 5-13). The raw data are shown in Fig. 5-12.

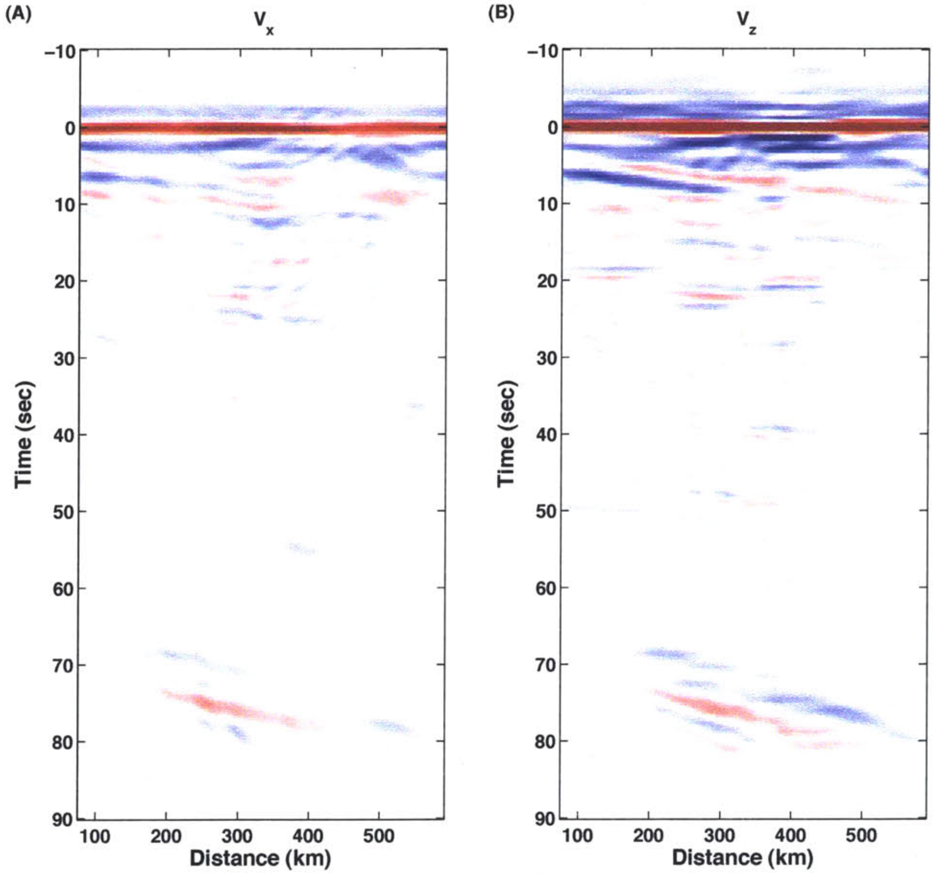


Figure 5-15 (A) Horizontal and (B) vertical component of the data after Wiener deconvolution. The input data (interpolated data) are shown in Fig. 5-14.

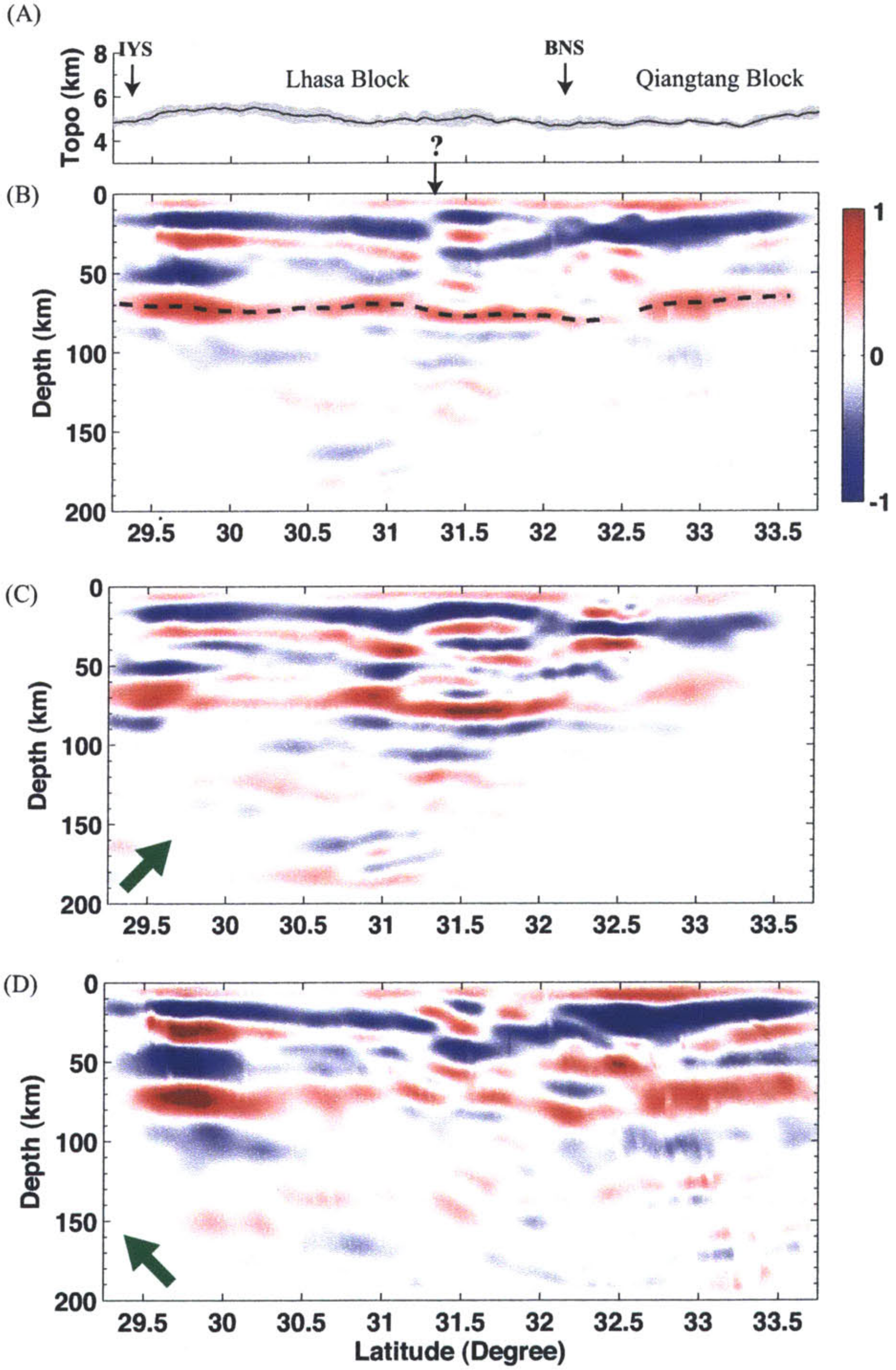


Figure 5-16 Hi-CLIMB RTM image with the curvelet interpolations. (A) Surface topography along the *x*-axis. The solid line is the mean elevation and the gray shading denotes the standard deviation across a swath of about 50 km in width. (B) Stacked image of SE and NW images. The Moho is delineated by a dashed line. (C) SE image contributed from the southeastern events. (D) NW image contributed from the northwestern events.

