Finite-frequency effects in global surface-wave tomography

Ying Zhou,[1,]* F. A. Dahlen,[1] Guust Nolet[1] and Gabi Laske[2]

[1]*Department of Geosciences, Princeton University, Princeton,* NJ 08544, *USA*
[2]*IGPP, Scripps Institution of Oceanography, University of California San Diego, La Jolla,* CA 92093-0225, *USA*

SUMMARY
We compare traditional ray-theoretical surface-wave tomography with finite-frequency tomography, using 3-D Born sensitivity kernels for long-period, fundamental-mode dispersion measurements. The 3-D kernels preserve sidelobes beyond the first Fresnel zone, and fully account for the directional dependence of surface-wave scattering, and the effects of time-domain tapering and seismic source radiation. Tomographic inversions of Love and Rayleigh phase-delay measurements and synthetic checkerboard tests show that (1) small-scale *S*-wave velocity anomalies are better resolved using finite-frequency sensitivity kernels, especially in the lowermost upper mantle; (2) the resulting upper-mantle heterogeneities are generally stronger in amplitude than those recovered using ray theory and (3) finite-frequency tomographic models fit long-period dispersion data better than ray-theoretical models of comparable roughness. We also examine the reliability of 2-D, phase-velocity sensitivity kernels in global surface-wave tomography, and show that phase-velocity kernels based upon a forward-scattering approximation or previously adopted geometrical simplifications do not reliably account for finite-frequency wave-propagation effects. 3-D sensitivity kernels with full consideration of directional-dependent seismic scattering are the preferred method of inverting long-period dispersion data. Finally, we derive 2-D boundary sensitivity kernels for lateral variations in crustal thickness, and show that finite-frequency crustal effects are not negligible in long-period surface-wave dispersion studies, especially for paths along continent–ocean boundaries. Unfortunately, we also show that, in global studies, linear perturbation theory is not sufficiently accurate to make reliable crustal corrections, due to the large difference in thickness between oceanic and continental crust.

Key words: Fréchet derivatives, global seismology, sensitivity, surface waves, tomography.

1 INTRODUCTION

The large-scale seismic structure of the upper mantle has been progressively mapped out by surface-wave tomography during the past three decades. Traditional surface-wave tomography is based upon JWKB ray theory, which is valid only if the lateral length scales of the heterogeneities are larger than the characteristic wavelength of the seismic waves. Therefore, only the largest-scale structures can be trusted in traditional long-period surface-wave tomography; furthermore, because structure in the lowermost upper mantle is mainly constrained by surface waves of the longest periods, the spatial resolution decreases with depth in ray-theoretical tomography. The limitations of ray theory have received growing attention in recent theoretical studies, and sensitivity kernels have been developed to account for finite-frequency effects upon seismic surface-wave propagation (e.g. Snieder & Nolet 1987; Li & Romanowicz 1995;

Friederich 1999; Spetzler *et al.* 2002; Zhou *et al.* 2004). Yomogida & Aki (1987) were the first to try to account for finite-frequency effects in surface-wave phase-velocity tomography, by introducing an *ad hoc* sensitivity within the first Fresnel zone, scaled exponentially to the phase delay between the great-circle ray and the detoured scattering path. Meier *et al.* (1997) proposed a two-step tomographic method for surface-wave tomography, in which JWKB waveform modelling is first used to obtain a preliminary 1-D *S*-wave velocity model along each source–receiver path, and the remaining residuals are then inverted based upon the Born scattering approximation. Yoshizawa & Kennett (2004) made use of simplified, 2-D, phase-velocity sensitivity kernels, with a uniform off-path sensitivity within the first one-third Fresnel zone, in a surface-wave tomographic study of Australia. On a global scale, Ritzwoller *et al.* (2002) have inverted for whole-earth group-velocity maps, using a simplified, uniform-across-the-path version of the 2-D sensitivity kernels developed by Spetzler *et al.* (2002).

