

## Imaging seismic velocity structure beneath the Iceland hot spot: A finite frequency approach

Shu-Huei Hung

Department of Geosciences, National Taiwan University, Taipei, Taiwan

Yang Shen

Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island, USA

Ling-Yun Chiao

Institute of Oceanography, National Taiwan University, Taipei, Taiwan

Received 12 November 2003; revised 13 May 2004; accepted 4 June 2004; published 11 August 2004.

[1] Tomographic models based on hypothetically infinite frequency ray interpretation of teleseismic travel time shifts have revealed a region of relatively low  $P$  and  $S$  wave speeds extending from shallow mantle to 400 km depth beneath Iceland. In reality, seismic waves have finite frequency bandwidths and undergo diffractive wave front healing. The limitation in ray theory leaves large uncertainties in the determinations of the magnitude and shape of the velocity anomaly beneath Iceland and its geodynamic implications. We developed a tomographic method that utilizes the banana-shaped sensitivity of finite frequency relative travel times from the paraxial kernel theory. Using available seismic data from the ICEMELT and HOTSPOT experiments, we applied the new method to image subsurface velocity structure beneath Iceland. Taking advantage that the sensitivity volume of broadband waveforms varies with frequency, we measured relative delay times in three frequency ranges from 0.03 to 2 Hz for  $P$  and 0.02 to 0.5 Hz for  $S$  waves. Given similar fit to data, the kernel-based models yield the root-mean-square amplitudes of  $P$  and  $S$  wave speed perturbations about 2–2.8 times those from ray tomography in the depths of 150–400 km. The kernel-based images show that a columnar low-velocity region having a lateral dimension of  $\sim 250$ –300 km extends to the base of the upper mantle beneath central Iceland, deeper than that resolved by the ray-based studies. The improved resolution in the upper mantle transition zone is attributed to the deeper crossing of broad off-path sensitivity of travel time kernels than in ray approximation and frequency-dependent wave front healing as an intrinsic measure of the distance from velocity heterogeneity to receivers. **INDEX TERMS:** 8180 Tectonophysics: Tomography; 7203 Seismology: Body wave propagation; 7218 Seismology: Lithosphere and upper mantle; 3260 Mathematical Geophysics: Inverse theory; 8121 Tectonophysics: Dynamics, convection currents and mantle plumes; **KEYWORDS:** finite frequency travel time, tomography, wave front healing, Fréchet kernel, Iceland

**Citation:** Hung, S.-H., Y. Shen, and L.-Y. Chiao (2004), Imaging seismic velocity structure beneath the Iceland hot spot: A finite frequency approach, *J. Geophys. Res.*, 109, B08305, doi:10.1029/2003JB002889.

### 1. Introduction

[2] Iceland and the adjacent Mid-Atlantic Ridge are characterized by topographic swell, an anomalously thick crust, and atypical geochemical signatures ([e.g., Schilling, 1973; White and McKenzie, 1989; Saunders et al., 1997; Darbyshire et al., 2000; Allen et al., 2002b; Foulger et al., 2003]. These features have been attributed hypothetically to the excess melt production of a hot, buoyantly upwelling mantle plume [e.g. Morgan, 1971; Ribe et al., 1995; Ito et al., 1996], although alternative interpretations have

lately received considerable attention [e.g., Korenaga and Kelemen, 2000; Foulger et al., 2001].

[3] Tomographic studies have revealed low compressional ( $P$ ) and shear ( $S$ ) wave speeds beneath Iceland at both regional and global scales [Tryggvason et al., 1983; Wolfe et al., 1997; Bijwaard and Spakman, 1999; Ritsema et al., 1999; Foulger et al., 2000; Zhao, 2001; Allen et al., 2002b; Montelli et al., 2004]. In particular, recent broadband seismic experiments on Iceland have provided high-resolution tomographic images of the upper mantle structure. Wolfe et al. [1997] inverted the ICEMELT travel time data set and found that a cylindrically shaped region of anomalously low velocities of up to  $-2\%$  for  $P$  and  $-4\%$  for  $S$  waves extends from the shallow upper mantle to 400 km depth beneath

central Iceland. *Allen et al.* [2002a] used a combined data set of fundamental mode surface wave phase velocity and long-period body wave delays from both ICEMELT and HOTSPOT stations. They found a laterally widespread low-velocity anomaly in the upper 200 km mantle, which lies above a cylindrical low-velocity zone extending down to 400 km depth. The  $P$  velocity model constructed from high-frequency body wave delays yields a more N–S elongated anomaly [*Allen et al.*, 2002a]. Constrained by a similar travel time data set, the alternative model proposed by *Foulger et al.* [2000] also showed that the low-velocity anomalies in the depth range of 250–400 km are elongated in the north-south direction. This elongation is interpreted as an indication of the upper mantle origin of the upwelling beneath the Iceland hot spot [*Foulger et al.*, 2000, 2001].

[4] Despite the extensive tomographic imaging of the mantle structure beneath Iceland, several aspects of the mantle dynamics beneath the hot spot remain unresolved or poorly resolved owing to a fundamental deficiency in the previous models, which treated seismic body waves as infinitely thin rays. As a result, the predicted arrival time is sensitive only to velocity variations along its geometrical ray path. This assumption is valid strictly only for infinite frequency waves. The observed seismic waves, however, have finite frequency bandwidths and are sensitive to three-dimensional (3-D) structure off the geometric ray path [*Dahlen et al.*, 2000; *Hung et al.*, 2000; *Zhao et al.*, 2000]. A more rigorous theory for finite frequency travel time shifts, hereinafter referred to as the Born-Fréchet kernel theory or banana-doughnut theory, yields a hollow banana-shaped sensitivity, which accounts for gradual diminution of accrued travel time anomalies along their propagation history by virtue of diffractive wave front healing. Given a velocity heterogeneity, ray approximation often overestimates travel time shifts because it does not account for wave front healing and other finite frequency effects of wave diffraction [e.g., *Wielandt*, 1987; *Hung et al.*, 2001; *Baig et al.*, 2003]. Consequently, tomographic inversions based on ray theoretical travel times likely underestimate the magnitude of velocity anomalies. Furthermore, because of the differences in the frequency and wave speed between  $P$  and  $S$  waves and thus differences in wave front healing, their travel times are sensitive to different 3-D volumes even for pairs of  $P$  and  $S$  delay times sharing the same geometrical ray path. Without robust estimates of  $P$  and  $S$  velocity heterogeneity strengths, other parameters and their geodynamic implications (such as  $\delta \ln V_S / \delta \ln V_P$ , an indicator of the presence of melt and compositional heterogeneity) cannot be further explored with confidence. It is thus important to reassess  $P$  and  $S$  velocity perturbations using a more realistic representation of wave propagation.

[5] In this paper we compare the models of  $P$  and  $S$  wave speed perturbations based on the conventional ray theory and the Born-Fréchet kernel theory. Taking advantage of the fact that the 3-D sensitivity kernel of a broadband seismic signal changes with frequency, we measure relative  $P$  and  $S$  travel time residuals across seismic stations in Iceland by cross correlation of the waveforms of  $P$  and  $S$  arrivals in three different frequency bandwidths. The banana-doughnut theory is then employed to construct the 3-D Fréchet kernels of individual arrivals for the inversion of compressional and shear velocity heterogeneity beneath Iceland. The finite

frequency approach yields a columnar low-velocity zone for both  $P$  and  $S$  waves that extends from shallow mantle to the base of the upper mantle. In the best resolved depth range (150–400 km) the amplitudes of the velocity perturbations are nearly twice those in the ray theoretical models.

## 2. Study Region and Data

[6] We use  $P$  and  $S$  arrivals from two seismic arrays in the ICEMELT and HOTSPOT experiments [*Bjarnason et al.*, 1996; *Wolfe et al.*, 1997; *Foulger et al.*, 2000, 2001; *Allen et al.*, 2002a, 2000b] and one permanent GSN station, BORG (Figure 1a). The two experiments deployed broadband seismometer instruments throughout Iceland and were operated during the years of 1995–1996 and 1996–1998, respectively. The data set comprises direct  $P$  and  $S$ , diffracted  $P_{\text{diff}}$  and  $S_{\text{diff}}$ , and  $PKIKP$  and  $SKS$  phases from teleseismic earthquakes that spanned a wide range of azimuths and epicentral distances ( $31^\circ$ – $176^\circ$ ) (Figure 1b).

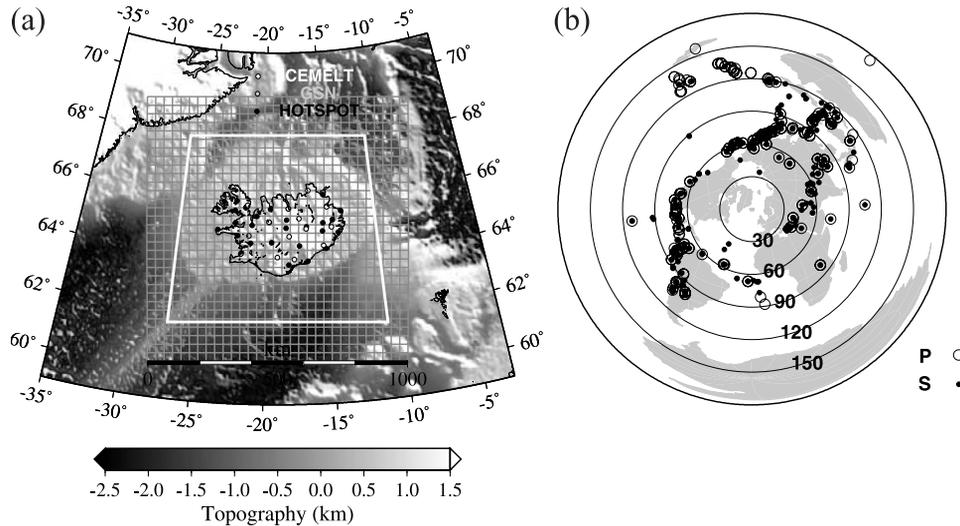
[7] Travel time delays are frequency-dependent as illustrated in Figure 2, which shows clear  $P$  arrivals at short and long periods from an event at an epicentral distance of  $\sim 71^\circ$ . The relative delays are significantly smaller in the long-period  $P$  waves, suggesting that wave diffraction takes place and that low-frequency travel time shifts are healed eminently because of their broad sensitivity. To account for frequency-dependent delays, we measure  $P$  and  $S$  travel times by multichannel cross correlation [*VanDecar and Crosson*, 1990] of band-pass-filtered waveforms in short, intermediate, and long periods (0.5–2.0, 0.1–0.5, and 0.03–0.1 Hz for  $P$  waves and 0.1–0.5, 0.05–0.1, and 0.02–0.05 Hz for  $S$  waves). The frequency bandwidths are chosen to isolate the strong microseismic noise peaked at  $\sim 6$  s in Iceland and to yield comparable wavelengths of the  $P$  and  $S$  arrivals for each corresponding frequency range. The  $S$  wave travel time is measured only in two higher-frequency bands in the HOTSPOT data set because of a limited low-frequency response of the HOTSPOT stations. The measurements of delays of the same phase arrival at different frequencies are justifiable in the banana-doughnut theory because the same phase arrival but at different frequencies no longer shares an identical sensitivity restricted on the ray path. Rather they “feel” the heterogeneity within each individual unique volume surrounding the ray path. A total of 5492 measured  $P$  delay times and 4428  $S$  delays are utilized to invert for spatial variations in  $P$  and  $S$  wave speed (or slowness) perturbations according to ray or 3-D finite frequency kernel formulations.