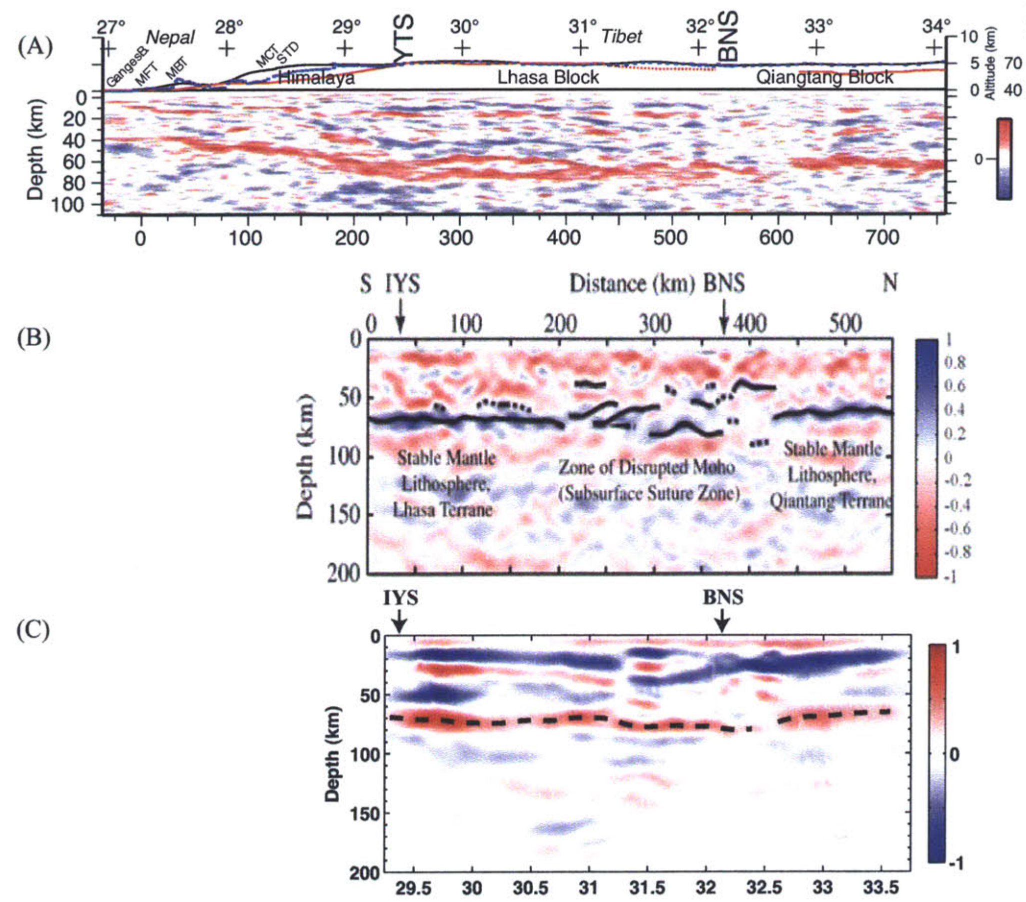


Figure 5-17 Comparison with previous studies on Hi-CLIMB data. (A) Receiver functions image (combination of *Ps*, *PpPs* and *PpSs* images) (from *Nábělek et al.* [2009]). All Hi-CLIMB stations are exploited for the CCP stacking. All depths are relative to sea level. (B) Migrated image by Gaussian beam method (from *Nowack et al.* [2010]). Note the blue (red) represents interfaces with increasing (decreasing) impedance, which is opposite to the convention in receiver function studies. (C) Passive source RTM image in this study.

Chapter 6

Multiple Seismic Reflectors in Earth's Lowermost Mantle[†]

The modern view of Earth's lowermost mantle considers a D" region of enhanced seismologically inferred heterogeneity bounded by the core mantle boundary (CMB) and an interface some 150-300 km above it, with the latter often attributed to the post-perovskite phase transition (in $MgSiO_3$). Deep-Earth exploration with *ScS* and *SKKS* wavefields, probing this region from above and below, respectively, suggest that this view be modified. Inverse scattering of *ScS* and *SKKS* data reveals seismic reflectors above the conventional D" region, beneath Central America and East Asia, two areas known for subduction of oceanic plates deep into Earth's mantle. The occurrence of multiple interfaces is inconsistent with expectations from a thermal response of a single

---

[†] (This chapter has been submitted as) Shang, X., S.-H. Shim, M. V. de Hoop and R. D. van der Hilst (2013), Multiple Seismic Reflectors in Earth's Lowermost Mantle, *Proc. Natl. Acad. Sci.*

isochemical post-perovskite transition but can be explained with post-perovskite transitions in differentiated slab materials. Our results imply that the lowermost mantle is more complex than hitherto thought and imply the presence of interfaces and compositional heterogeneity beyond the D" region *sensu stricto*.

Key words: lowermost mantle, post-perovskite transitions, multiple seismic reflectors, compositional heterogeneity

6.1 Introduction

The lowermost mantle, extending several hundred kilometers above the ~2,900 km deep core mantle boundary (CMB), is of considerable interest because it comprises the boundary layer of thermo-chemical mantle convection across which heat is conducted from the core into the mantle. Almost three decades after the detection of an interface some 150-300 km above the CMB [*Lay and Helmberger*, 1983], the so-called D" region below it is still a challenging target for cross-disciplinary research. Both the seismic discontinuity that marks the top of the D" region [*Lay and Helmberger*, 1983] and the heterogeneity below it [*Wysession et al.*, 1998; *Garnero*, 2000] have been attributed to a perovskite (Pv) to post-perovskite (pPv) transition in the dominant mantle silicate [*Murakami et al.*, 2004; *Oganov and Ono*, 2004; *Shim et al.*, 2004; *Hernlund et al.*, 2005; *Hutko et al.*, 2008], and this association has inspired estimation of temperatures above and heat flux across the CMB [*Lay et al.*, 2006; *van der Hilst et al.*, 2007]. The D" interface remains enigmatic, however, and recent high pressure–temperature experiments suggest that seismic observations concerning its depth and thickness are inconsistent with those expected for a pPv transition unless the chemical composition of the regions where they occur differs significantly from standard bulk composition models such as pyrolite [*Catalli et al.*, 2009; *Andrault et al.*, 2010; *Grocholski et al.*, 2012]. The transition thickness can be reduced through a nonlinearity in the phase fraction profile [*Catalli et al.*, 2009] and lattice preferred orientation of pPv [*Ammann et al.*, 2010], but the pPv transition pressure in pyrolite may be too high for it to occur in the lower mantle [*Grocholski et al.*, 2012]. Candidate compositions for a seismically detectable pPv transition at pressures less than the CMB include mid-oceanic ridge basalt (MORB) and harzburgite components of differentiated oceanic lithosphere transported to the lowermost mantle by subduction. Furthermore, silica may transform from modified stishovite to seifertite in Si-rich parts of the lowermost mantle [*Grocholski et al.*, 2013]. Inspired by these results, we search for multiple interfaces in and above the conventional D" region using seismic waves that sample the lowermost mantle.

6.2 Deep Earth exploration seismology

For large scale seismic exploration of the Earth's lowermost mantle we adapted a 3-D inverse scattering technique – a generalized Radon transform (GRT) – from its original use in controlled-source hydrocarbon exploration to imaging with *ScS* and *SKKS* waves emitted by naturally occurring earthquakes (Fig. 6-1, left). The GRT can extract subtle signal from large volumes of waveform data and facilitates the discovery of hitherto unknown structures because it does not rely on a priori assumptions about the location or shape of geological targets. The method applied here, an improved version of what we used in our previous studies [*Wang et al.*, 2006; *Wang et al.*, 2008], is summarized in the Supporting Information (SI, Text S1).

The *ScS* and *SKKS* waveform data are entirely independent in that (i) they are associated with different source-receiver combinations, (ii) sample the mantle (and core) along different propagation paths, (iii) arrive in different time windows, and (iv) concern different wave polarizations and, thus, sensitivies to (radial) anisotropy. We have shown that despite these differences they yield similar images of the first-order structural features in the lowermost 400 km of the mantle beneath Central America [*Wang et al.*, 2006; *Wang et al.*, 2008], where inferred depth variations of the presumed D" reflector correlate with tomographic wave speed anomalies [*van der Hilst et al.*, 2007; *Sidorin et al.*, 1999; *Sun et al.*, 2006].

We search for impedance contrasts up to 600 km above the CMB beneath two geographical areas that are well sampled by *ScS* or *SKKS* (Fig. 6-1, left) and where subduction has occurred for hundreds of millions of years. The first area is Central America (Fig. 6-1, upper right; Fig. 6-2, right) which has long been a type locality for studies of deep subduction of the Farallon plate [*van der Hilst et al.*, 1997; *Li et al.*, 2008; *Grand et al.*, 1997; *Grand*, 2002; *Ribe et al.*, 2007] and D" imaging with *ScS* waves [*Wysession et al.*, 1998; *Garnero*, 2000; *van der Hilst et al.*, 2007; *Thomas et al.*, 2004; *Hutko et al.*, 2006]. The second is East Asia, which is well sampled by *SKKS* waves (Fig. 6-1, lower right), and where tomography [*van der Hilst et al.*, 1997; *Li et al.*, 2008;

Grand et al., 1997; *Grand*, 2002] reveals a large high-speed anomaly (Fig. 6-2, left), presumably produced by deep subduction of the Izanagi and Pacific plates from the east and the Tethys and Indo-Australian plates from the south [*van der Hilst et al.*, 1997; *Grand et al.*, 1997].