In this paper, we investigate finite-frequency effects in global, long-period, surface-wave, phase-delay tomography, focusing upon three independent but related topics. First, we compare global

---

*Now at: Seismological Laboratory, California Institute of Technology, Pasadena, CA 91125, USA. E-mail yingz@gps.caltech.edu

tomographic results based upon traditional surface-wave ray theory and 3-D, finite-frequency sensitivity kernels. The finite-frequency kernels applied in this study are based upon the single-scattering (Born) approximation, formulated in the framework of surface-wave mode summation (Zhou *et al.* 2004). The resulting 3-D kernels fully account for the effects of directional scattering, the effects of seismogram windowing or tapering, and the effects of seismic source radiation. The preferred models of *S*-wave velocity and radial anisotropy in the upper mantle obtained using these finite-frequency sensitivity kernels are documented and discussed in a second paper (Zhou *et al.* 2005). In this paper, we show that, at the same level of model roughness, finite-frequency tomography fits a global dispersion data set better than ray-theoretical tomography, and that, because wave front healing and other diffractional effects are properly accounted for, anomalies with lateral length scales smaller than the dominant wavelength are better resolved in finite-frequency tomography. Second, we point out that surface-wave scattering due to lateral heterogeneities is directional dependent, and that this dependence varies with depth; as a result, phase-velocity maps that are obtained using approximated or simplified 2-D sensitivity kernels may have significant artefacts, and be no more reliable than those obtained using traditional ray theory. Third and finally, we develop finite-frequency sensitivity kernels for lateral variations in crustal thickness, and discuss the fundamentally intractable problem of making simple and accurate long-period, surface-wave crustal corrections.

2 THEORY

2.1 Ray-theoretical tomography

Traditional surface-wave tomography is based upon JWKB ray theory—a high-frequency approximation. Strictly speaking, ray theory is valid only if the angular frequency of the waves is infinitely high. In practice, the surface-wave JWKB approximation is acceptable if the lateral length scale of the 3-D velocity heterogeneity is large compared to the characteristic wavelength. In this case, measured surface-wave phase delays $\delta\phi(\omega)$ can be written as a linear path integral over the source–receiver great-circle ray:

$$\delta\phi(\omega) = -\frac{\omega}{c(\omega)} \int_0^\Delta \frac{\delta c(\omega, l)}{c(\omega)} \, dl, \tag{2.1}$$

where ω is the angular frequency, $c(\omega)$ is the unperturbed, spherical-earth surface-wave phase velocity, l is the distance along the great-circle ray, and Δ is the source–receiver epicentral distance. The frequency-dependent perturbation in surface-wave phase velocity δc can be expressed in terms of the perturbations in the shear-wave and compressional-wave velocities β and α through relations of the form

$$\delta c = \int_0^R \left(\frac{\partial c}{\partial \beta} \right) \delta\beta \, dr, \quad \text{Love waves}, \tag{2.2}$$

$$\delta c = \int_0^R \left[\left(\frac{\partial c}{\partial \beta} \right) \delta\beta + \left(\frac{\partial c}{\partial \alpha} \right) \delta\alpha \right] dr, \quad \text{Rayleigh waves}, \tag{2.3}$$

where the depth-dependent Fréchet derivatives $\partial c/\partial \beta$ and $\partial c/\partial \alpha$ (Dahlen & Tromp 1998, Chapter 11) are functions of radius r and angular frequency ω, and the integration extends from the centre of the Earth, $r = 0$, to the surface, $r = R$.

2.2 Finite-frequency tomography

In finite-frequency tomography, surface-wave phase delays are sensitive not only to heterogeneities along the source–receiver great-circle ray, but also to anomalies that are off the reference ray. In this paper, we apply the 3-D, finite-frequency sensitivity kernels for fundamental-mode, surface-wave dispersion measurements developed by Zhou *et al.* (2004). The measured phase delay at a specific angular frequency ω is written as a volumetric integration over all velocity heterogeneities in the 3-D earth $\oplus$, so that eqs (2.1)–(2.3) are replaced by:

$$\delta\phi(\omega) = \iiint_\oplus K_\beta(\omega, \mathbf{x}) \frac{\delta\beta}{\beta}(\mathbf{x}) \, d^3\mathbf{x}, \quad \text{Love waves} \tag{2.4}$$

$$\delta\phi(\omega) = \iiint_\oplus \left[K_\beta(\omega, \mathbf{x}) \frac{\delta\beta}{\beta}(\mathbf{x}) + K_\alpha(\omega, \mathbf{x}) \frac{\delta\alpha}{\alpha}(\mathbf{x}) \right] d^3\mathbf{x},$$
$$\text{Rayleigh waves}, \tag{2.5}$$

where we have ignored the effects of density perturbations for simplicity, based on the following considerations: (1) if we assume mantle anomalies are dominantly thermal, velocity variations are much more significant than density variations; and (2) fundamental-mode surface-wave phase delays are most sensitive to velocity perturbations (Zhou *et al.* 2004).