## 3. Theory of Travel Time Tomography

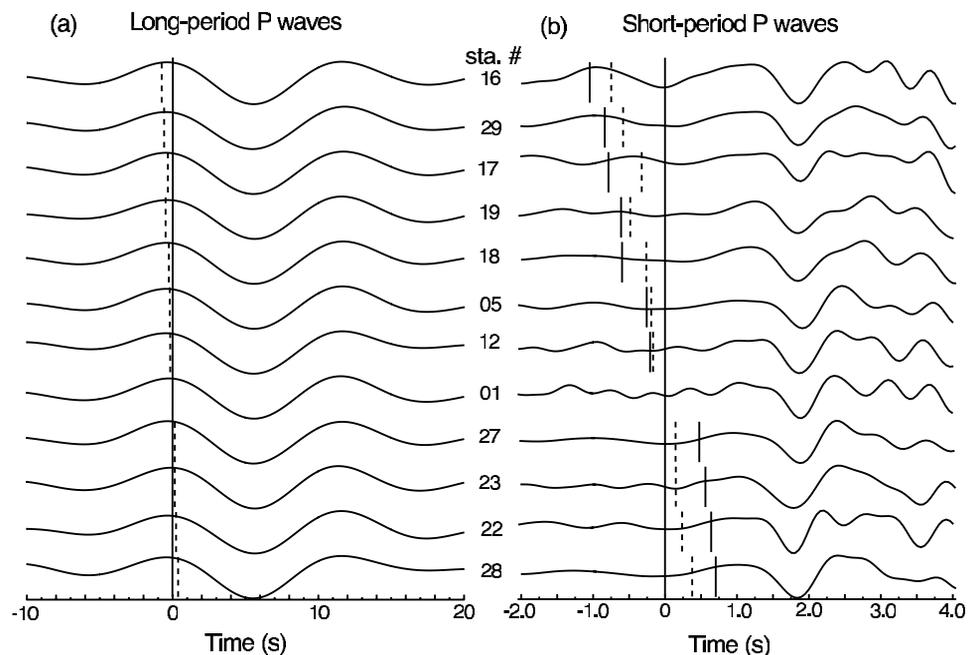
### 3.1. Linearized Ray Theory

[8] In conventional travel time tomography the linearized ray theory is commonly utilized to interpret a travel time shift relative to a value predicted for a radially symmetric Earth model (e.g., the IASP91 or PREM model). Propagation of a seismic wave within the heterogeneous Earth is regarded as a ray traveling along its infinitely thin geometrical path. A travel time shift  $\delta t$  that relates to infinitesimal wave speed perturbations encountered by a phase arrival is expressed as

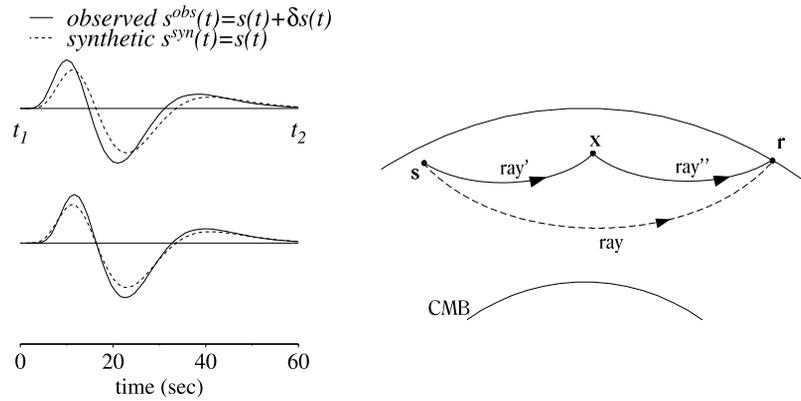
$$\delta t = - \int_{\text{ray}} c(\mathbf{x})^{-2} \delta c(\mathbf{x}) dl, \quad (1)$$



**Figure 1.** (a) Location of study area and seismic stations in Iceland. The data in this study are from two temporary broadband seismic arrays of the ICEMELT and HOTSPOT experiments and one GSN station, BORG. The outer box spans  $\sim 1000$  km in both latitude and longitude directions and is divided into equally spaced grids for model parameterization in travel time tomography. The inner box delineates the region within which seismic velocity perturbations are best resolved and discussed in the paper. (b) Azimuthal coverage of the earthquakes with useful body wave arrivals. Open circles represent the events having measured  $P$  travel times, and solid circles are those providing  $S$  data.



**Figure 2.** Example of frequency-dependent  $P$  arrivals observed at the HOTSPOT stations from an earthquake at an epicentral distance of  $\sim 71^\circ$ . Broadband pulses have been band-pass-filtered (a) at 0.03–0.1 Hz to measure low-frequency travel time shifts and (b) at 0.5–2 Hz for high-frequency measurement. The waveforms are aligned according to the time picks of the multichannel cross correlation (MCCC) method [VanDecar and Crosson, 1990]. The predicted arriving times from the 1-D IASP91 model have been subtracted. Note the relative time shifts of low-frequency  $P$  arrivals (dashed lines) are about 65% of those of high-frequency  $P$  arrivals (solid lines).



**Figure 3.** (left) Illustration of the finite frequency travel time shift of a body wave phase  $\delta t$  measured by cross correlation of the observed waveform  $s(t) + \delta s(t)$  (solid lines) with the corresponding synthetic seismogram in a 1-D radially symmetric Earth model (dashed lines). (right) Cartoon cross section of a spherically symmetric Earth showing the geometrical ray from a source  $s$  to a receiver  $r$ . The Born single-scattering theory along with the paraxial approximation accounts for the scattered waves of the same type as the unperturbed central ray which propagates along a composite path  $\text{ray}'$  and  $\text{ray}''$  from the source to a point heterogeneity  $x$  and then to the receiver. If the scattered waves arrive closely and interfere with the unperturbed wave, they contribute to the sensitivity kernel of a cross-correlation travel time.

where  $dl$  is the incremental arc length along the unperturbed ray path,  $c(\mathbf{x})$  is the unperturbed wave speed at a point  $\mathbf{x}$  on the ray from a 1-D reference Earth model, and  $\delta c/c$  is the fractional wave speed perturbation to be imaged.

### 3.2. Banana-Doughnut Theory

[9] In reality, wave front healing, scattering, and other diffraction phenomenon around velocity heterogeneities render the travel time of an actual finite frequency wave sensitive to 3-D wave speed perturbation off the ray path. The travel time shift measured by broadband waveforms can differ significantly from the prediction of ray theory when the cross-path dimension of the 3-D velocity variation is comparable to the characteristic width of the Fresnel zone of the wave [e.g., *Wielandt, 1987; Hung et al., 2001; Baig et al., 2003*]. Recent investigation into the theoretical underpinnings of finite frequency travel times has put forth an efficient method to correct the deficiencies of ray approximation [e.g., *Marquering et al., 1999; Dahlen et al., 2000; Zhao et al., 2000*]. On the basis of Born single-scattering approximation in conjunction with body wave propagation, a new formulation of the Born-Fréchet kernels for seismic travel time expresses explicitly the influence of velocity heterogeneity off the ray path upon a finite frequency travel time shift; it replaces the 1-D line integral in equation (1) with a 3-D volume integration over the entire Earth where wave speed perturbations are not zero [*Dahlen et al., 2000*],

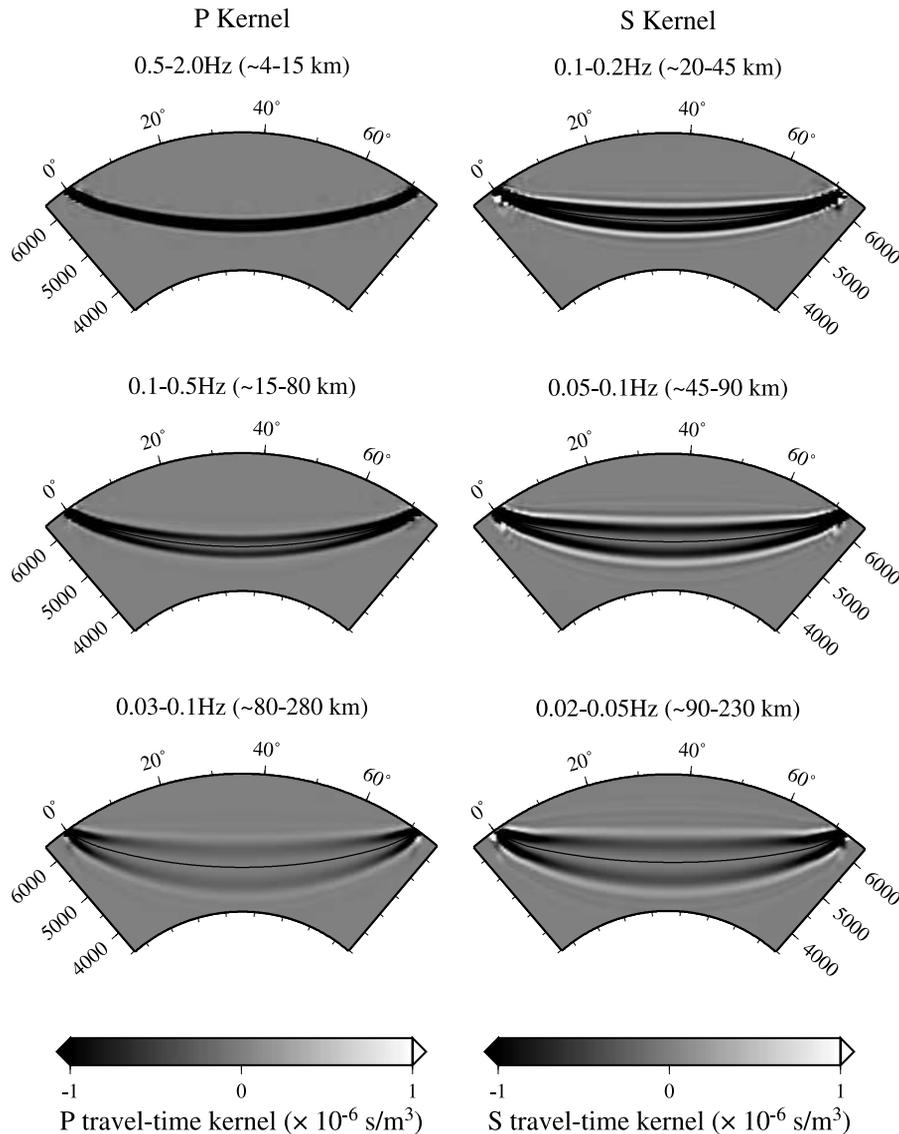
$$\delta t = \iiint_{\oplus} K(\mathbf{x}) \delta c(\mathbf{x}) / c(\mathbf{x}) d^3 \mathbf{x}, \quad (2)$$

where the quantity  $K(\mathbf{x})$  is the 3-D Fréchet sensitivity kernel for a shift  $\delta t$  that is measured by cross correlation of an observed pulse with its spherical Earth synthetics. Taking advantage of the fact that the primary sensitivity of a travel time shift is confined within a slender region surrounding the ray, the paraxial kernel which only incorporates the

contribution from forward scattering paths of the same type as the unperturbed ray yields a computationally efficient formulation [*Dahlen et al., 2000*]:

$$K = -\frac{1}{2\pi c} \left( \frac{\mathcal{R}}{c_r \mathcal{R}' \mathcal{R}''} \right) \frac{\int_0^\infty \omega^3 |s_{\text{syn}}(\omega)|^2 \sin(\omega \Delta T) d\omega}{\int_0^\infty \omega^2 |s_{\text{syn}}(\omega)|^2 d\omega}, \quad (3)$$

where quantity  $\Delta T$  is the additional travel time required to take the detour path through the single point scatterer, the integration point  $\mathbf{x}$  in equation (2) within a radially symmetric Earth (Figure 3). The scaling constants  $c$  and  $c_r$  are the wave speeds of the 1-D background model at the scatterer  $\mathbf{x}$  and receiver  $\mathbf{r}$ ; the quantities of  $\mathcal{R}$ ,  $\mathcal{R}'$ , and  $\mathcal{R}''$  are geometrical spreading factors for the unperturbed ray, the forward source-to-scatterer ray, and the backward receiver-to-scatterer ray, respectively. Frequency dependence of cross-correlated travel times is manifested by the presence of the power spectrum of the spherical Earth pulse  $|s_{\text{syn}}|^2$  in equation (3). Figure 4 shows the ray plane cross-sectional views of 3-D sensitivity kernels of finite frequency travel times for a teleseismic wave at an epicentral distance of  $\sim 70^\circ$ . Each particular kernel is constructed for a  $P$  or  $S$  arrival in the three designated frequency bands relative to the predicted time in the IASP91 model. The difference between ray and kernel interpretations is particularly evident for the low- and intermediate-frequency arrivals, as illustrated by the broad region of off-path sensitivity enclosing a visible minimum sensitivity hole. The phase arrival of a typical teleseismic long-period wave of 20 s places very little constraint on the velocity structure close to the ray path; rather, it is influenced mostly by the heterogeneity off the ray by a few hundred kilometers in the upper mantle and up to 1000 km at the turning point (Figure 4). The region spanned by the two side lobes of the strongest sensitivity is usually referred as to the Fresnel zone, whose width varies with the propagation distance and