For Central America we used (approximately) 130,000 *ScS* traces from 1,900 earthquakes (m_b>5.0; 1990-2009) recorded at one or more of a total of 2,700 seismograph stations (Fig. 6-1, top center). For Asia we used 120,000 *SKKS* traces from 11,000 events recorded at 1,700 stations (Fig. 6-1, bottom center). The range of epicentral distances is 0–80° for *ScS* and 100–180° for *SKKS*. The data used (~20% of all available traces) passed selection criteria based on signal-to-noise and multi-channel cross correlation values. Source signatures are estimated through principal component analysis (PCA) and removed from the data through Wiener deconvolution [*Chen et al.*, 2010; *Rondenay et al.*, 2005]. With GRT we estimate (from scattered energy) elasticity contrasts at nodes of a 3-D grid (10 km vertical, 1° lateral spacing), and spatial alignment of such contrasts indicates the presence of an interface. We use a tomographic model [*Grand*, 2002] to correct for mantle heterogeneity; other models may induce slight differences but these are too small to be of concern here. More information about data selection and processing is provided in SI (Text S2).

6.3 Structural complexity in the lowermost mantle

We illustrate the main structural features in the lowermost mantle beneath the regions under study by means of 2D (vertical) GRT sections through 3D image volumes. The nearly 1,500 km long Central America section (Fig. 6-3A) cuts across the main (tomographically inferred) high wavespeed anomaly (Fig. 6-2, right). This segment parallels section A-A' in our previous study [*van der Hilst et al.*, 2007], but it is shorter because detection of weak structures above D" requires more stringent imaging conditions and sampling criteria than was necessary for the imaging of CMB and D". The first of the East Asia sections shown here (B-B'; Fig. 6-3B) also cuts mainly across seismically fast regions, whereas the other (C-C', Fig. 6-3C) samples slow regions as well. In these profiles, black (red) pulses indicate positive (negative) impedance contrasts with increasing depth; their amplitudes are normalized (with respect to the CMB reflection) and only a qualitative indicator of the reflector strength. The background colors depict tomographic variations in shear wavespeed, which are used to construct the images.

Both beneath Central America and East Asia, well-aligned black pulses mark the CMB as well as a laterally continuous interface labeled 'X' (fat magenta lines, Fig. 6-3). Consistent with previous results [*Hutko et al.*, 2008; *van der Hilst et al.*, 2007], 'X' occurs some 250-300 km above the CMB beneath Central America (Fig. 6-3A), but beneath East Asia, where such a feature has not been imaged before, it is positioned closer to the CMB. In all sections, the data also reveal a weaker, and previously unknown, shallower interface (hereinafter 'Y'), some 450-500 km above the CMB. Stacking across the sections suppresses scatterers that do not form laterally coherent structures and reveals clearly the CMB, X, and Y as the main interfaces (Fig. 6-3, right).

Beneath Central America (Fig. 6-3A) ‘Y’ is laterally intermittent but clearly visible over at least 1,000 km horizontal distance. The data reveal impedance contrasts between CMB and ‘X’, as described before [*Hutko et al.*, 2008; *van der Hilst et al.*, 2007], and suggest that one or more wavespeed drops (red pulses, labeled ‘z’) occur between ‘X’ and ‘Y’. Locally, ‘z’ may look like a sidelobes of nearby positive pulses, but in most places it clearly appears as a separate signal. The Asia section across the fast anomaly (Fig. 6-3B) is qualitatively similar, with a weak, laterally intermittent interface visible some 200 km above ‘X’. Also here, negative pulses appear between CMB and ‘X’ and between ‘X’ and ‘Y’ (in particular around 300 km above CMB in the right half of the section). Fig. 6-3C suggests that the character of these structures changes when moving from (tomographically inferred) high to low wavespeeds in the lowermost mantle: interface ‘Y’ is laterally continuous in the easternmost 1,200 km “fast” part of Section C-C’, intermittent in the center (between 1,000 and 1,500 km horizontal distance), and absent in the seismically ‘slow’ region further west. In the latter, no coherent scattering are visible above interface ‘X’, located here ~200 km above CMB.

6.4 Phase transitions in differentiated subducted lithosphere?

Our application of modern imaging techniques to ever growing data sets confirms the widespread presence of a positive impedance contrast 150-300 km above the CMB beneath Central America and maps it for the first time beneath East Asia. Consistent with previous results, we interpret this horizon ‘X’ as the top of the D" region and associate it with the post-perovskite (Pv→pPv) transition. For reference, we also show (black dashes) the hypothetical phase boundary predicted by *Sidorin et al.* [1999], who assumed that (i) tomographically inferred wavespeed variations have a thermal origin, (ii) a (pressure-induced) mineralogical boundary exists (with a positive pressure-temperature dependence of 6 MPa/K), and (iii) this boundary can be extrapolated globally [*Sidorin et al.*, 1999; *Sun et al.*, 2006]. This prediction is hereafter referred to as the "thermal model". Along the Central America section (A-A') and in the western part of Asia section C-C' interface X coincides with the hypothetical phase boundary expected from the thermal model. In contrast, it does not correlate with thermal predictions in the high wavespeed parts of the Asia sections (e.g., B-B').

It is also beginning to discover hitherto unknown structures in the lowermost mantle. The detection of laterally continuous scatter surfaces above the D" interface may be at the edge of current resolution, but tests with synthetic data demonstrate that it cannot be attributed to noise, multiple scattering near sources or receivers (such as depth phases), reverberations within the D" layer (SI, Fig. 6-S2), or multiples of *SKKS* (that is, SK_nS, with n>2) [*Wang et al.*, 2008]. If 'X' marks the top of D", the images suggest that interfaces – and by implication compositional or phase boundaries – exist several hundred km above the D" region.

Our first-order observations – (i) poor correlation between tomographic wave speed variations and depths to the D" discontinuity and (ii) existence of multiple reflectors – are inconsistent with a single (pressure induced, temperature controlled) phase transition in magnesium silicate ($MgSiO_3$) perovskite in a compositionally homogeneous mantle. Forward (Pv→pPv) and reverse (pPv→Pv) transitions can produce multiple interfaces in the steep thermal gradient near the CMB [*Hernlund et al.*, 2005; *Hutko et al.*, 2008; *van der Hilst et al.*, 2007] but not several hundreds of km above it.

With hundreds of millions of years of subduction along the eastern seaboard of Asia and beneath Central America, it is possible that some of the multiple reflectors are due to buckling of slabs above the CMB [*Ribe et al.*, 2007], as has been suggested to occur beneath Central America [*Sun et al.*, 2006; *Thomas et al.*, 2004]. Another explanation involving subduction concerns the presence of material that is chemically distinct from a pyrolitic bulk composition. Variations in mineralogy and iron and aluminum content can influence the propagation speed of seismic waves [*Kiefer et al.*, 2002; *Jackson et al.*, 2004; *Tsuchiya and Tsuchiya*, 2006] as well as the pPv transition depth [*Catalli et al.*, 2009; *Grocholski et al.*, 2012; *Tateno et al.*, 2005] and detectability [*Grocholski et al.*, 2012]. Of particular interest are predictions from experimental mineral physics that in a heterogeneous mixture of harzburgite, basalt, and bulk mantle the Pv→pPv transition in the MORB fraction can occur several hundreds of kilometers above the transitions in either the harzburgite components or pyrolite (Fig. 6-4) [*Grocholski et al.*, 2012]. The shallower pPv transition would be weak because MORB contains much less magnesium silicate (30%) than harzburgite or pyrolite (60–80%) [*Hirose et al.*, 2005] and because the pPv transition depth interval in MORB is greater than harzburgite (but still smaller than in pyrolite [*Grocholski et al.*, 2012]).

Because of complex mineralogy and phase chemistry, uncertainties in absolute pressure in mineral physics data [*Shim*, 2008], controversy in seismic velocities of bulk mantle composition (pyrolite) and (mixtures of) recycled materials (MORB, harzburgite) [*Stixrude and Lithgow-Bertelloni*, 2012; *Tsuchiya*, 2011], and the fact that weaker impedance contrasts are only now beginning to emerge as robust features, one-to-one mapping of phase transitions and seismological boundaries is still premature. But combining evidence from mineral physics and seismic imaging (Fig. 6-4), we suggest that 'X' marks the pPv transition either in average mantle (if effects from LPO and element partitioning decrease pPv transition thickness to within detectable limits) or in the harzburgitic component of differentiated subducted lithosphere, and that 'Y' marks the transition in the subducted MORB component. Negative impedance contrasts ('z' and closer to CMB in Fig. 6-3) may reflect transformations in silica [*Hirose et al.*, 2005], local existence of partial melt [*Liebske and Frost*, 2012], or reverse pPv transformations in basalt, harzburgite, or bulk mantle. These are exciting targets of future joint mineral physics research and deep Earth exploration seismics and may lead to further revision of the canonical view of a lowermost mantle with compositional and structural heterogeneity restricted to the D" layer.