The 3-D sensitivity kernels $K_\beta(\omega, \mathbf{x})$ and $K_\alpha(\omega, \mathbf{x})$ are computed using the single scattering approximation; the effects of surface-wave cross-branch mode coupling can be fully taken into account by a double summation over all surface-wave modes at the frequency of interest (Zhou *et al.* 2004). However, our previous studies have suggested that the importance of surface-wave cross-branch mode coupling depends upon the desired spatial resolution, and that mode-coupling effects may be neglected at the current stage of global dispersion tomography, because (1) the spatial resolution is limited by global path coverage and (2) the number of measurements affected by significant surface-wave mode coupling is likely to be small in hand-picked data sets (Zhou *et al.* 2004). For these reasons, we ignore the effects of cross-branch mode coupling in this tomographic comparison study; this makes the computation of the finite-frequency sensitivity kernels very efficient. The factors that determine the geometry of the phase-delay kernels have been discussed in detail by Zhou *et al.* (2004); in brief, the cross-path width of the region of significant sensitivity is proportional to the square root of the period of the wave. The frequency dependence of the shear-velocity sensitivity kernel $K_\beta(\omega, \mathbf{x})$ is illustrated in Fig. 1, where it is seen that 5-mHz Love waves show a cross-path sensitivity that is $\sqrt{3}$ times broader than that of 15-mHz waves. The 3-D sensitivity kernels applied in this study are computed using the spherically symmetric reference earth model 1066A (Gilbert & Dziewonski 1975).

3 TOMOGRAPHY

3.1 Data and model parametrization

To illustrate the importance of accounting for finite-frequency wave propagation effects in surface-wave tomography, we invert a global data set of fundamental-mode Love-wave and Rayleigh-wave dispersion (Laske & Masters 1996), which include waves that have propagated along minor arcs (G1, R1) and major arcs (G2, R2), as well as multiorbits (G3, G4, R3, R4). The phase-delay measurements are made using a multitaper technique to reduce the bias in spectral
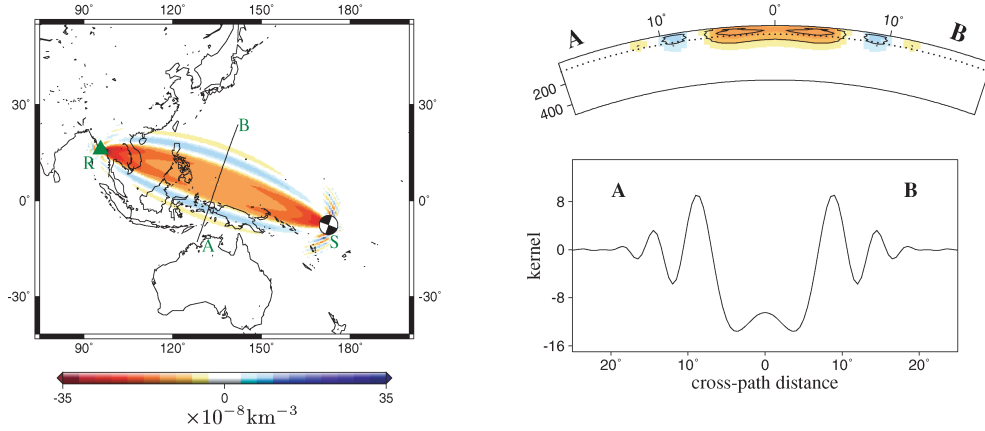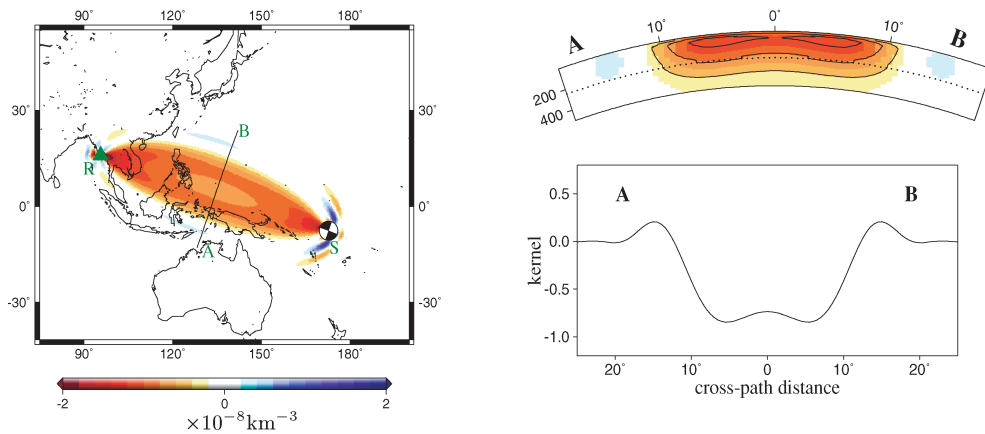
(a) 15-mHz Love wave K_β