**Figure 4.** Born-Fréchet kernels for finite frequency travel times of teleseismic (left)  $P$  and (right)  $S$  phases at an epicentral distance of  $\sim 70^\circ$ , measured by cross correlation of the observed waveforms with the corresponding spherical Earth synthetics in three designated frequency ranges, 0.5–2.0, 0.1–0.5, and 0.03–0.1 Hz for  $P$  waves and 0.1–0.2, 0.05–0.1, and 0.02–0.05 Hz for  $S$  waves. In the ray plane cross sections, all the kernels exhibit a bent elliptical shape with zero sensitivity right on the geometrical ray paths, shown by solid lines. The maximum sensitivity is in the two side-lobed, dark gray-to-black regions, which can be regarded as the first Fresnel zones, and their widespread cross-path scales are approximately proportional to  $\sim \sqrt{\lambda L}$ , where  $\lambda$  is the characteristic wavelength of each individual  $P$  or  $S$  pulse;  $L$  is the propagation distance to the receiver. The kernels of high-frequency  $P$  and  $S$  arrivals thus have a slender hollow banana-shaped geometry, while low-frequency travel times are essentially influenced by wave speed heterogeneity farther off the ray and differ substantially from ray theoretical predictions.

the dominant wavelength of an observed pulse used for waveform cross-correlation measurements. For the ICEMELT and HOTSPOT data, the  $P$  wave observed in the low-frequency range of 0.03–0.1 Hz has a dominant period of  $\sim 12$  s and a Fresnel zone width of  $\sim 140$  and  $\sim 270$  km at the distances of 200 and 600 km, respectively, from the station. The width of the 1-s  $P$  delay kernel is slimmed down to  $\sim 40$ –75 km at the same distances. Likewise, the sensitivity of a  $S$  arrival measured in the

shortest detectable period of  $\sim 8$  s occupies a volume with a dimension ranging from  $\sim 85$  km at a distance of 200 km from the station to  $\sim 160$  km at 600 km. For a 20-s period  $S$  wave traveling the same distances its Fresnel zone widths are 130 and 260 km, respectively.

[10] In regional travel time tomography the relative arrival times of a phase at a number of stations in a seismic array are measured to constrain spatial variations of underlying mantle velocity perturbations. The finite volume

sensitivity of a relative delay between two nearby stations 1 and 2,  $\delta t_1 - \delta t_2$ , is simply the difference of the individual Fréchet kernels for the travel time shifts,  $\delta t_1$  and  $\delta t_2$ , i.e.,

$$K_{\delta t_1 - \delta t_2} = K_{\delta t_1} - K_{\delta t_2}. \quad (4)$$

### 3.3. Model Parameterization and Regularization

[11] There are only a finite amount of observed travel time data in tomographic inversions, though the pursued unknown variables, i.e., the fractional wave speed perturbations, are continuous. The general data rule for the inverse problem can be characterized by [e.g., *Chiao and Kuo*, 2001; *Chiao and Liang*, 2003]

$$d_i = \int_D g_i(\mathbf{x})m(\mathbf{x})d^3\mathbf{x} + e_i, \quad (5)$$

where  $d_i$  and  $e_i$ ,  $i = 1 \dots N$ , represent the  $i$ th travel time data and its error, respectively;  $\mathbf{x}$  is the position vector in 3-D vector model space,  $D$ ; and the data kernel,  $g_i(\mathbf{x})$ , is the Fréchet derivative of a functional that maps the continuous model function  $m(\mathbf{x})$  into the discrete data  $d_i$ . In linearized ray theory described in equation (1) the model function to be solved for is often assigned to be the slowness perturbation,  $\delta(1/c)$ , so that the data kernel,  $g_i$ , simply corresponds to the differential path length,  $dl$ . When the doctrine of travel time inversion is based on the banana-doughnut theory of equation (2), the Fréchet derivative is related to the 3-D integral of the finite frequency kernel within an infinitesimal 3-D volume, i.e.,  $Kd^3\mathbf{x}$ .

[12] In regional tomography the model function of fractional wave speed or slowness perturbations is commonly parameterized on a spatial grid of nodes. Making  $m_l$  the model parameter at the  $l$ th grid node leads to a concise matrix form for the discrete inverse problem:

$$d_i = A_{il}m_l. \quad (6)$$

For the models derived from ray theory,  $A_{il}$  is the difference of the total path lengths throughout a specific volume or voxel that contributes to the  $l$ th node between the two arrivals in the  $i$ th paired relative travel time measurement. When the data rule of the inversion calls for the 3-D finite frequency kernels,  $A_{il}$  is the differential value of the integrated volumetric kernels,  $Kd^3\mathbf{x}$ , contributing to the  $l$ th node. Numerical integration of the kernel value within a single voxel is achieved by the weighted sum of 27 sampling points according to the Gaussian quadrature formula [e.g., *Zienkiewicz and Taylor*, 1989]. The data vector,  $\mathbf{d}$ , is composed of  $N$  measured relative travel time shifts,  $[\delta t_1, \delta t_2, \dots, \delta t_N]^T$  and stipulated as the result of the discrete inner product of the matrix  $\mathbf{A}$  and the model vector of dimension  $L$ ,  $\mathbf{m} = [m_1, m_2, \dots, m_L]^T$ .

[13] Our tomographic model is a  $9^\circ \times 9^\circ$  cap centered at a pole located at ( $65^\circ\text{N}$ ,  $19^\circ\text{W}$ ), which spans  $\sim 1000$  km in both latitude and longitude directions and extends in depth to 1000 km from the surface. The model space is parameterized in regular grids with a dimension  $L$  of  $32 \times 32 \times 32$ , which results in an approximately equally spaced discretization with  $\sim 32$  km in all directions (Figure 1a).

[14] The standard damped least squares (DLS) solution for the model obtained with the minimum norm constraint is given by [e.g., *Menke*, 1984]

$$\hat{\mathbf{m}} = (\mathbf{A}^T\mathbf{A} + \theta^2\mathbf{I})^{-1}\mathbf{A}^T\mathbf{d}, \quad (7)$$

where  $\mathbf{I}$  is an  $L \times L$  identity matrix;  $\theta^2$  is a nonnegative damping factor specified prior to the inversion to provide the intended harshness of the minimum norm criterion. The inverse of the massive matrix is approximated by the iterative solution of the LSQR algorithm [*Paige and Saunders*, 1982].

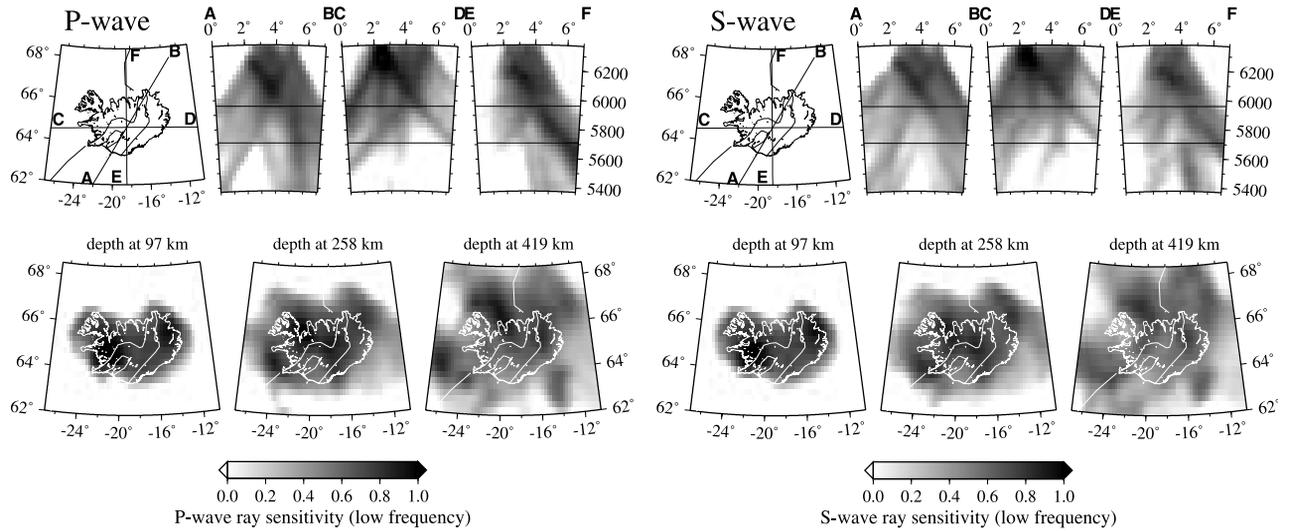
[15] Over parameterization is usually recommended to avoid the aliasing or spectral leakage [e.g., *Trampert and Snieder*, 1996; *Chiao and Kuo*, 2001]. Because of the infinitely thin width across a ray a large portion of the model parameters is sparsely sampled as a result of non-uniform coverage of ray geometry. The unresolvable model components require an adequate regularization scheme to capture the well-resolved features while suppressing the poorly constrained small-scale variations. One way to impose a smoothness regularization scheme other than simple norm damping involves a convolutional quelling acting on the matrix  $\mathbf{A}$ , which enforces an a priori length scale of the model correlation to annihilate the roughness of the solution. The DLS solution in equation (7) can be, in general, substituted by

$$\hat{\mathbf{m}} = \mathbf{W}(\mathbf{W}^T\mathbf{A}^T\mathbf{A}\mathbf{W} + \theta^2\mathbf{I})^{-1}\mathbf{W}^T\mathbf{A}^T\mathbf{d}, \quad (8)$$

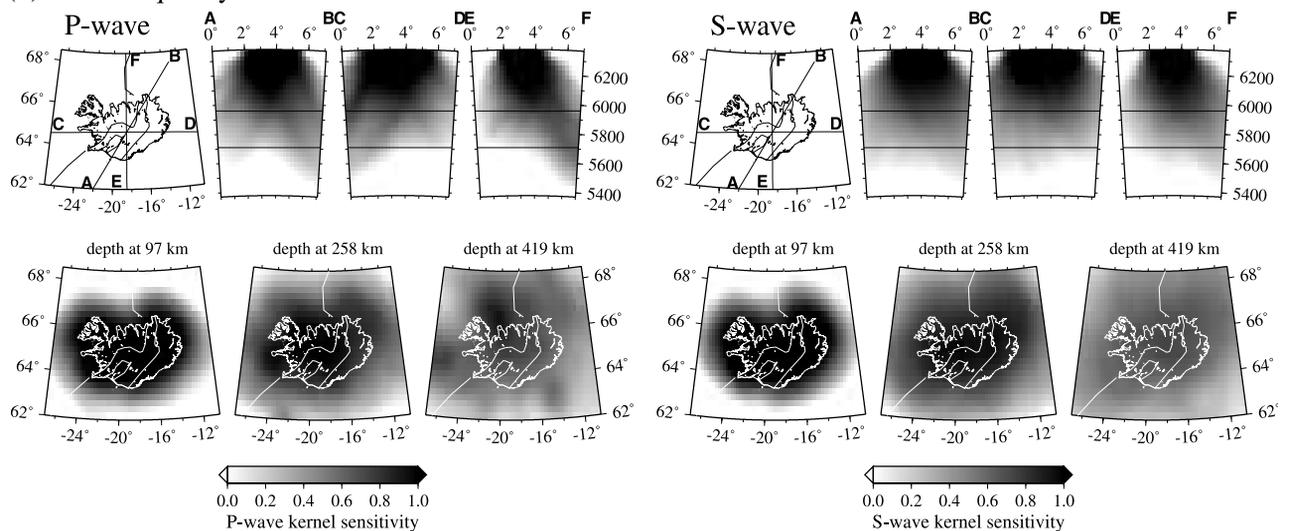
where  $\mathbf{W}$  is the convolution operator that prescribes a preferential correlation length on the resolved model [*Meyerholtz et al.*, 1989]. If  $\mathbf{W}$  is an identity matrix, then no a priori correlation length is imposed, and the model solution becomes identical to the simple DLS estimate. For ray-based inversions we employ a Gaussian correlation function with a prescribed number of grid intervals (1 grid) as the standard deviation of the Gaussian function. This approach is similar to the ad hoc ‘‘fat’’ ray approximation with no hole inside, which only accounts for the widespread cross-path sensitivity but not wave front healing of a finite frequency travel time. Because the broad sensitivity founded on the Born-Fréchet kernel theory provides a natural smoothness regularization, only the simple norm damping is allowed in the solutions obtained from finite frequency kernel theory to avoid ad hoc smoothness enforcement.