6.5 Methods

We used a generalized Radon transform (GRT) to recover unknown elastic reflectors in the lowermost mantle from the scattered *ScS* and *SKKS* wavefields (SI Text S1, Fig. 6-S1). In essence, the GRT maps (as in reverse time migration) the scattered seismic wavefield (recorded at the surface) back to sub-surface contrasts in elasticity. For each spatial node of a 3D grid, GRT exploits data redundancy through the integration of waveform data over a wide range of scattering angles and azimuth. In theory, point scatterers can be resolved in the Rayleigh diffraction limit (which depends on frequency), but in practice spatial resolution depends on how a subsurface point is illuminated (that is, the wave slownesses and the range of scatter angles over which data are integrated), which depends on source-receiver distribution. To ensure robustness, we require that target points be sampled from a sufficient range of angles (> 15°) around the stationary point (specular reflection), which enables us to resolves structure at lateral scales of 500 km or larger. The GRT imaging does not rely on a priori assumptions about the location or shape of geological targets, which facilitates discovery of hitherto unknown structures. More details of the method and data preprocessing are given in *SI* Texts S1 and S2. For the mineral physics data, the absolute pressure (or depth) is uncertain by at least ±5 GPa (or ±100 km) [*Shim*, 2008] and depends on the pressure scale used. We note that the use of the gold pressure scale by Tsuchiya [*Tsuchiya*, 2003], originally used in *Grocholski et al.* [2012], gives depths in between those inferred from the gold scales used in Fig. 6-4B. The relative pressure scale (or difference in pressure or depth) is better constrained (±1 GPa), and all mineral physics data shown in Fig. 6-4B are constrained using the same pressure scale (gold). Therefore, depth differences among the phase boundaries in Fig. 6-4B are more reliable for comparison with seismic data than absolute depths.

Acknowledgments

This research was supported by US National Science Foundation (NSF) grants EAR-0757871 (CSEDI program), DMS-0724778 (CMG program), and EAR-1301813 (to S.-

H.S.). All waveform data used in this study are retrieved from Data Management Center of the Incorporated Research Institutions for Seismology (IRIS-DMC). We acknowledge the pioneering work by Dr. Ping Wang (formerly at MIT, now at CGG Veritas, Houston) on GRT imaging with *SKKS* and *ScS* data. The research was conducted by X.S., under supervision of R.v.d.H and M.V.d.H. S.-H.S. assisted with the interpretation.

References

Ammann, M., J. Brodholt, J. Wookey, and D. Dobson (2010), First-principles constraints on diffusion in lower-mantle minerals and a weak D" layer, *Nature*, *465*(7297), 462-465.

Andrault, D., M. Munoz, N. Bolfan-Casanova, N. Guignot, J.-P. Perrillat, G. Aquilanti, and S. Pascarelli (2010), Experimental evidence for perovskite and post-perovskite coexistence throughout the whole D" region, *Earth Planet. Sci. Lett.*, *293*(1), 90-96.

Catalli, K., S. H. Shim, and V. Prakapenka (2009), Thickness and Clapeyron slope of the post-perovskite boundary, *Nature*, *462*(7274), 782-U101.

Chen, C. W., D. Miller, H. Djikpesse, J. Haldorsen, and S. Rondenay (2010), Array-conditioned deconvolution of multiple-component teleseismic recordings, *Geophys. J. Inter.*, *182*(2), 967-976.

Dziewoński, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*(4), 297-356.

Garnero, E. (2000), Lower mantle heterogeneity, *Annu. Rev. Earth Planet. Sci.*, *28*, 509-537.

Grand, S. P., R. D. van der Hilst, and S. Widiyantoro (1997), High resolution global tomography: a snapshot of convection in the Earth, *GSA Today*, *7*(4).

Grand, S. P. (2002), Mantle shear-wave tomography and the fate of subducted slabs, *Phil. Trans. R. Soc. A*, *360*(1800), 2475-2491.

Grocholski, B., K. Catalli, S.-H. Shim, and V. Prakapenka (2012), Mineralogical effects on the detectability of the postperovskite boundary, *Proc. Natl. Acad. Sci.* , *109*(7), 2275-2279.

Grocholski, B., S.-H. Shim, and V. Prakapenka (2013), Stability, metastability, and elastic properties of a dense silica polymorph, seifertite implications for Shergotty meteorite and lowermost mantle structures., *J. Geophy. Res.*, in revision.

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*(7035), 882-886.

Hirose, K., N. Takafuji, N. Sata, and Y. Ohishi (2005), Phase transition and density of subducted MORB crust in the lower mantle, *Earth Planet. Sci. Lett.*, *237*(1), 239-251.

Hirose, K., N. Sata, T. Komabayashi, and Y. Ohishi (2008), Simultaneous volume measurements of Au and MgO to 140GPa and thermal equation of state of Au based on the MgO pressure scale, *Phys. Earth Planet. Inter.*, *167*(3), 149-154.

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*(7091), 333-336.

Hutko, A. R., T. Lay, J. Revenaugh, and E. J. Garnero (2008), Anticorrelated seismic velocity anomalies from post-perovskite in the lowermost mantle, *Science*, *320*(5879), 1070.

Jackson, J. M., J. Zhang, and J. D. Bass (2004), Sound velocities and elasticity of aluminous $MgSiO_3$ perovskite: Implications for aluminum heterogeneity in Earth's lower mantle, *Geophys. Res. Lett.*, *31*(10).

Jamieson, J. C., J. N. Fritz, and M. H. Manghnani (1982), *High-pressure research in geophysics* edited by S. Akimoto and M. H. Manghnani, pp. 27–48, Springer.

Kiefer, B., L. Stixrude, and R. Wentzcovitch (2002), Elasticity of (Mg,Fe)SiO_3-perovskite at high pressures, *Geophys. Res. Lett.*, *29*(11), 34-31.

Lay, T., and D. Helmberger (1983), A lower mantle S-wave triplication and the shear velocity structure of D", *Geophys. J. R. Astron. Soc*, *75*, 799-838.

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*(5803), 1272-1276.

Li, C., R. D. van der Hilst, E. R. Engdahl, and S. Burdick (2008), A new global model for P wave speed variations in Earth's mantle, *Geochem. Geophy. Geosyst.*, *9*(5), Q05018.

Liebske, C., and D. J. Frost (2012), Melting phase relations in the MgO-$MgSiO_3$ system between 16 and 26GPa: Implications for melting in Earth's deep interior, *Earth Planet. Sci. Lett.*, *345*, 159-170.

Murakami, M., K. Hirose, K. Kawamura, N. Sata, and Y. Ohishi (2004), Post-perovskite phase transition in $MgSiO_3$, *Science*, *304*(5672), 855-858.

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*(6998), 445-448.

Ribe, N., E. Stutzmann, Y. Ren, and R. van der Hilst (2007), Buckling instabilities of subducted lithosphere beneath the transition zone, *Earth Planet. Sci. Lett.*, *254*(1-2), 173-179.

Rondenay, S., M. G. Bostock, and K. M. Fischer (2005), Multichannel inversion of scattered teleseismic body waves: Practical considerations and applicability, *Geophys. Monograph-American Geophys. Union*, *157*.

Shim, S., T. Duffy, R. Jeanloz, and G. Shen (2004), Stability and crystal structure of $MgSiO_3$ perovskite to the core-mantle boundary, *Geophys. Res. Lett.*, *31*, 1029-1039.

Shim, S.-H. (2008), The post-perovskite transition, *Annu. Rev. Earth Planet. Sci.*, *36*, 569-599.

Sidorin, I., M. Gurnis, and D. V. Helmberger (1999), Evidence for a ubiquitous seismic discontinuity at the base of the mantle, *Science*, *286*(5443), 1326-1331.

Stixrude, L., and C. Lithgow-Bertelloni (2012), Geophysics of Chemical Heterogeneity in the Mantle, *Annu. Rev. Earth Planet. Sci.*, *40*, 569-595.

Sun, D., T. R. A. Song, and D. Helmberger (2006), Complexity of D" in the presence of slab-debris and phase changes, *Geophys. Res. Lett.*, *33*(12), L12S07.

Tateno, S., K. Hirose, N. Sata, and Y. Ohishi (2005), Phase relations in $Mg_3Al_2Si_3O_{12}$ to 180 GPa: Effect of Al on post-perovskite phase transition, *Geophys. Res. Lett.*, *32*(15).

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

Tsuchiya, T. (2003), First-principles prediction of the PVT equation of state of gold and the 660-km discontinuity in Earth's mantle, *J. Geophys. Res.*, *108*(B10), 2462.

Tsuchiya, T., and J. Tsuchiya (2006), Effect of impurity on the elasticity of perovskite and postperovskite: Velocity contrast across the postperovskite transition in (Mg,Fe, Al)(Si,Al)O_3, *Geophys. Res. Lett.*, *33*(12).