(b) 5-mHz Love wave K_β

Figure 1. 3-D phase-delay sensitivity kernel K_β for (a) 15-mHz and (b) 5-mHz fundamental-mode Love waves. Map views of the 15- and 5-mHz sensitivity kernels K_β are plotted at 80 and 250 km depth, respectively (dotted lines in the cross-sections). The cross-path variations in sensitivity at those depths are plotted below the cross-sections. The seismic source is strike-slip, with Love-wave radiation symmetric about the source–receiver great-circle path (see beachball). The epicentral distance is 80°. Letters S, R, A and B denote the source, receiver and endpoints of the cross-section, respectively. Surface-wave mode coupling effects are ignored.

estimates (Laske & Masters 1996). This data set includes a total of ~12 000 wave trains, including both Love waves and Rayleigh waves. The frequencies of the dispersion measurements range from 5 to 15 mHz, with an interval of 1 mHz. All measurements are residuals with respect to model 1066A (Gilbert & Dziewonski 1975).

We parametrize the surface of the Earth using a set of spherical triangular grid points (Baumgardner & Frederickson 1985). The triangulation starts with 20 equal spherical triangles, and is iteratively refined by connecting the midpoint of the three sides of each triangle (Fig. 2). The spherical triangles used in this study are 16-fold, with 2562 vertexes and 5120 triangles. The lateral spacing between neighbouring grid points is $4.3° \pm 0.3°$. We assume that the seismic velocity within any spherical triangle can be approximated by linear spatial interpolations through the velocities at the three vertexes. In the radial (depth) direction, the uppermost 580 km of the upper mantle is parametrized using nine grid points, with a nominal grid spacing of 60 km in the uppermost 400 km; velocity perturbations are interpolated linearly between neighbouring depth grid points.

3.2 The inverse problem

We invert for the fractional *S*-wave velocity perturbations $\delta\beta/\beta$ at the $2562 \times 9 = 23\,058$ grid points. The *P*-wave velocity perturbations are scaled to the *S*-wave velocity perturbations using the relation $\delta\alpha/\alpha = 0.5\,(\delta\beta/\beta)$, adapted from laboratory measurements based on petrological models of the upper mantle (Montagner & Anderson 1989). We have experimented with scaling parameters in the range $0 \leq (\delta\alpha/\alpha)/(\delta\beta/\beta) \leq 1$, and find that the tomographic results do not vary significantly. The discrete forms of both ray-theoretical tomography, eqs (2.1)–(2.3), and finite-frequency tomography, eqs (2.4)–(2.5), can be written in the canonical form

$$\mathbf{Am} = \mathbf{b}, \tag{3.1}$$

where $\mathbf{A}$ is the sensitivity kernel matrix, $\mathbf{m}$ is the vector of unknown parameters $\delta\beta/\beta$, and $\mathbf{b}$ is the phase-delay data vector. To compute the finite-frequency kernel matrix $\mathbf{A}$, the numerical integrations in eqs (2.4)–(2.5) are carried out with spatial samplings that are fine enough to avoid possible aliasing in the sensitivity kernels; all of
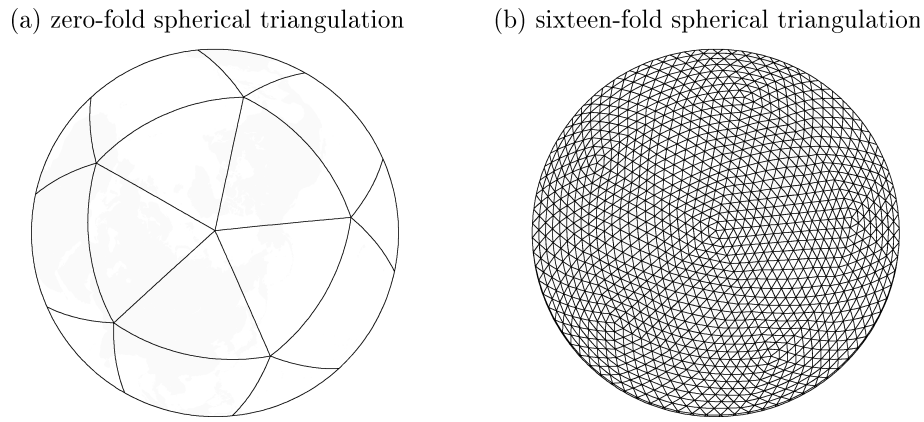
Figure 2. Icosahedral spherical triangulation (Baumgardner & Frederickson 1985). (a) The starting twenty equal spherical triangles. (b) The refined final spherical triangles after a 16-fold triangulation; the spatial separation of neighbouring grid points is $4.3° \pm 0.3°$.