[16] Under the ray assumption the measured travel time anomaly is invariably attributed to the velocity heterogeneity on the ray path no matter what frequencies are utilized to determine the arriving time of an observed broadband pulse. Thus the relative delays obtained in different frequency ranges must be solved separately under the ray approximation for wave speed perturbations. Because the majority of useful teleseismic events have unambiguous  $P$  and  $S$  arrivals in the designated low-frequency range, among all the ray-based models those constrained by the longest-period data set are best resolved and are compared with the kernel-derived models. In contrast, the kernel-based tomography combines the measured travel time anomalies at all available frequencies, fully utilizing the broadband nature of the recorded arrivals.

## (a) ray-theoretical models



## (b) finite-frequency kernel models

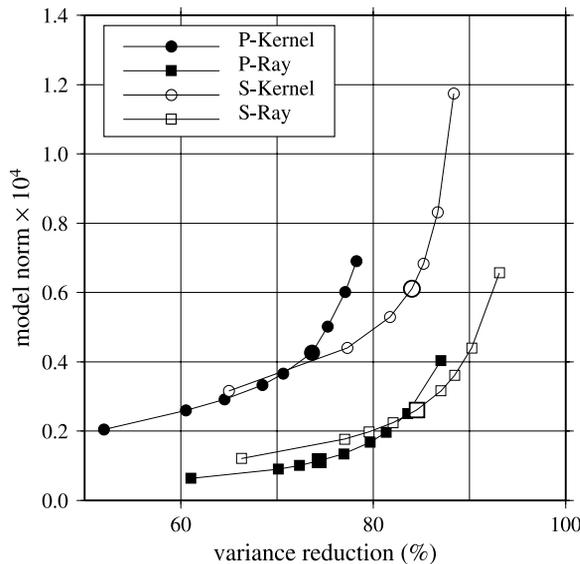


**Figure 5.** Diagonal elements of the matrix  $\mathbf{A}\mathbf{A}^T$ , where  $\mathbf{A}$  is the Gram matrix for the inversion, represent the overall sensitivity of  $P$  or  $S$  data contributing to each grid node for the (a) ray theoretical and (b) finite frequency kernel models. The values are plotted in a logarithm scale on three vertical AB, CD, and EF slices through central Iceland as denoted on the insert Iceland map and three horizontal slices at 97, 258, and 419 km depth. The solid lines on the Iceland map delineate the neovolcanic rift zones and the Mid-Atlantic Ridge. Sampling in the ray theory directly reflects the geographical distribution of source-receiver paths, while sampling by 3-D finite volume kernels is more evenly distributed because of the broad cross-path sensitivity.

[17] In Figure 5 we present the spatial distribution of the diagonal values of the product of the Gram matrix  $\mathbf{A}$  and its transpose  $\mathbf{A}^T$ ,  $\text{diag}(\mathbf{A}\mathbf{A}^T)$ , a measure defined here to compare the overall sensitivity inherent to the ray and kernel tomography provided by the observed  $P$  and  $S$  travel time data. Each element of  $\text{diag}(\mathbf{A}\mathbf{A}^T)$  is equivalent to the total sum of the squares of the path lengths of 1-D rays or of the squared values of the volumetric kernels contributing to each node. The variation of these values reflects the robustness of the constraint on the velocity perturbation at each node. The plots are presented in a logarithm scale on six 2-D cross sections where the resolved velocity anomaly is shown subsequently. As expected, sampling in the ray

theory reflects directly the path geometry. Restricted by the aperture of the land-based seismic array, no crossing rays exist below 400 km beneath Iceland to allocate the velocity heterogeneity in the deeper mantle [e.g. *Aki et al.*, 1977; *Ritsema and Allen*, 2003]. In contrast, the uppermost 400 km beneath central Iceland are sampled relatively evenly by the kernels. Though the sampling density is reduced in the upper mantle transition zone, the space with no passing rays can actually be “sensed” by broad off-path sensitivity of long-period waves passing nearby.

[18] There is no absolutely objective and unique criterion to compare the models based on different data rules. Here we choose the models with equally good fit or the same



**Figure 6.** Trade-off between model norm and variance reduction for the  $P$  velocity (solid symbols) and  $S$  velocity (open symbols) models based on ray theoretical (squares) and finite frequency kernel (circles) tomography. Larger symbols represent the models in Figure 7, which have variance reductions of  $\sim 75\%$  for  $P$  wave and  $\sim 84\%$  for  $S$  wave.

variance reduction to compare and discuss the main differences between the resolved velocity structures. First, we vary the norm damping factors and perform the trade-off analysis of model norm versus variance reduction. Figure 6 shows the trade-off relations of the model  $L_2$  norm versus variance reduction for both  $P$  and  $S$  models obtained by ray and finite frequency kernel theory. We decided to compare the models within the pivot regime on the trade-off curves, which provide coherent images without significant contamination from observational errors and give approximately equal variance reductions of 75% for  $P$  and 84% for  $S$  data.

#### 4. $P$ and $S$ Velocity Models

[19] Because the ray paths of teleseismic body wave arrivals are almost vertical and rarely cross each other in shallow depths beneath the stations, relative travel time data are incapable of imaging shallow velocity structures. The seismic signature accrued along the path in the crustal portion hence needs to be cautiously corrected. There are two common approaches to carry out this correction. The first is to incorporate an additional free term at each station that stands for the time shift of an arrival associated with crustal velocity anomaly and lateral variation in station elevation and Moho topography [e.g., Wolfe *et al.*, 1997; Foulger *et al.*, 2000]. The station terms are solved simultaneously in the inversion to absorb as much the delay signature from shallow heterogeneity as possible. The second is to subtract the crustal travel time anomaly predicted by a velocity model directly from the observed travel time. The corrected delay times are then employed to solve for underlying mantle structures [e.g., Allen *et al.*, 2002a]. We have exercised both tactics in our tomographic

inversion and have found that the difference in resolved mantle velocity anomalies is trivial in the target region (below 150 km, and see discussions below).

#### 4.1. Strengths of Velocity Anomalies

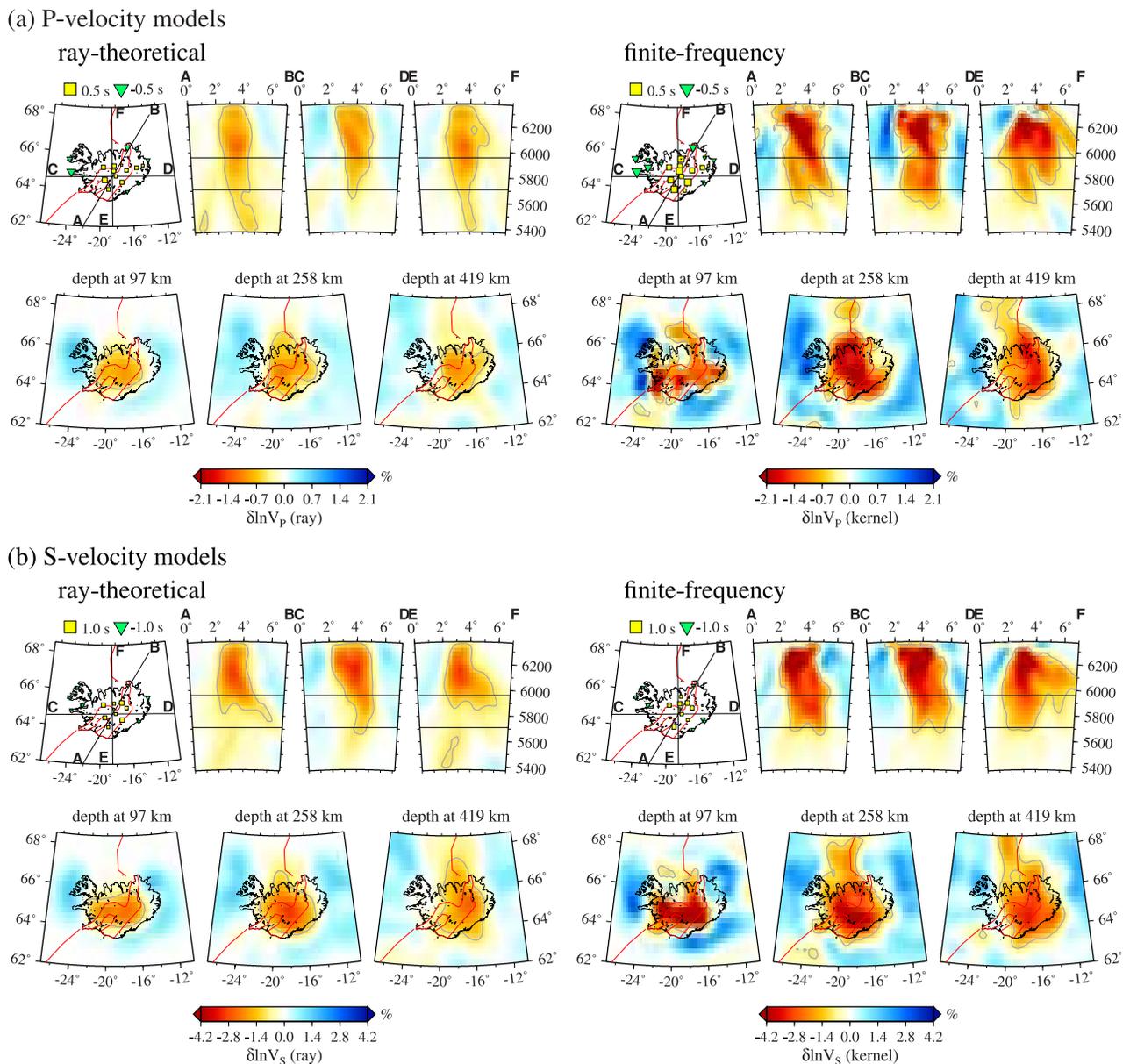
[20] Figure 7 compares the resolved mantle  $P$  and  $S$  velocity anomaly between the ray and kernel tomography using the minimum norm regularization chosen in Figure 6. The diagram of each model depicts the images of fractional wave speed perturbations on the same horizontal and vertical slices as in Figure 5. The simultaneously inverted station terms are shown by the squares and triangles on the insert Iceland map. Both models yield a nearly vertical columnar zone of low  $P$  and  $S$  velocity anomaly centered in Iceland as outlined by the contours of  $-0.5\%$   $\delta \ln V_P$  and  $-1.0\%$   $\delta \ln V_S$ , respectively. The anomaly defined by the iso- $\delta \ln V$  contours extends continuously from the depth of 150 km to at least 400 km. It has an east-west dimension of  $\sim 250$ – $300$  km in the kernel-based models and is slightly narrower in the ray-based ones. The low-velocity region appears to be elongated northward with a reduced magnitude in the depth range of 250–400 km. This is most obvious in the vertical profile EF and horizontal slices at 258 and 419 km depth in the kernel-derived  $P$  and  $S$  models.

[21] Given approximately equal fit to the observed data, the magnitudes of both  $P$  and  $S$  velocity anomalies derived from finite frequency travel time kernels are significantly greater than those from ray approximation and in previous studies. The root-mean-square (RMS) of fractional  $P$  and  $S$  wave speed perturbations for each layer and the ratios of the kernel-derived RMS magnitude to the ray-derived one are plotted as a function of depth in Figure 8. The magnitude of the kernel-derived  $P$  velocity anomaly is on average 2–3 times as large as that of the ray-derived anomaly within the depths of 100–400 km. The magnitude of the  $S$  wave speed perturbations increases approximately by a factor 1.6–2.3 in the finite frequency model relative to the corresponding ray-based model.

#### 4.2. Upper Mantle Transition Zone Structure

[22] In the ray-based models the seismic structure tends to be smeared in the direction of incoming rays. This is evident in a comparison of the vertical slices of the ray-based models in Figure 7 and those in Figure 5.