Tsuchiya, T. (2011), Elasticity of subducted basaltic crust at the lower mantle pressures: Insights on the nature of deep mantle heterogeneity, *Phys. Earth Planet. Inter.*, *188*(3), 142-149.

van der Hilst, R., S. Widiyantoro, and E. Engdahl (1997), Evidence for deep mantle circulation from global tomography, *Nature*, *386*(6625), 578-584.

van der Hilst, R., M. De Hoop, P. Wang, S. H. Shim, P. Ma, and L. Tenorio (2007), Seismostratigraphy and thermal structure of Earth's core-mantle boundary region, *Science*, *315*(5820), 1813-1817.

Wang, P., M. de Hoop, R. van der Hilst, P. Ma, and L. Tenorio (2006), Imaging of structure at and near the core mantle boundary using a generalized radon transform: 1. Construction of image gathers, *J. Geophys. Res.*, *111*.

Wang, P., M. V. de Hoop, and R. D. van der Hilst (2008), Imaging the lowermost mantle (D") and the core-mantle boundary with SKKS coda waves, *Geophys. J. Inter.*, *175*(1), 103-115.

Wysession, M. E., T. Lay, J. Revenaugh, Q. Williams, E. Garnero, R. Jeanloz, and L. H. Kellogg (1998), The D" discontinuity and its implications, in *The core-mantle boundary region*, edited by M. Gurnis, M. E. Wysession, E. Knittle and B. A. Buffett, pp. 273-298, American Geophysical Union.

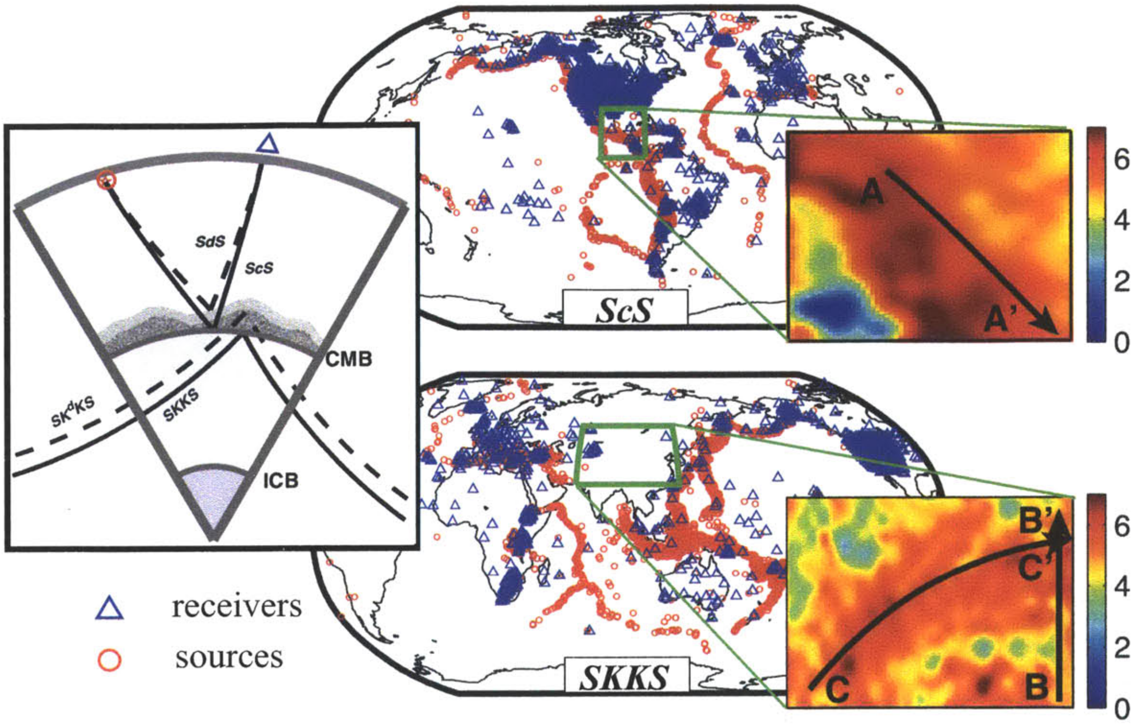


Figure 6-1 Left: Generic ray geometry of *ScS* (blue paths) and *SKKS* (red paths) phases, which sample the lowermost mantle from above and below, respectively. Solid lines depict *ScS* and *SKKS* reflections at the Core Mantle Boundary (CMB); dashed lines depict possible paths of scattering at structure above the CMB, which is used to generate the sections shown in Fig. 6-3. Center and Right: Top: Distribution of epicenters (red circles) and stations (blue triangles) that yield the *ScS* data used in the construction of the common image point gathers. The range of epicentral distances is 0–80°. The green rectangle depicts the study region in Central America (5-30°N; 80- 105°W). Inset: natural logarithm of the number of *ScS* midpoints in 2°× 2° geographical bins. Bottom: Same, but for *SKKS* in East Asia (25–55°N; 65–125°E). The range of epicentral distances for *SKKS* is 100–180°.

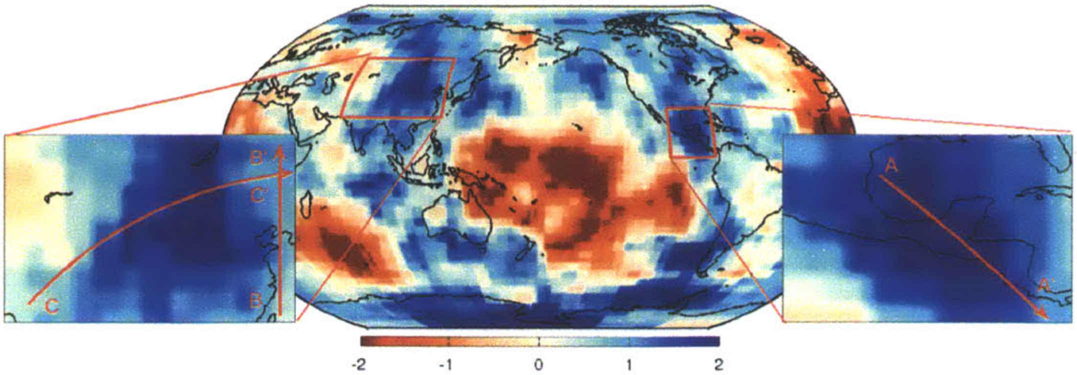


Figure 6-2 Tomographically inferred lateral variations (in %) in shear speed at approximately 200 km above the core mantle boundary [*Grand*, 2002]. Red boxes indicate regions under study: East Asia (left) and Central America (right). Red arrows in these boxes depict locations of cross sections A-A' (in Central America) and B-B' and C-C' (in East Asia), for which the inverse scattering results are shown in Fig. 6-3.