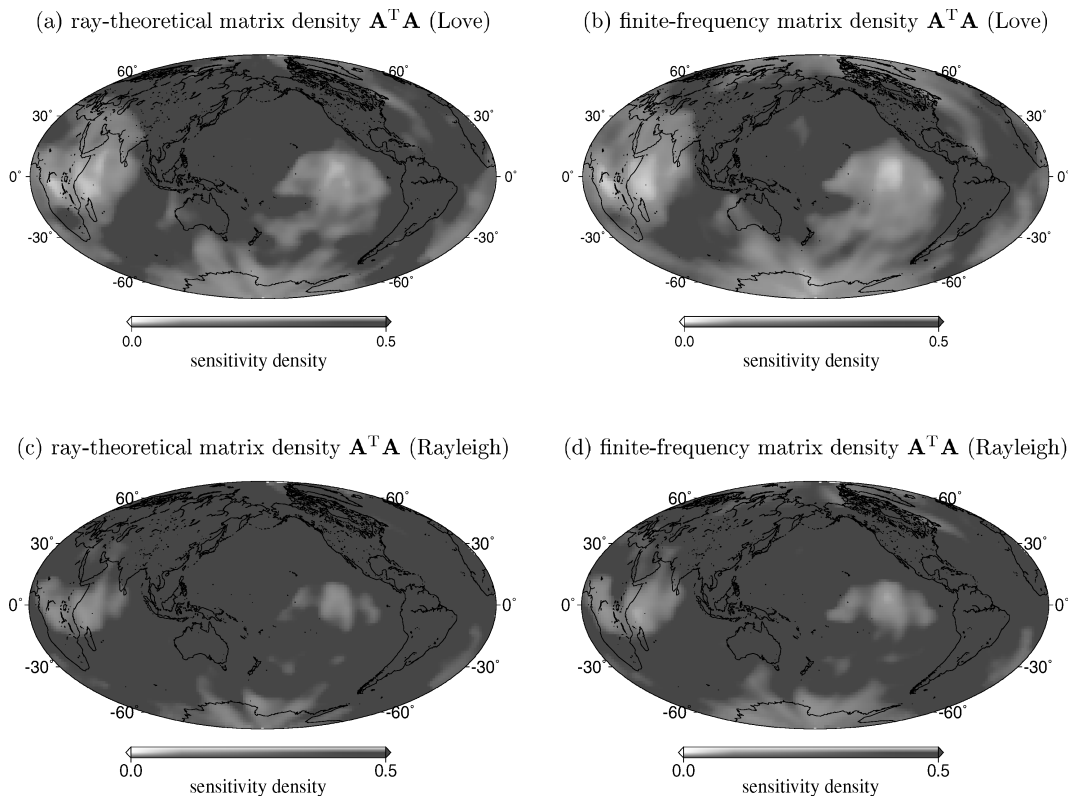**Figure 3.** Diagonal elements of the matrix products $\mathbf{A}^\mathrm{T}\mathbf{A}$ in ray-theoretical (left) and finite-frequency (right) tomography, plotted at a depth of 180 km. In the top two maps, (a) and (b), the matrix $\mathbf{A}$ includes sensitivities of Love waves with frequencies from 5 to 15 mHz and wave trains G1, G2, G3 and G4. In the bottom two maps, (c) and (d), the matrix $\mathbf{A}$ includes sensitivities of Rayleigh waves with frequencies from 5 to 15 mHz and wave trains R1, R2, R3 and R4.

the sidebands as well as spatial sensitivity variations within the first Fresnel zone are fully accounted for in the computation. The diagonal elements of the ray-theoretical and finite-frequency matrix products $\mathbf{A}^\mathrm{T}\mathbf{A}$, a measure of the spatial coverage of the data, are plotted in Fig. 3. Naively, we might expect the matrix density $\mathbf{A}^\mathrm{T}\mathbf{A}$ to be more uniform in finite-frequency tomography, because of the off-ray spreading of the sensitivity. In practice, since the majority of the great-circle ray paths are not near the maximum of the outgoing source radiation, the finite-frequency spatial coverage is comparable to that of ray theory.