[23] Vertical smearing is reduced in kernel-based tomography, as demonstrated in resolution tests using the exact path coverage provided by the observed delay times. In Figure 9 we compare the lateral and vertical resolutions in the ray and kernel tomography by recovering a hypothetical low-velocity model with the maximum  $-3\%$   $P$  and  $-6\%$   $S$  wave speed perturbations. The contours outline the shape of the input velocity anomaly, which characterizes two main features revealed in the tomographic images: a  $\sim 300$ -km wide,  $\sim 800$ -km-long low-velocity channel roughly parallel to the spreading center in the shallow mantle above 150 km (more discussion in section 4.3) and an underlying low-velocity conduit having a diameter of 200 km. To illustrate vertical resolution in the deep mantle, we limit the height of the cylindrical low-velocity conduit to 200 km and place another low-velocity conduit of the same dimensions in the depth from 500 to 700 km beneath central Iceland. Gaussian random errors with the standard deviations of 0.05 s for  $P$

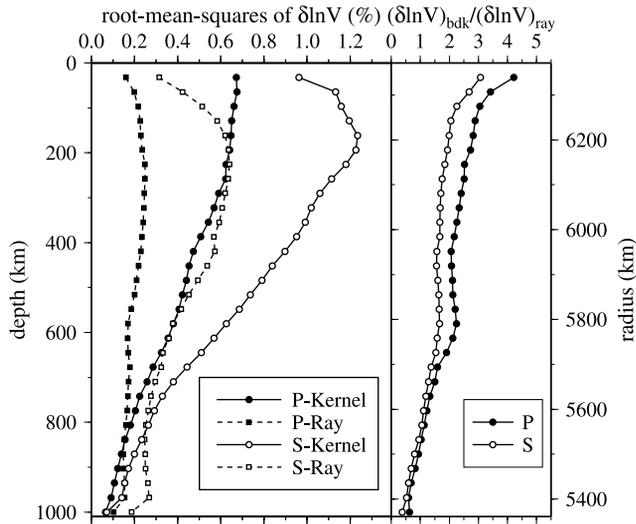


**Figure 7.** Comparison of the tomographic images of (a)  $P$  and (b)  $S$  wave speed perturbations in the (left) ray theoretical and (right) finite frequency kernel models. Velocity variation in each model is depicted on the 2-D slices shown in Figure 5. The contours shape the region in which the  $P$  and  $S$  velocities are at least 0.5% and 1.0% lower than those in the surrounding mantle. All the models reveal a columnar zone of low-velocity anomaly of  $\sim 250$ – $300$  km in diameter beneath central Iceland. In the finite frequency models this low-velocity feature extends to the base of the upper mantle, while the maximum resolvable depth for the ray-based models is  $\sim 400$  km. The maximum amplitudes of both  $P$  and  $S$  velocity perturbations in the kernel-based models are  $\sim 1.8$ – $2.3$  times those in the ray theoretical models. The simultaneously solved station terms, which account for the travel time shift associated with local shallow heterogeneity, are indicated by squares for the positive or relatively low-velocity anomaly and triangles for the negative or relatively high-velocity anomaly. The size of the symbols is proportional to the magnitude of the station corrections.

data and 0.1 s for  $S$  data are added to synthetic travel time delays generated from ray theoretical and finite frequency kernel formulations. The LSQR inversion procedures with the same minimum norm damping criteria and extra convolutional-quelling smoothing for ray tomography are then employed to recover the input velocity anomaly. In addition,

we also simultaneously solve the station correction terms in the same fashion as for the observed data to check if they impair the recovery of the input structure, particularly in the uppermost mantle.

[24] The kernel-based tests show a better recovery of both the amplitude and geometry of the input  $P$  and  $S$



**Figure 8.** (left) Root-mean-square (RMS) of fractional wave speed perturbations for  $P$  (solid symbols) and  $S$  (open symbols) models as a function of depth or radius. (right) Ratios of the RMS of the kernel-obtained velocity perturbations to those of the ray-based perturbations as a function of depth. Only the results within the best resolved region outlined by white lines in Figure 1a are included in the RMS calculation.

velocity structure than the ray-based ones. While the amplitude of the recovered structure in the transition zone is significantly reduced compared to the input, the geometry of the anomaly is well determined in the kernel models with little vertical smearing. Vertical smearing in the transition zone is clearly problematic in the ray inversions. Because the synthetic time delays obtained from finite frequency kernels are healed and in general smaller than those unhealed times overpredicted by infinitely frequency rays, the contrast in the recovered heterogeneity strength between the kernel and ray tomography is not as pronounced as that found in the inversion of observed data.

[25] To assess horizontal smearing, which may broaden the lateral extent of the low-velocity region and lead to an erroneous estimate of the dimension of the anomalous feature, we duplicated the resolution test in Figure 9 but with half the width of the ridge-parallel structure and half the diameter of the low-velocity cylinders (Figure 10). The resolution for the narrower velocity anomaly deteriorates substantially. Nonetheless, the geometry of the input anomaly in the transition zone is reasonably well recovered in the kernel-based models, and there is no significant broadening due to horizontal smearing.

[26] We conclude that the anomalously low  $P$  and  $S$  velocity structures in the upper mantle transition zone imaged by Born-Fréchet kernel tomography are required by the finite frequency travel times. The low velocities in the transition zone are in general agreement with the inference of an anomalously hot and thin transition zone beneath Iceland from observations of  $P$ -to- $S$  conversions at the 410- and 660-km discontinuities [Shen *et al.*, 1998, 2002]. We note that the limited aperture of the arrays renders the finite frequency kernels of the Iceland data sets

insensitive to the velocity structure in the lower mantle (Figure 5).

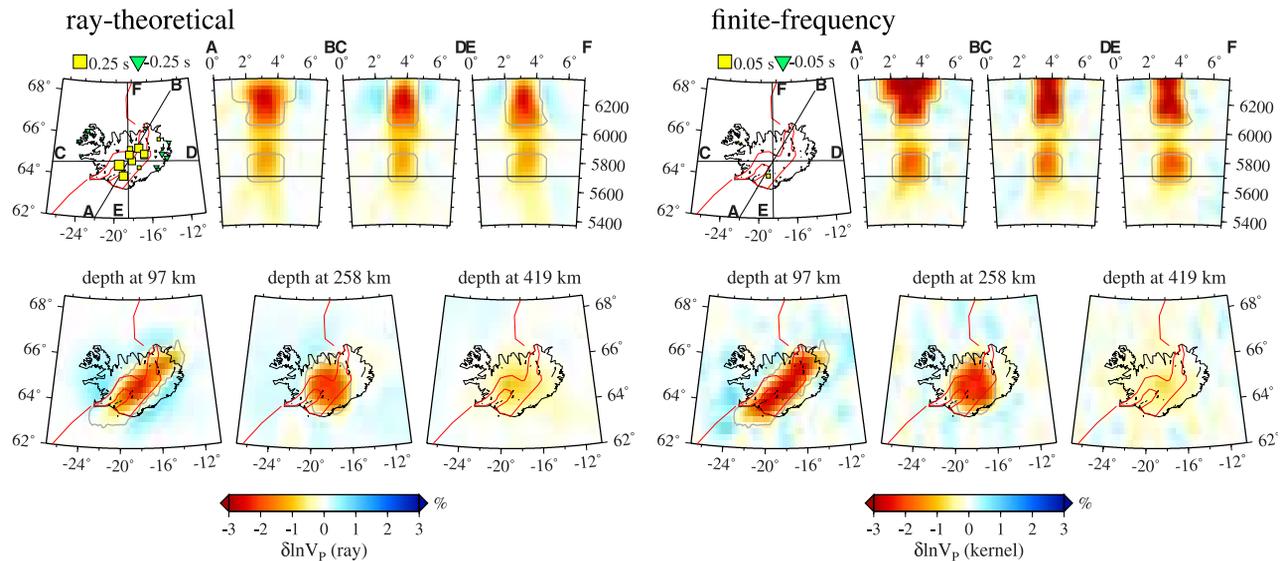
### 4.3. Shallow Upper Mantle Velocity Structure

[27] Though all the models indicate that the compressional and shear wave speeds in the uppermost mantle are relatively slow beneath Iceland as illustrated by the images at 97 km depth (Figure 7), the lateral variations of velocity perturbations are quite different between the ray- and kernel-based solutions. The smooth, circular-shaped low-velocity feature beneath central Iceland in the ray-based models does not vary much between the two horizontal slices at depths of 97 and 258 km except that the diameter of the velocity anomaly is slightly smaller in the shallower depth and becomes more east-west elongated for the  $S$  wave model. In contrast, the deep, columnar low-velocity zones in the finite frequency  $P$  and  $S$  models are transformed to shallow features that are more coincident with the locations of the volcanic zones in Iceland. The ray-based  $P$  and  $S$  body wave models from Wolfe *et al.* [1997] and surface wave phase velocity models from Li and Detrick [2003] resolved similar anomalously slow structures in the shallow upper mantle.

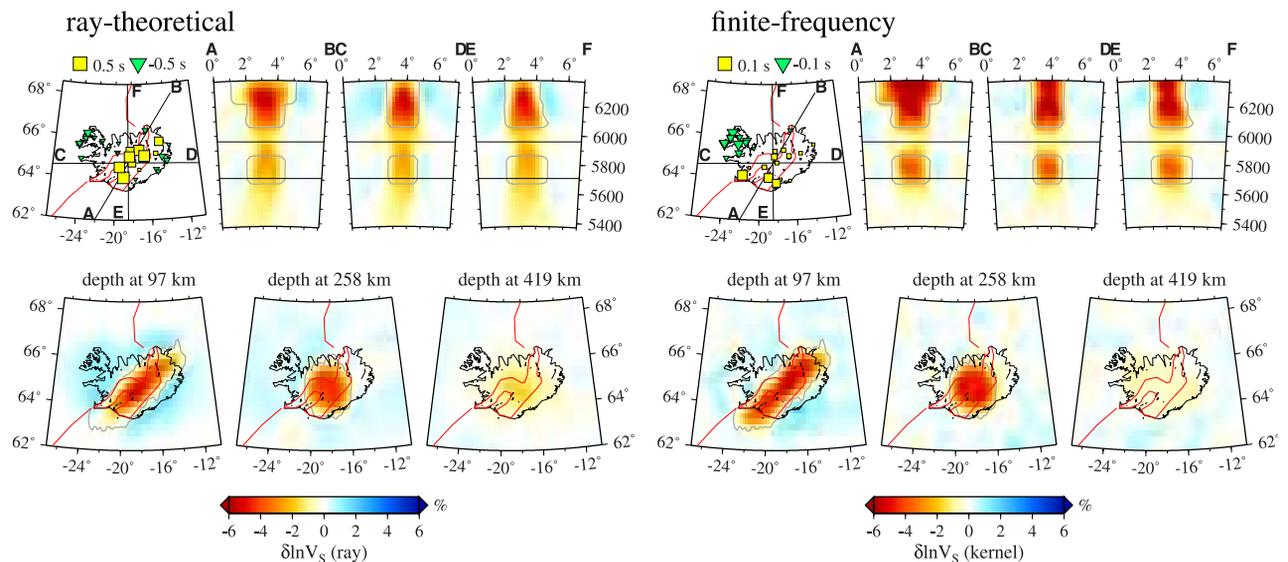
[28] The station terms solved to absorb the travel time fluctuation due to shallow structures yield a consistent pattern among all the models, a relative delay at the stations in central Iceland associated with a thicker crust and/or a lower velocity in the crust and uppermost mantle. If the crustal effects are not completely extracted from the observed relative delays by station terms, the residual will be mapped into mantle velocity structures [Allen *et al.*, 2002a]. To assess the crustal effects on the mantle velocity structure, we then make an a priori crustal corrections on the observed  $P$  and  $S$  arriving times, to the first-order approximation, by calculating the travel times taken by a vertical ray passing through the crustal layer beneath the stations. We adopt a 3-D crustal  $S$  wave model by Allen *et al.* [2002b] and a linear scaling relation of  $V_P/V_S = 1.78 + 0.004d$ , where  $d$  is the depth in km to obtain the crustal  $P$  velocity. We use the corrected data set to perform the inversion employed with identical minimum norm damping and smooth quelling procedures. Figure 11 presents the resulting  $P$  and  $S$  velocity perturbations on the same slices shown in Figure 7. Since the travel time anomaly due to the crustal velocity variation is assumed to be removed to the first-order prior to the inversion, the top 30 km which represents the crust is not included in the model parameterization.