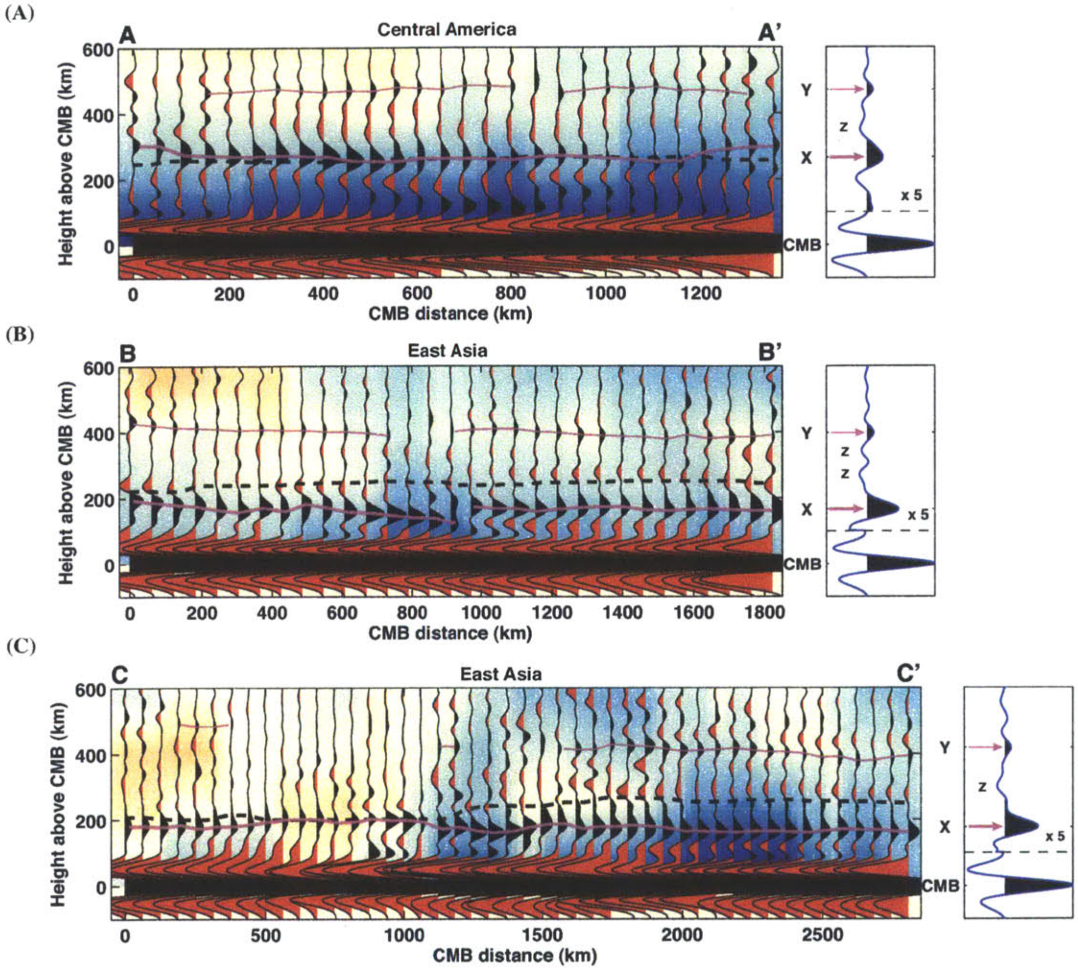


Figure 6-3 Reflectivity profiles superimposed on smooth (tomographically inferred) variations in shear wavespeed [*Grand*, 2002]—scale as in Fig. 6-2. (A) Section A-A' across Central America: [24.5°N, 96.6°W] to [6.5°N, -83°W]; (B) Section B-B' across East Asia: [25°N, 123°E] to [55°N, 123°E]; (C) Section C-C' across East Asia: [27°N, 70°E] to [50°N, 123°E]. Black (red) pulses depict positive (negative) impedance contrasts. The magenta lines depict coherent reflectors inferred from data (thick dashes depict interface 'X'; thin dashes depict interface 'Y'); black dashes depict the prediction by *Sidorin et al.* [1999] from purely thermal considerations. At the right: stacks (along interface 'X') showing the main (positive) impedance contrasts CMB, 'X', and 'Y', as well as minor negative impedance contrasts located between 'X' and 'Y' (labeled 'z'). In these stacks, the part above 100 km above CMB (the horizontal dashed line) is amplified by a factor of 5. Upon stacking, laterally coherent reflectors are enhanced and incoherent scattering suppressed.

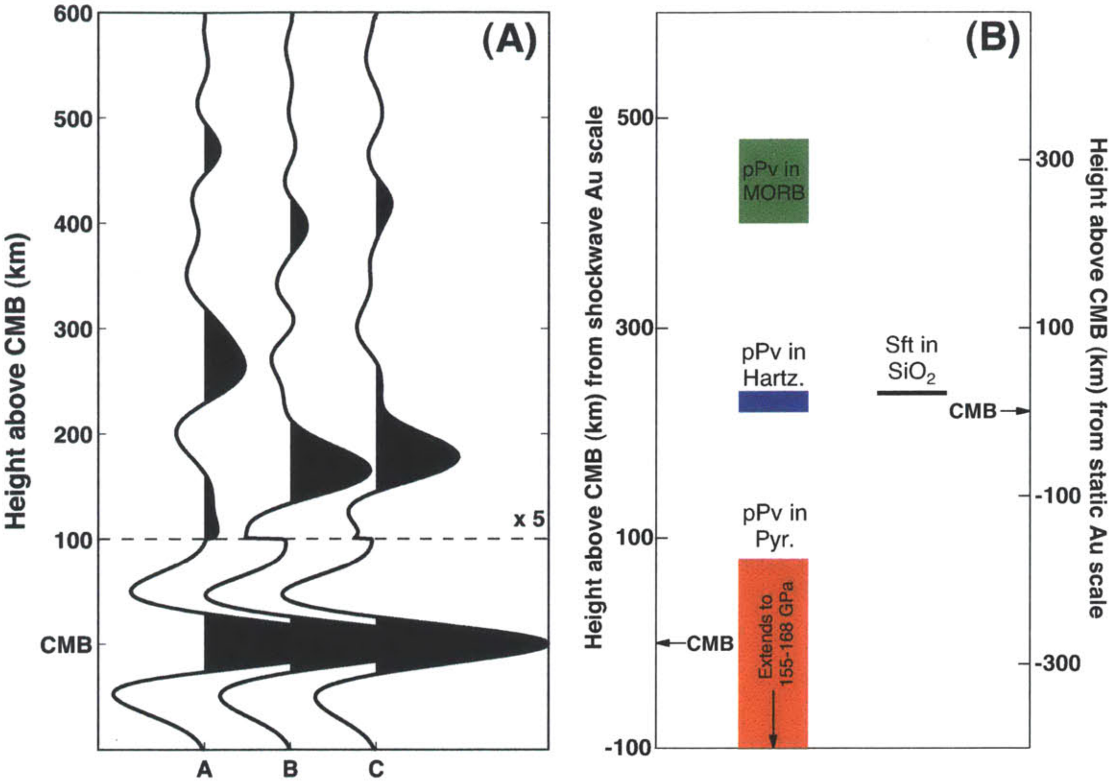


Figure 6-4 Left: stacks along interface 'X' in sections A-A', B-B', and C-C' (shown in Fig. 6-3). Right: depth ranges for post-perovskite transitions in mid ocean ridge basalt (MORB), harzburgite (Harz), and pyrolite (Pyr), after *Grocholski et al.* [2012], and transitions in silica from modified stishovite to seifertite (Sft), after *Grocholski et al.* [2013], at ~2,500 K. Plotted are the lower bounds of the pPv transition thickness. The pressures at the phase boundaries are calculated using either a shock wave (gold) scale from *Jamieson et al.* [1982], scale bar on the left of (B), or a static compression (gold) scale from *Hirose et al.* [2008], on the right. Pressure was converted into depth using PREM [*Dziewoński and Anderson*, 1981]. Absolute pressure (depth) is uncertain by at least ±5 GPa (±100 km), but pressure differences are constrained better (±1 GPa). Given these uncertainties, the depth difference between 'X' and 'Y' agrees remarkably well with the difference in pPv transition depth between MORB and harzburgite.

Supporting Information

SI Text S1: Methodology

Inverse scattering in seismology refers to estimating the characteristics of an (subsurface) object from elastic waves that scatter from it. In the generalized Radon transform (GRT) used here we use the Born approximation (that is, single scattering) and assume that the scatterers (including interfaces) can be represented as perturbations relative to a smooth back ground medium [*Miller et al.*, 1987; *Beylkin*, 1990; *de Hoop and Bleistein*, 1997]. For this purpose, the medium parameter, for instance wave speed $c(\mathbf{x})$, can be decomposed into two parts $c(\mathbf{x}) = c_0(\mathbf{x}) + \delta c(\mathbf{x})$, where $\delta c(\mathbf{x})$ is a perturbation relative to a smoothly varying background medium $c_0(\mathbf{x})$. Accordingly, the wave field $u(\mathbf{x})$ can be divided into two parts $u(\mathbf{x}) = u_0(\mathbf{x}) + \delta u(\mathbf{x})$, where $u_0(\mathbf{x})$ is the direct wave due to propagation in $c_0(\mathbf{x})$ and $\delta u(\mathbf{x})$ the scattered wave due to $\delta c(\mathbf{x})$. The background smooth model can be estimated from reference Earth models and seismic tomography. Inverse scattering maps the scattered wave field $\delta u(\mathbf{x})$ back into images of perturbation of medium properties $\delta c(\mathbf{x})$.

The scattered wavefield $\delta u(\mathbf{x})$ depends on the position of imaging point $\mathbf{y}$, source $\mathbf{x}^s$ and receiver $\mathbf{x}^r$ (Fig. 6-S1). For a given $(\mathbf{y};\mathbf{x}^s,\mathbf{x}^r)$, we define the slowness vector $\mathbf{p}^s(\mathbf{y})$ for the incident ray from source to the imaging point, as well as $\mathbf{p}^r(\mathbf{y})$ from the receiver side. These two slowness vector can define migration dip $\mathbf{v}^m(\mathbf{y}) = \mathbf{p}^m(\mathbf{y})/||\mathbf{p}^m(\mathbf{y})||$, with $\mathbf{p}^m(\mathbf{y}) =$

$\mathbf{p}^s(\mathbf{y}) + \mathbf{p}^r(\mathbf{y})$ and scattering angle θ and azimuth ψ (Fig. 6-S1).