Due to limited and uneven global path coverage, as well as contamination of the data by noise, the tomographic equation is under-determined, and the inverse matrix $(\mathbf{A}^\mathrm{T}\mathbf{A})^{-1}$ is singular or ill-posed. In this study, we apply a Laplacian smoothing to regularize the least-squares inverse problem, $(\mathbf{Am} - \mathbf{b})^\mathrm{T}(\mathbf{Am} - \mathbf{b}) = \text{minimum}$, and we simultaneously invert for perturbations to the source origin times $\mathbf{t} = (\ldots, t_k, \ldots)$, so that the tomographic problem becomes

$$\chi^2 + \epsilon \, ||\mathbf{Sm}||^2 + \epsilon_t ||\mathbf{t}||^2 = \text{minimum},$$

$$\text{where} \quad \chi^2 = \sum_{i=1}^{N} \left(\frac{A_{ij}m_j + H_{ik}t_k - b_i}{\sigma_i} \right)^2. \tag{3.2}$$

The quantity N is the total number of phase-delay measurements, and χ^2 is the data misfit function, with σ_i being the observational

error associated with the *i*th measurement b_i. The product $H_{ik}t_k$ represents the phase delay in the *i*th observation due to a shift t_k in the origin time of the *k*th earthquake. Summation over the model parameter index *j* and the earthquake index *k* in eq. (3.2) is understood. The quantity

$$||\mathbf{Sm}|| \approx \left(\iiint_\oplus |\nabla^2(\delta\beta/\beta)|^2 \, d^3\mathbf{x} \right)^{1/2} \tag{3.3}$$

is an approximate, discretized measure of the root-mean-square (rms) Laplacian of the fractional velocity perturbation. The smoothing parameter ϵ governs the trade-off between the model Laplacian of the inverted model and the misfit to the data; We choose an origin-time damping parameter ϵ_t such that the perturbations t_k are not greater than ± 4 s; because the surface waves used in this study are long-period (67–200 s), the differences between tomographic models inverted with and without origin-time corrections are negligible. The sparse linear tomographic problem, eq. (3.2), is solved using the LSQR algorithm (Paige & Saunders 1982), iterated until full convergence. If the observations are contaminated by noise with a Gaussian distribution, the data misfit χ^2 of the recovered earth model should approach the number of observations, that is, $\chi^2/N = 1$.

We apply crustal corrections to the global phase-delay data set, based on a global crust model, CRUST2.0 (Laske *et al.* 2001), which specifies a seven-layer crust in each 2° by 2° grid. Maps of the local phase-delay corrections due to the CRUST2.0 crustal variations are computed at each of the discrete frequencies of measurement,

5–15 mHz, and ray-theoretical crustal corrections are then calculated for each phase-delay observation by integration along the ray path. In Section 6, we show that finite-frequency crustal effects upon long-period surface waves are not negligible; however, boundary topography sensitivity kernels that are based upon the Born approximation are not sufficiently accurate to account for the large global variations in Moho depth. The observational errors σ_i in eq. (3.2) are estimated from phase-delay measurements of closely spaced earthquake pairs, based upon the assumption that the surface waves excited by these repeating earthquakes sample approximately the same region of the upper mantle, and therefore should experience nearly equal phase delays. The rms of the estimated observational errors is about 45–75 per cent of the rms of the measured phase delays (Zhou *et al.* 2005).