[29] Figure 12 shows the correlation coefficients as a function of depth for the models with two different crustal corrections. The structure is overall well correlated below 150 km depth with the correlation coefficient over 90%. The correlation at the uppermost three layers above 100 km depth deteriorates. The low degree of similarity between the  $P$  models at shallow depths is partly due to errors in crustal correction applied prior to the inversion since we assume that the  $P$  crustal velocity is completely correlated with the  $S$  velocity by a linear scaling relation. In the finite frequency models the drop in correlation above 100 km results partly from ray theoretical crustal corrections applied identically to finite frequency travel time data measured in different frequency ranges. The other source of errors may arise from the deficiency of far-field Fréchet kernels at shallow

## (a) synthetic P models



## (b) synthetic S models



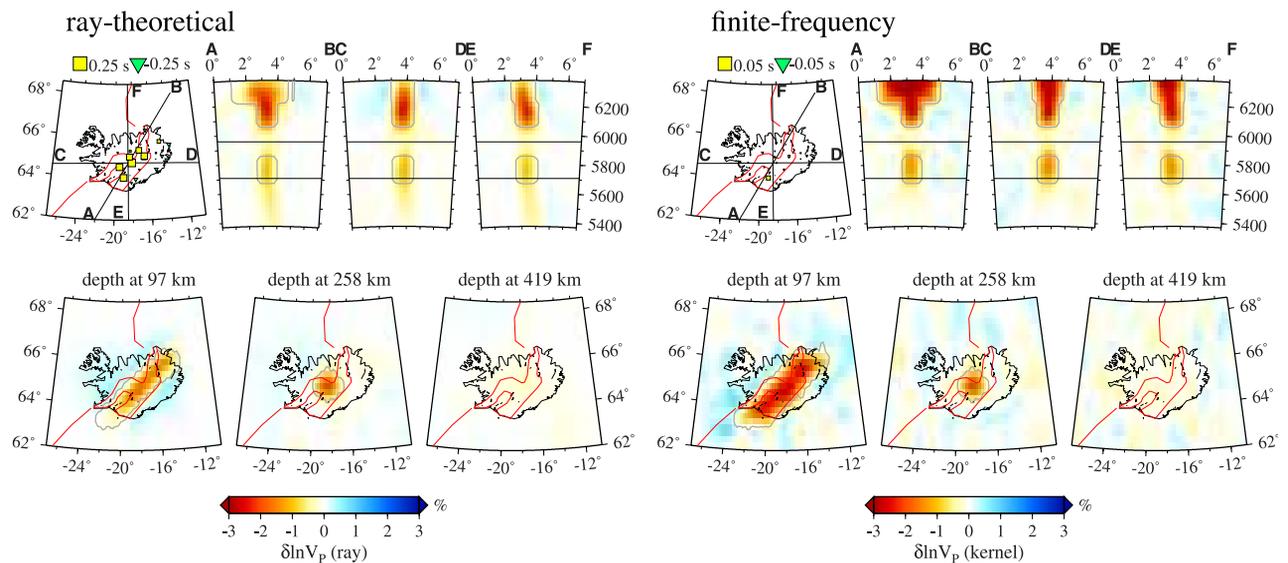
**Figure 9.** Resolution tests using the same path coverage of (a)  $P$  and (b)  $S$  data for the ray and kernel inversion. The input model outlined by the contoured lines has a 300-km-wide, low-velocity channel parallel to the neovolcanic zone in the shallow mantle above 150 km. The two underlying cylindrical anomalies ( $-3\%$  for  $P$  and  $-6\%$  for  $S$ ) beneath central Iceland have a height of 200 km and are placed 350 km apart down to the depth of 700 km. The magnitudes and shapes of the input low-velocity anomalies are better recovered from near surface to the base of the upper mantle in the finite frequency models. The models based on the ray theoretical travel times show obvious vertical smearing of the input transition zone structure.

depths where local field terms are predominant. On the basis of the complete Green's tensor or displacement response from an impulsive force, *Favier et al.* [2004] showed that  $P$  and  $S$  travel time kernels are no longer zero on the ray paths near the source and receiver and that the contribution of middle field terms can extend down to a depth of one wavelength  $\lambda$  for  $P$  wave and  $2\lambda$  for  $S$  wave. Nevertheless, the zone significantly influenced by near and middle field terms is within a depth of  $0.5\lambda$  from the receiver, where the neglect of local field response can introduce a 10% error in  $P$  and  $S$  delays in the near field [*Favier et al.*, 2004]. For the

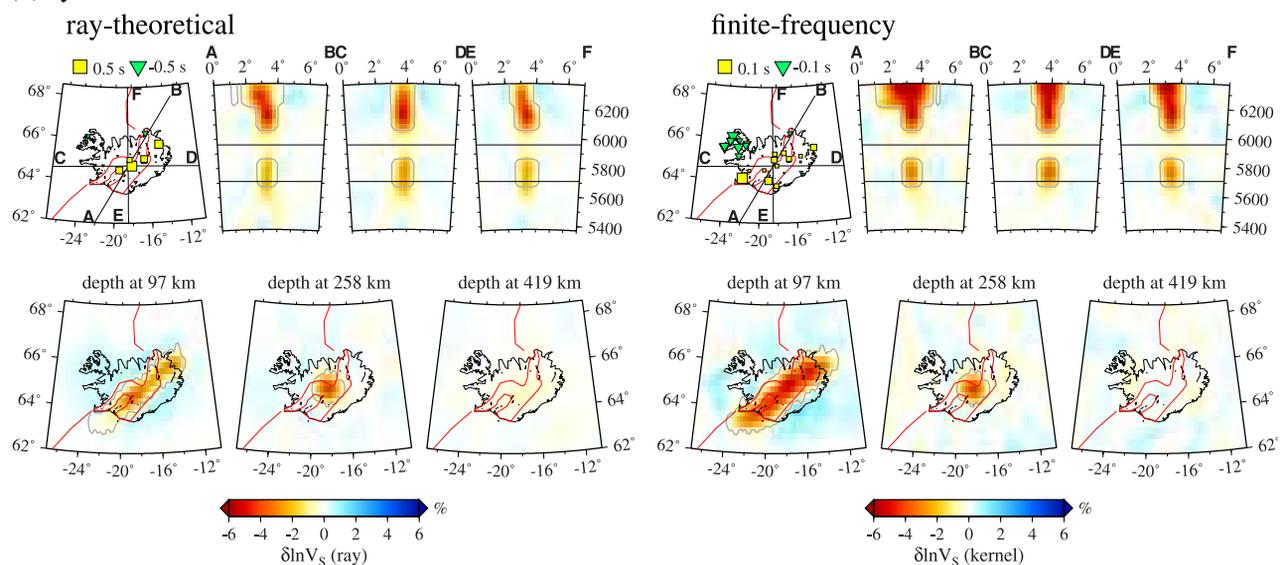
$P$  and  $S$  waves used in this study the zone of significant local field net influence is restricted at depth less than  $\sim 50$  km.

[30] We note that the kernel-derived station terms in the resolution tests (Figure 9) are at least 5 times smaller than those in the ray-based tests and in the inversions of real data (Figure 7), suggesting that the station terms in the kernel models are more decoupled from the shallow mantle structure than those in ray inversions. Consequently, the banana-doughnut theory recovers shallow velocity heterogeneity in the uppermost mantle much better than ray theory (Figures 9

## (a) synthetic P models



## (b) synthetic S models



**Figure 10.** Resolution tests using the same path coverage of (a)  $P$  and (b)  $S$  data for the ray and kernel inversions. The input model shares the same configuration but only half the lateral dimension of the low-velocity zone in Figure 7.

and 10). Thus the roughly spreading-center-parallel low-velocity feature at  $\sim 100$  km depth is probably not an artifact resulting from far-field approximation.

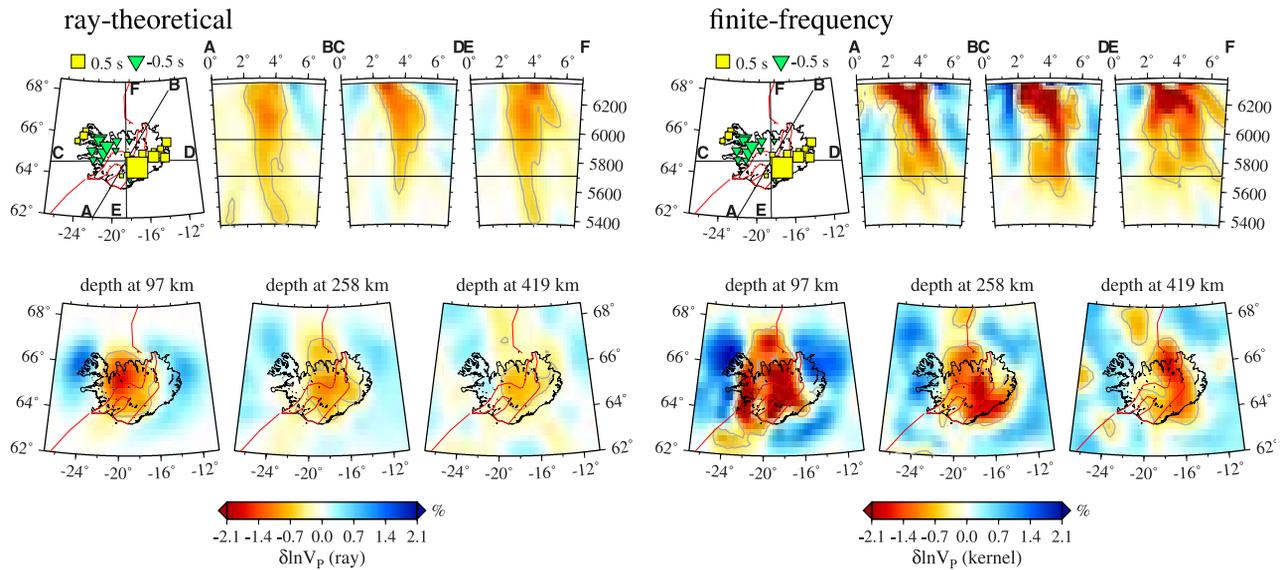
## 5. Discussion

[31] The differences between the velocity models obtained by the ray theoretical and finite frequency kernel inversions are evident and significant. The finite frequency models yield considerably greater magnitudes of seismic wave speed perturbations and better resolutions of the geometry of velocity anomalies, especially in deeper mantle (Figures 7, 9, and 10). The improvement can be attributed to the proper representation of the ubiquitous phenomenon of wave front healing of seismic waves by the 3-D Fréchet sensitivity kernel. In the ray theoretical scenario a single observed travel

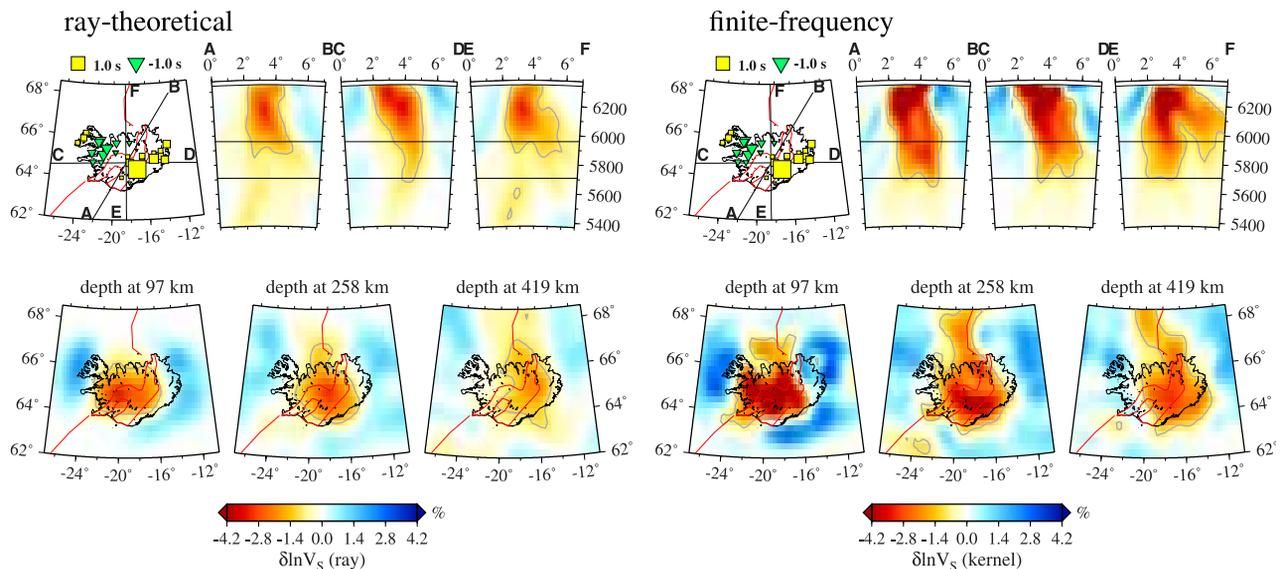
time advance or delay cannot pinpoint the location of a fast or slow anomaly along the source-to-receiver path. Consequently, velocity models lack resolution in places with no crossing rays to locate the velocity anomaly. In banana-doughnut theory, on the other hand, relative travel time delays between two nearby stations in an array from the same earthquake could “feel” the same velocity structure by their broadening sensitivity in deep mantle in places where their infinitely thin rays do not necessarily cross. Moreover, the extent of wave front healing imprinted on the observed travel time shift varies with the propagation distance from velocity anomaly to observation point as a function of the wavelength and therefore provides an intrinsic signature to “target” the heterogeneity (Figure 13).