In order to do multi-scale analysis and avoid caustics, the recorded wavefield $\delta u(\mathbf{y}; \mathbf{x}^s, \mathbf{x}^r)$ can be mapped into angle domain $\delta u(\mathbf{y}; \mathbf{v}^m, \theta, \psi)$ [*Wang et al.*, 2006; *Wang et al.*, 2008; *de Hoop and Bleistein*, 1997], with scatter angle θ and azimuth ψ. The structural reflectivity at $\mathbf{y}$ can be approximated as

$$I(\mathbf{y}) = \int \delta u(\mathbf{y}; v^m, \theta, \psi) \left\| \mathbf{p}^m(\mathbf{y}) \right\|^3 dv^m \, d\psi \, d\theta \tag{6.S1}$$

where data $\delta u(\mathbf{x})$ is integrated over scatter angles θ and azimuths ψ.

In practice, the spatial resolution of GRT depends on how an image point is illuminated (integration range of Eq. 6.S1). The imaging result would be biased by uneven source-receiver distribution (e. g. dominant earthquake direction, poor illumination coverage). In order to ensure robustness and mitigate inversion artifacts, we regularize it by two steps. Firstly, all image points are required to be sampled from a sufficient range of angles around the stationary point (specular reflection). For each vertical profile in Fig. 6-3 (the same latitude and longitude with different depth), we check the illumination coverage using the point on core-mantle boundary (CMB). The survived points must have specular reflected data (migration dip angle is zero) and a wide range of migration dip angle (here we use 15 degree). Secondly, the contribution from different earthquakes is balanced by simple ray-count normalization. After trace normalization, the ray count is proportional to the incident wave energy. Before summing over scatter angle θ in Eq. 6.S1, the ray count is calculated for each angle bin (3 degree) and each partial image is normalized by the ray count.

SI Text S2: Preprocessing of the Data

There are several steps in the selection and preprocessing of the data. First, for all data we remove the instrument response and rotate the data to radial and transverse components. We use transverse component (SH wave) for *ScS* data and radial component (SV wave) for *SKKS* data. Second, all data are band-passed between 10 and 50 seconds with a 4-pole Butterworth filter, and then normalized with respect to the reference phase (*ScS* or *SKKS*). Third, we discard data on the basis of a simple quality criterion obtained from multi-channel cross-correlation [*VanDecar and Crosson*, 1990]. For each earthquake, we organize the data in 10° (epicentral) distance bins, and from each seismic record in a bin we extract a 50 s time window around the theoretical *ScS* (or *SKKS*) arrivals. After energy normalization, we then cross correlate each trace with all others in the bin, which yields a qualitative estimate of the average correlation coefficient for each trace. Traces with averaged correlation coefficient lower than a set threshold (we use 0.6) are discarded. Roughly 20% of all data meets this quality criterion. The trace polarity can be corrected according to the sign of averaged correlation coefficient. After multichannel cross-correlation, traces are aligned with reference phases. Fourth, principle component analysis (PCA) is applied to estimate the source signature [*Rondenay et al.*, 2005], which is then removed from the data to enhance image resolution by Wiener deconvolution [*Chen et al.*, 2010]. The water level in Wiener deconvolution is selected automatically and adaptively based on the noise spectrum of array data. Finally, travel times are corrected for ellipticity [*Kennett and Gudmundsson*, 1996] and for 3D wavespeed variations using a tomographic model for mantle shear wave speed [*Grand*, 2002].

SI Text S3: Tests with Synthetic Data

Since the quality of the GRT imaging operator depends on the aperture of angles θ and azimuths ψ (SI Text S1, above) and the validity of the single-scattering approximation, tests with synthetic data are necessary to examine the effects of uneven sampling, noise, interfering main phases (such as depth phases), and multiple scattering (e.g., near source or receiver scattering and internal reverberations). In all tests described below we computed synthetic waveforms for the real source-receiver geometry (including focal depths) and focal mechanism obtained from Global CMT catalog. All synthetic seismograms are calculated with the WKBJ method [*Chapman*, 1978], using a radially stratified wave speed model (ak135 model [*Kennett et al.*, 1995]), on which a 3% *S* wave velocity jump is superimposed at 250 km above CMB. Before the inversion, all synthetic data are band-pass filtered between 5 to 50 s.

Figure 6-S2 shows the effects on the *SKKS* image gather of random noise (100% noise level), depth phases, and multiple reverberations. Even when the signal cannot be identified in most individual traces, the reflector at 250 km above CMB is clearly recovered after stacking over scattering angle (Fig. 6-S2a). The depth phase can produce small waveform distortion (Fig. 6-S2b). However, it does not produce significant artifacts in the image. The reverberation in the crust and D″ layer can also be a potential artifact source in that such phases could interfere with the coda wave of *SKKS*. Applying the GRT to the synthetic data with such signals shows that they do not in general

contaminate the image profile (Fig. 6-S2c,d).

SKKS multiples SK_nS (n>2), which reverberate within the liquid outer core, arrive later and could interfere with the reflections from lowermost mantle interfaces in time domain. In angle gathers, their slownesses differ from *SKKS* (and SKS^dSKS), therefore they can be distinguished by a clear residual moveout and suppressed by parabolic Radon transforms [*Wang et al.*, 2008].

References Cited in Supplementary Information

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

Chapman, C. (1978), A new method for computing synthetic seismograms. *Geophys. J. R. Astron. Soc. 54*, 481–518.

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

Miller, D., M. Oristaglio, and G. Beylkin (1987), A new slant on seismic imaging: Migration and integral geometry, *Geophysics, 52*(*7*), 943-964.

Kennett, B., E. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Inter. 122,* 108–124.

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

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

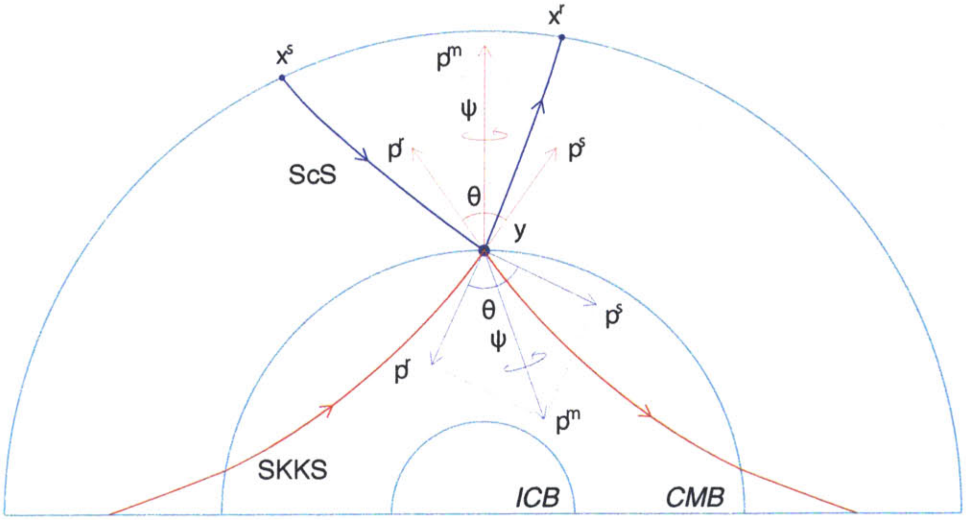


Figure 6-S1 Illustration of ray path geometry considered in the generalized Radon transform of *ScS* (blue paths) and *SKKS* (red paths) phases, for subsurface image point $\mathbf{y}$. Slowness vectors are given by $\mathbf{p}$. The slowness of the incident ray (from source $\mathbf{x}^s$ to $\mathbf{y}$) direction is $\mathbf{p}^s$, $\mathbf{p}^r$ denotes the scattered path (from $\mathbf{y}$ to receiver $\mathbf{x}^r$), and $\mathbf{p}^m = \mathbf{p}^s + \mathbf{p}^r$ defined the migration dip direction (which controls radial resolution). Scattering angle and azimuth are given by θ and ψ, respectively. CMB and ICB denote core-mantle boundary and inner-core boundary, respectively.