[32] *Hung et al.* [2001] and *Baig et al.* [2003] conducted a comprehensive analysis of the effects of wave front

## (a) P-velocity models



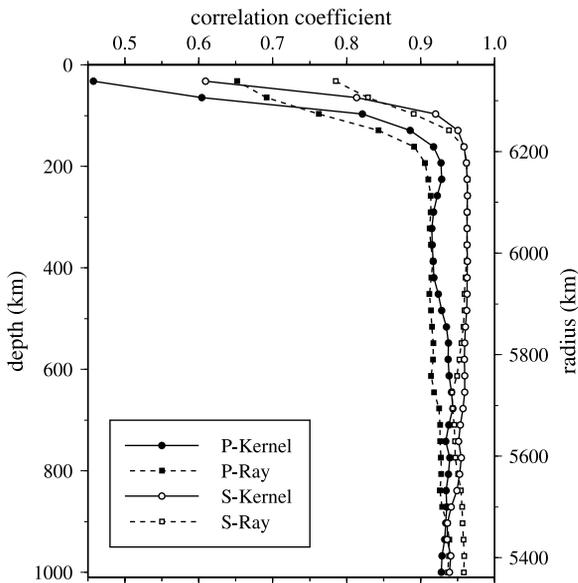
## (b) S-velocity models



**Figure 11.** Comparison of the tomographic images of fractional (a)  $P$  and (b)  $S$  wave speed perturbations between the (right) ray theoretical and (left) finite frequency kernel models. Instead of using the station terms to absorb travel time shifts associated with crustal and shallow mantle structure, we use the  $S$  wave crustal model of *Allen et al.* [2002b] and a linear scaling relation for  $P$  and  $S$  wave velocities to obtain 1-D average velocity structures beneath each station. The delays of nearly vertical teleseismic rays accrued within the local crust beneath the stations are then removed from observed  $P$  and  $S$  travel time residuals prior to the inversion. The contours delineate the geometry of  $-0.5\%$   $P$  and  $-1.0\%$   $S$  velocity perturbations. Except for the image on the constant depth slice at 97 km, the resolved velocity variations are quite similar to those shown in Figure 7. Note that the top white area on the vertical slices represent the 30-km-thick crust and is excluded in tomography.

healing on cross-correlation measured travel times in the media of the sphere-shaped or randomly distributed velocity perturbations characterized by a variety of scale lengths. The travel time shifts measured directly by synthetic pulses from “ground truth” numerical modeling of finite frequency wave propagation [*Hung and Forsyth, 1998*] are compared with those predicted from linearized ray theory and Born-Fréchet kernel theory. Ray theoretical travel

times, in general, overestimate the cross-correlation travel times whenever the heterogeneity scale length is less than the width of the first Fresnel zone,  $\sim\sqrt{\lambda L}$ . The finite frequency effect in a more Earth-like stochastic model is less evident than in that with a single spherical anomaly [*Baig et al., 2003*]. Though the extent of wave diffraction from the cylindrical velocity structure beneath Iceland probably lies between them, the following examples for a



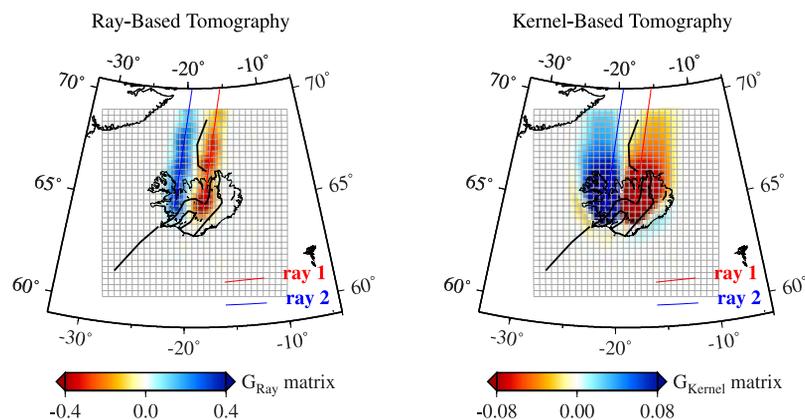
**Figure 12.** Correlation functions on each constant depth nodal grid between the models using station terms and those having a priori crustal corrections. The  $P$  velocity models from the ray and kernel tomography are denoted by solid circles and squares, respectively, while the  $S$  models are shown by open symbols. Though shallow crustal heterogeneity is treated very differently in the models, the resulting velocity structures are well correlated except for those at the four grid nodes above 100 km. The poor correlation at shallow depths particularly for the kernel-based  $P$  model can be attributed to the improper crustal correlation and the inaccuracy of the kernels in the near-field range.

spherical anomaly provide a measure of the healing effect. The ground truth delay of a 20-s  $S$  wave (wavelength  $\lambda \sim 90$  km) obtained at a distance of 400 km following passage through a low-velocity sphere of 250 km in

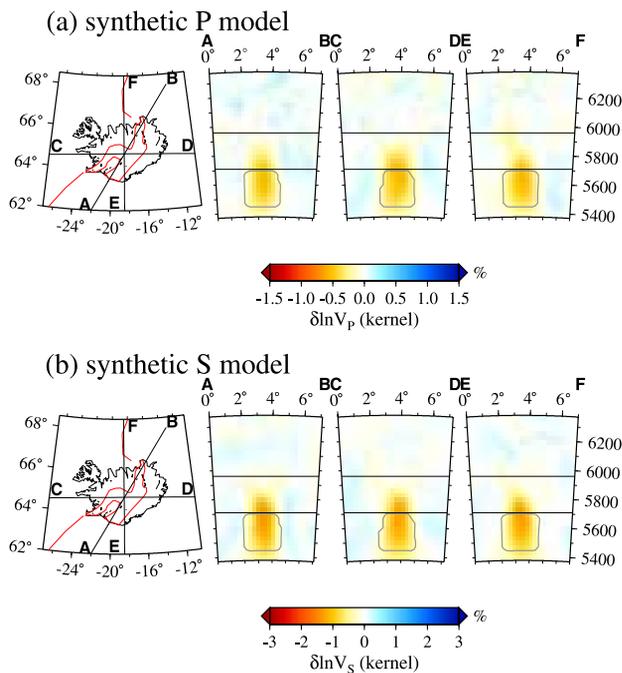
diameter is  $\sim 60\%$  less than the ray prediction or the travel time shift observed right behind the velocity anomaly. Optimistically, a ray-based tomography can then recover only 40% of the magnitude of this velocity anomaly. The travel time of a  $P$  wave at the same period is reduced even more severely because of its higher speed and longer wavelength. The healing of a 10-s  $S$  wave is substantially reduced; only less than 10% of the delay is diminished during the propagation distance of 400 km.

[33] The banana-doughnut travel times, which account for the loss by finite frequency wave diffraction, yield excellent agreements with the ground truth travel times over a wide range of propagation distances and anomaly dimensions and magnitudes. As a result, seismic tomography based on 3-D Fréchet kernels holds the promise of a fuller recovery of velocity heterogeneity.

[34] The maximum amplitudes of anomalously low-velocity structures resolved from our ray inversion of long-period travel time data are  $-1.5\%$  for  $P$  velocity and  $-3.5\%$  for  $S$  velocity in the best resolved depth range of 200–400 km. The results are similar to those found in the previous regional models which all implicitly utilize the infinite frequency assumption while interpreting finite frequency travel time data measured by waveform cross correlation. Though the damping and regularization schemes may be imposed differently in the various inversions by different investigators, the strengths of resolved low-velocity anomalies from ray tomography are not higher than  $-2\%$  and  $-4\%$  for  $P$  and  $S$  waves, respectively. If seismic wave speed perturbation is caused entirely by thermal effect, the  $-2\%$  and  $-4\%$   $P$  and  $S$  velocity anomalies from ray-based tomographic images can be accounted for by an excess temperature of  $\sim 150$ – $200^\circ\text{C}$  [Wolfe et al., 1997; Allen et al., 2002a], in general agreement with those inferred from basalt chemistry [Schilling, 1991; White et al., 1995; Shen and Forsyth, 1995; Saunders et al., 1997; Presnall et al., 2002] and transition zone thickness beneath Iceland [Shen et al., 1998]. With the same velocity-temperature relation the



**Figure 13.** Sensitivity of a long-period  $P$  relative delay between two nearby stations recording the same event from the north. (left) Plot of conventional 1-D ray tomography and (right) plot of 3-D kernel inversion. The images portray the variations of the maximum values of 1-D “fat” ray or 3-D banana-doughnut-shaped sensitivity among all the depths projected onto the parameterized grids on the surface of the study area. Though the incoming teleseismic rays are roughly parallel, the overlapped cross-path sensitivity of a single, finite frequency relative delay provides a natural signature to target underlying velocity structure.



**Figure 14.** Resolution tests that assess the resolution in the lower mantle using the finite frequency data kernels. The input model is a single, 200-km-wide cylindrical conduit placed between the depths of 700 and 900 km as outlined by the contours. It has the magnitude of  $-1.5\%$  and  $-3\%$  for  $P$  and  $S$  velocity perturbation, respectively. The station terms are not incorporated in the resolution test because of the absence of shallow velocity heterogeneity. Compared to that in the uppermost mantle and the transition zone, the resolution in the lower mantle is significantly degraded in both the shape and magnitude of the recovered velocity anomaly.

maximum amplitudes of the low-velocity anomaly in the kernel-based model ( $-3.5\%$  for  $P$  and  $-6.2\%$  for  $S$  wave speed within the depths of 200–400 km), however, would require an excess temperature of  $\sim 250$ – $350^\circ\text{C}$ . Alternatively, the enhanced strength of  $P$  and  $S$  wave speed perturbations is indicative of higher attenuation or the presence of partial melt in the deep ( $>150$  km) mantle beneath the Iceland hot spot. A detailed discussion of the geodynamic implication for the thermal state, the presence of partial melt, and the chemical composition projected from the imaged seismic velocity perturbations will be deferred to a later study integrating results from various seismic investigations [e.g., Shen *et al.*, 2002; Allen and Tromp, 2002] including a joint kernel-based inversion of  $P$  and  $S$  data for the ratio of  $\delta\ln V_S/\delta\ln V_P$ .

[35] Last, whether the columnar low-velocity zone beneath Iceland is tilted and extends continuously to the lower mantle remains uncertain from the tomographic point of view. Because of the limited aperture of the seismic network, even the finite frequency travel time data provides poor sampling of the lower mantle beneath Iceland (see Figure 5). In Figure 14 we evaluate the recoverability of a blob of velocity heterogeneity in the uppermost lower mantle. The input structure is a single

low-velocity cylinder of 200 km in diameter placed between 700 and 900 km beneath central Iceland. Considering the decrease of  $\delta\ln V/\delta T$  in the lower mantle [Karato, 1993], we reduce the magnitudes of the input  $P$  and  $S$  velocity perturbations to  $-1.5\%$  and  $-3.0\%$ , half those in the previous tests for the resolution in the upper mantle. The recovered anomaly is severely smeared vertically, confirming the loss of resolution in the lower mantle. Two other factors may further render it difficult to detect the possible presence of a low-velocity anomaly in the uppermost mantle beneath Iceland with the existing data. If the low-velocity zone is tilted in such a way that the lower mantle velocity heterogeneity associated with the surface hot spot is not directly below central Iceland, the recovery would be worse. An uplifted 660-km discontinuity associated with an excess temperature in the transition zone [Shen *et al.*, 1998, 2002] could cause a localized high-velocity zone due to the phase change, which partially balances the thermal effects. We conclude that if a plume-like, low-velocity structure is present in the uppermost lower mantle beneath Iceland, it can be revealed unequivocally only by a network of ocean bottom seismic stations off Iceland.

## 6. Conclusions

[36] Several conclusions can be drawn from this comparative study between the linear ray and finite frequency kernel tomography. When the banana-doughnut-shaped sensitivity accounting for wave front healing and other finite frequency diffractive effects is utilized to interpret seismic travel time variations, the magnitudes of tomographically constrained velocity perturbations on average increase by a factor of 2–3 times for  $P$  wave and of 1.6–2.3 times for  $S$  wave in the best resolved depths of 100–400 km. The maximum amplitudes of low  $P$  and  $S$  wave speed perturbations below 200 km are  $-3.5\%$  and  $-6.2\%$ , respectively; much greater than  $-2\%$  and  $-4\%$  found in the ray theoretical models. The enhanced velocity heterogeneity strength suggests that the mantle is more attenuating than previously assumed or that partial melt is present to nearly 400 km, much deeper than previously suggested [e.g., White *et al.*, 1995].

[37] The kernel-derived tomographic images are best characterized by a columnar low-velocity zone for both  $P$  and  $S$  waves beneath central Iceland, similar to those revealed by the ray-based models. The lateral dimension of such anomalous structure is slightly wider in the kernel-derived models,  $\sim 250$ – $300$  km across Iceland. A much lower-magnitude, northward elongated feature is discernible in the depth range of 250–400 km. As a result of the widespread off-path sensitivity for cross-correlation travel times, the finite frequency tomography is able to image the structure below 400 km where the rays essentially do not cross to locate the velocity anomaly. Both resulting  $P$  and  $S$  models indicate that this anomalous region extends continuously from the shallow mantle to the base of the upper mantle. Limited by the array aperture of Iceland stations, the velocity structure in the lower mantle cannot be well resolved from the current travel time data set. In the shallow mantle above 120 km the kernel-based models reveal a low-velocity region beneath the neovolcanic zones.

[38] **Acknowledgments.** We thank G. R. Foulger, an anonymous referee, and Associate Editor E. Garnero for their thoughtful comments, which greatly improved the manuscript. Financial support for this work was provided by National Science Council of Taiwan, under grants NSC-902611-M-002-020 and NSC-902611-M-002-025 (S.H.H.) and U.S. National Science Foundation (Y.S. and S.H.H.).

## References

- Aki, K., A. Christofferson, and E. S. Husebye (1977), Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.*, **82**, 277–296.
- Allen, R. M., and J. Tromp (2002), Constraining the Iceland low-velocity anomaly to test causal hypotheses, *Eos Trans. AGU*, **83**(47), Fall Meet. Suppl., Abstract S71D-12.
- Allen, R. M., et al. (2002a), Imaging the mantle beneath Iceland using integrated seismological techniques, *J. Geophys. Res.*, **107**(B12), 2325, doi:10.1029/2001JB000595.
- Allen, R. M., et al. (2002b), Plume-driven plumbing and crustal formation in Iceland, *J. Geophys. Res.*, **107**(B8), 2163, doi:10.1029/2001JB000584.
- Baig, A. M., F. A. Dahlen, and S.-H. Hung (2003), Traveltimes of waves in three-dimensional random media, *Geophys. J. Int.*, **153**, 467–482.
- Bijwaard, H., and W. Spakman (1999), Tomographic evidence for a narrow whole mantle plume below Iceland, *Earth Planet. Sci. Lett.*, **166**, 121–126.
- Bjarnason, I. T., C. J. Wolfe, S. C. Solomon, and G. Gudmundson (1996), Initial results from the ICEMELT experiment: Body-wave delay times and shear-wave splitting across Iceland, *Geophys. Res. Lett.*, **23**, 459–462.
- Chiao, L.-Y., and B.-Y. Kuo (2001), Multiscale seismic tomography, *Geophys. J. Int.*, **145**, 517–527.
- Chiao, L.-Y., and W.-C. Liang (2003), Multiresolution parameterization for geophysical inverse problems, *Geophysics*, **145**, 517–527.
- Dahlen, F. A., S.-H. Hung, and G. Nolet (2000), Fréchet kernels for finite-frequency traveltimes—Theory, I, *Geophys. J. Int.*, **141**, 157–174.
- Darbyshire, F. A., R. S. White, and K. F. Priestley (2000), Structure of the crust and uppermost mantle of Iceland from a combined seismic and gravity study, *Earth Planet. Sci. Lett.*, **181**, 409–428.
- Favier, N., S. Chevrot, and D. Komatitsch (2004), Near-field influence on shear wave splitting and travel time sensitivity kernels, *Geophys. J. Int.*, **156**, 467–482.
- Foulger, G. F., et al. (2000), The seismic anomaly beneath Iceland extends down to the mantle transition zone and no deeper, *Geophys. J. Int.*, **142**, F1–F5.
- Foulger, G. R., et al. (2001), Seismic tomography shows that upwelling beneath Iceland is confined to the upper mantle, *Geophys. J. Int.*, **146**, 504–530.
- Foulger, G. R., Z. Du, and B. R. Julian (2003), Icelandic-type crust, *Geophys. J. Int.*, **155**, 567–590.
- Hung, S.-H., and D. W. Forsyth (1998), Modeling anisotropic wave propagation in oceanic inhomogeneous structures using the parallel multi-domain pseudospectral method, *Geophys. J. Int.*, **133**, 726–740.
- Hung, S.-H., F. A. Dahlen, and G. Nolet (2000), Fréchet kernels for finite-frequency traveltimes—II. Examples, *Geophys. J. Int.*, **141**, 175–203.
- Hung, S.-H., F. A. Dahlen, and G. Nolet (2001), Wavefront healing: A banana-doughnut perspective, *Geophys. J. Int.*, **146**, 289–312.
- Ito, G., J. Lin, and C. W. Gable (1996), Dynamics of mantle flow and melting at a ridge-centered hotspot: Iceland and the Mid-Atlantic Ridge, *Earth Planet. Sci. Lett.*, **144**, 53–74.
- Karato, S.-I. (1993), Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, **20**, 1623–1626.
- Korenaga, J., and P. B. Kelemen (2000), Major element heterogeneity in the mantle source of the North Atlantic igneous province, *Earth Planet. Sci. Lett.*, **184**, 251–268.
- Li, A., and R. Detrick (2003), Azimuthal anisotropy and phase velocity beneath Iceland: Implication for plume-ridge interaction, *Earth Planet. Sci. Lett.*, **214**, 153–165.
- Marquering, H., F. A. Dahlen, and G. Nolet (1999), Three dimensional sensitivity kernels for finite-frequency traveltimes: The banana-doughnut paradox, *Geophys. J. Int.*, **137**, 805–815.
- Menke, W. (1984), *Geophysical Data Analysis: Discrete Inverse Theory*, Academic, San Diego, Calif.
- Meyerholtz, K. A., G. L. Pavlis, and S. A. Szpakowski (1989), Convolutional quelling in seismic tomography, *Geophysics*, **54**, 570–580.
- Montelli, R., G. Nolet, G. Masters, F. A. Dahlen, E. R. Engdahl, and S.-H. Hung (2004), Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, **303**, 338–343.
- Morgan, W. J. (1971), Convection plumes in the lower mantle, *Nature*, **230**, 42–43.
- Paige, C. C., and M. A. Saunders (1982), LSQR: An algorithm for sparse linear-equations and sparse least-squares, *Trans. Math. Software*, **8**, 43–71.
- Presnall, D. C., G. H. Gudfinnsson, and M. J. Walter (2002), Generation of mid-ocean ridge basalts at pressures from 1 to 7 GPa, *Geochim. Cosmochim. Acta*, **66**, 2073–2090.
- Ribe, N., U. R. Christensen, and J. Theissing (1995), The dynamics of plume-ridge interaction, 1: Ridge-centered plumes, *Earth Planet. Sci. Lett.*, **134**, 155–168.
- Ritsema, J., and R. M. Allen (2003), The elusive mantle plume, *Earth Planet. Sci. Lett.*, **207**, 1–12.
- Ritsema, J., H. J. van Heijst, and J. H. Woodhouse (1999), Complex shear wave velocity structure imaged beneath Africa and Iceland, *Science*, **286**, 1925–1928.
- Saunders, A. D., J. G. Fitton, A. C. Kerr, M. J. Norry, and R. W. Kent (1997), The North Atlantic igneous province, in *Large Igneous Provinces: Continental, Oceanic, and Planetary Flood Volcanism*, *Geophys. Monogr. Ser.*, vol. 100, edited by J. J. Mahoney and M. F. Coffin, pp. 45–85, AGU, Washington, D. C.
- Schilling, J.-G. (1973), Iceland mantle plume: Geochemical study of Reykjanes Ridge, *Nature*, **242**, 565–571.
- Schilling, J.-G. (1991), Fluxes and excess temperatures of mantle plumes inferred from their interaction with migrating mid-ocean ridges, *Nature*, **352**, 397–403.
- Shen, Y., and D. W. Forsyth (1995), Geochemical constraints on initial and final depths of melting beneath mid-ocean ridges, *J. Geophys. Res.*, **100**, 2211–2237.
- Shen, Y., S. C. Solomon, I. T. Bjarnason, and C. J. Wolfe (1998), Seismic evidence for a lower mantle origin of the Iceland mantle plume, *Nature*, **395**, 62–65.
- Shen, Y., et al. (2002), Seismic evidence for a tilted mantle plume and north-south flow beneath Iceland, *Earth Planet. Sci. Lett.*, **197**, 261–272.
- Trampert, J., and R. Snieder (1996), Model estimations biased by truncated expansions: Possible artifacts in seismic tomography, *Science*, **271**, 1257–1260.
- Tryggvason, K., E. S. Husebye, and R. Stefansson (1983), Seismic image of the hypothesized Icelandic hot spot, *Tectonophysics*, **100**, 97–118.
- VanDecar, J. C., and R. S. Crosson (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seismol. Soc. Am.*, **80**, 150–169.
- White, R. S., and D. McKenzie (1989), Magmatism at rift zones - the generation of volcanic continental margins and flood basalts, *J. Geophys. Res.*, **94**, 7685–7729.
- White, R. S., J. Brown, and J. R. Smallwood (1995), The temperature of the Iceland plume and the origin of outward-propagating V-shaped ridges, *J. Geol. Soc. London*, **152**, 1039–1045.
- Wielandt, E. (1987), On the validity of the ray approximation for interpreting delay times, in *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*, edited by G. Nolet, pp. 85–98, D. Reidel, Norwell, Mass.
- Wolfe, C. J., I. T. Bjarnason, J. C. VanDecar, and S. C. Solomon (1997), Seismic structure of the Iceland mantle plume, *Nature*, **385**, 245–247.
- Zhao, D. (2001), Seismic structure and origin of hotspots and mantle plumes, *Earth Planet. Sci. Lett.*, **192**, 251–265.
- Zhao, L., T. H. Jordan, and C. H. Chapman (2000), Three-dimensional Fréchet differential kernels for seismic delay times, *Geophys. J. Int.*, **141**, 558–576.
- Zienkiewicz, O. C., and R. L. Taylor (1989), *The Finite Element Method*, vol. I, *Basic Formulation and Linear Problems*, 4th ed., 648 pp., McGraw-Hill, New York.

L.-Y. Chiao and S.-H. Hung, Department of Geosciences, National Taiwan University, No. 1, Roosevelt Road Sec. 1, Taipei, 106, Taiwan. (chiao@ntu.edu.tw; shung@ntu.edu.tw)

Y. Shen, Graduate School of Oceanography, University of Rhode Island, 29 Fish Road, Narragansett, RI 02882, USA. (yshen@gso.uri.edu)