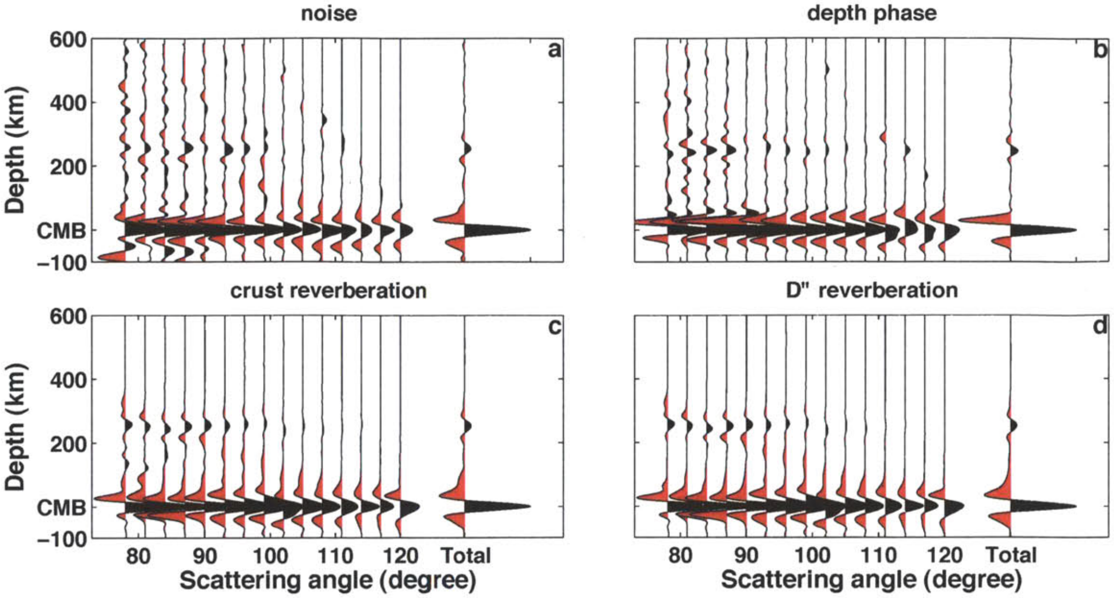


Figure 6-S2 Effects on the *SKKS* GRT image gathers of the presence of (a) white noise (100% additive noise), (b) depth phases, (c) multiple reverberation in the crust, and (d) multiple reverberation between D" discontinuity and CMB. In the tests, a 3% shear wave velocity jump at 250 km above CMB is superimposed on a layered model (ak135) [*Kennett et al.*, 1995].

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this thesis, we develop and apply inverse scattering transforms to construct images of the lithosphere and the lowermost mantle with broadband teleseismic data.

In **Chapter II** we present passive source reverse time migration (RTM), a new type of wave equation prestack depth migration, for subsurface imaging with multi-component data from dense arrays. In contrast to conventional methods using wave conversions, RTM does not rely on simplifying assumptions about the geometry of the geological structures of interest and can handle strong heterogeneities without difficulty. In passive source RTM, multi-component array data are first propagated backward by solving the elastic wave equation directly. After polarization separation, a modified cross correlation imaging condition between P and S wave constitutes is applied to obtain an inverse scattering transform. From synthetic experiments, it is evident that for complex geological structures RTM is superior to the traditional CCP receiver function stacking, provided that data from dense seismograph arrays are available.

A semi-automatic workflow of teleseismic data preprocessing for inverse scattering is proposed in **Chapter III**. After instrument response removal, the traces from one event are aligned with a reference phase by multi-channel cross-correlation. The noisy traces are marked and removed automatically in the same time. The polarity of traces is corrected according to the sign of cross-correlation coefficients. After alignment, the source signature is estimated by principle component analysis. Finally, the source signature is deconvolved by Wiener deconvolution, which is data-adaptive and superior to traditional water level deconvolution. Synthetic and real data examples show that this workflow is less subjective and more efficient for massive data preprocessing. Codes will be made available to the broader community.

Due to the absence of suitable dense array data, wavefield interpolation is a necessary step before RTM application. In **Chapter IV** we demonstrate the feasibility of sparsity promoting curvelet interpolation on undersampled teleseismic datasets. We investigate the effect of undersampling ratio, from 0.5 to 0.15, on typical teleseismic wavefield recovery. In the presence of noise, the Pareto curve (L-curve in geophysical literatures) is exploited to find the optimal trade-off between the data fit and sparsity of the solution. Both synthetic and real data examples demonstrate that for typical teleseismic acquisition geometry with 50% to 85% missing traces, the curvelet-based interpolation works remarkably well.

In **Chapter V** we apply passive source RTM to a sparsely sampled teleseismic dataset. Synthetic examples demonstrate that for a typical teleseismic imaging of lithosphere, dense sampling (e.g. ~2 km spacing for 1 Hz wave) is required to suppress spatial alias in RTM image. With the aid of curvelet interpolation, passive source RTM becomes to be feasible (e.g. 6-10 km spacing is enough for 1 Hz wave). Hi-CLIMB array in Tibet is used as a real data example of passive source RTM with curvelet interpolation. RTM image reveals clear and continuous Moho discontinuity across the Bangong-Nujiang suture (BNS) at the depth of about 70 km, which is not well imaged by CCP stacking and Gaussian beam migration in previous studies.

Due to the sparsity of seismographs globally, it is more appropriate to use a high frequency approximation of the inverse scattering. We use a generalized Radon transform (GRT) to probe the deep mantle structures. In **Chapter VI** we use about 130,000 *ScS* traces and 120,000 *SKKS* traces to image the lowermost mantle beneath central America and east Asia respectively, two areas known for subduction of oceanic plates deep into Earth's mantle. Inverse scattering of *ScS* and *SKKS* data reveals seismic reflectors above the conventional D" region beneath central America and east Asia. The occurrence of multiple interfaces is inconsistent with expectations from a thermal response of a single isochemical post-perovskite transition but can be explained with post-perovskite transitions in differentiated slab materials. Our results imply that the lowermost mantle is more complex than hitherto thought and that the presence of interfaces and compositional heterogeneity beyond the D" region *sensu stricto.*

7.2 Future Work

In this thesis, we mainly focus on the 2D application of passive source RTM. It is straightforward but not trivial to extend to 3D cases. With 2D areal dense networks on the surface rather than 1D linear arrays, subsurface structures can then be constrained in more details, without resorting to cumbersome 2.5D approximations. In teleseismic imaging, several key ingredients, such as the 3D fast curvelet transform [*Ying et al.*, 2005], a large-scale ℓ_1-norm optimizer [e.g. *van den Berg and Friedlander*, 2008], and a 3D elastic wave propagator [e.g. *Komatitsch and Tromp*, 1999; *Zhang et al.*, 2012] are all available. There are no theoretical obstacles but computational cost for 3D RTM application could be formidable.

To overcome the illumination problem in teleseismic imaging, partially due to uneven distribution of earthquakes, multiples can be incorporated in current RTM scheme. Due to distinct ray path and moveout signature from primary waves, multiples increase the illumination aperture and image resolution. In reflection seismology, surface related (and internal) multiples are employed for the imaging of the salt base and flank [e.g. *Berkhout and Verschuur*, 2006; *Malcolm et al.*, 2009]. In teleseismic study, people have begun to use it for crust imaging [e.g. *Tseng et al.*, 2009]. However, including sharp boundaries in the background model is still an on-going research topic for RTM.

Passive source RTM is suitable for scatterer detection, but the accuracy of the resulting images is dependent on the smooth background model. On the other hand, travel time tomography is sensitive to large-scale volumetric anomalies. In the future we aim to combine passive source RTM and tomography in an explicit (non-linear) joint inversion of teleseismic wavefields for wavespeed determination and interface detection and location.

References

Berkhout, A., and D. Verschuur (2006), Imaging of multiple reflections, *Geophysics*, *71*(4), SI209-SI220.

Komatitsch, D., and J. Tromp (1999), Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Inter.*, *139*(3), 806-822.

Malcolm, A. E., B. Ursin, and M. V. de Hoop (2009), Seismic imaging and illumination with internal multiples, *Geophys. J. Inter.*, *176*(3), 847-864.

Tseng, T.-L., W.-P. Chen, and R. L. Nowack (2009), Northward thinning of Tibetan crust revealed by virtual seismic profiles, *Geophys. Res. Lett.*, *36*(24), L24304.

van den Berg, E., and M. P. Friedlander (2008), Probing the Pareto frontier for basis pursuit solutions, *SIAM J. Sci. Comput.*, *31*(2), 890-912.

Ying, L., L. Demanet, and E. Candes (2005), 3D discrete curvelet transform, paper presented at Optics & Photonics 2005, International Society for Optics and Photonics.

Zhang, W., Z. G. Zhang, and X. F. Chen (2012), Three-dimensional elastic wave numerical modelling in the presence of surface topography by a collocated-grid finite-difference method on curvilinear grids, *Geophys. J. Inter.*, *190*(1), 358-378.