4 FINITE-FREQUENCY EFFECTS IN 3-D TOMOGRAPHY

In this section, we compare *S*-wave velocity models obtained by inversion of the same phase-delay measurements using both finite-frequency and ray-theoretical kernel matrices **A**. Fig. 4 shows the trade-offs between the model roughness (normalized model Laplacian) $||\mathbf{Sm}||/||\tilde{\mathbf{m}}||$ and the data misfit χ^2/N, where the quantity $||\tilde{\mathbf{m}}|| = ||\mathbf{m}||/\sqrt{N}$ is the rms of the velocity perturbation $\delta\beta/\beta$. The roughness of the model is insensitive to the perturbation norm $||\tilde{\mathbf{m}}||$. In both ray-theoretical and finite-frequency tomography, the normalized data misfit χ^2/N is greater than one, which indicates
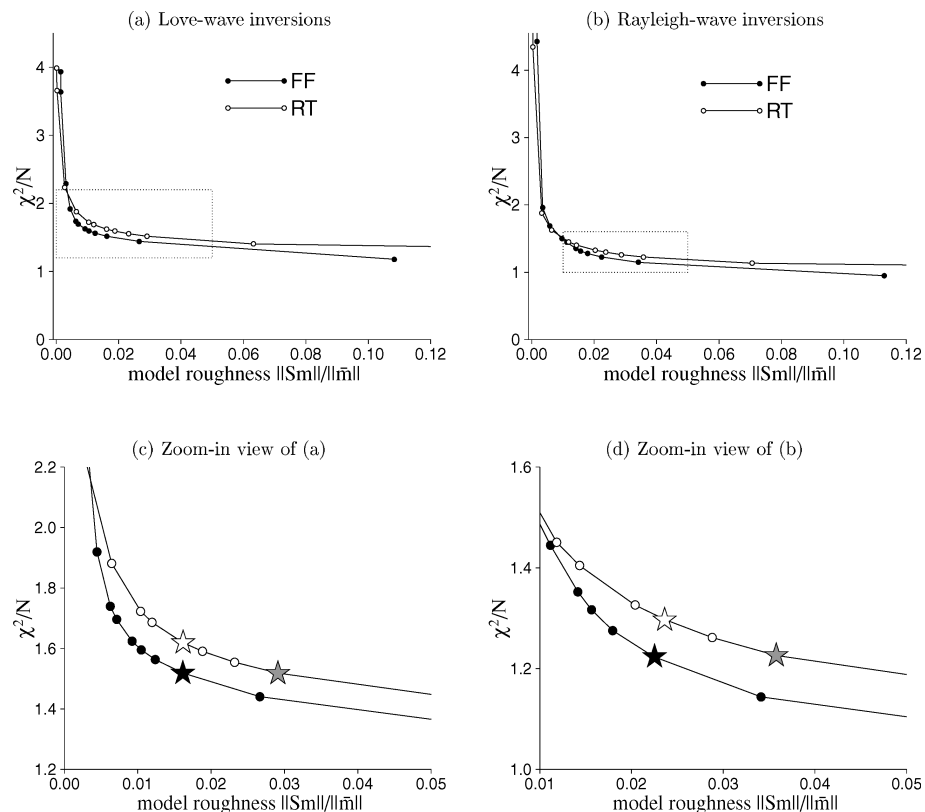

Figure 4. Misfit versus roughness trade-off curves for finite-frequency (FF) and ray-theoretical (RT) tomography. In (a) and (b) both the Love and Rayleigh 'no-smoothing' ($\epsilon = 0$) finite-frequency models have roughness $||\mathbf{Sm}||/||\tilde{\mathbf{m}}|| \approx 0.11$; the corresponding ray-theoretical 'no-smoothing' models are out of the scale, with roughnesses $||\mathbf{Sm}||/||\tilde{\mathbf{m}}|| \approx 0.52$ and $||\mathbf{Sm}||/||\tilde{\mathbf{m}}|| \approx 0.60$, for Love and Rayleigh waves, respectively. Plots (c) and (d) show blowups of the regions outlined by the dotted boxes in (a) and (b). The black stars indicate the chosen 'optimal' finite-frequency models discussed in the paper; the white stars are ray-theoretical models with the same roughness $||\mathbf{Sm}||/||\tilde{\mathbf{m}}||$, and the grey stars are ray-theoretical models with the same misfit χ^2/N. Map views of the starred tomographic models are plotted in Figs 5–8.

that the errors assigned to the phase-delay measurements are underestimated. In tomographic practice, it is often difficult to estimate the actual error in the data; furthermore, the assumption that the noise is Gaussian may be invalid. For these reasons, we do not rescale the estimated errors in the data to force $\chi^2/N = 1$, but instead accept models with misfits χ^2/N slightly greater than one.

4.1 3-D model comparison: kernels versus rays

The 'optimal' finite-frequency tomographic model discussed in this paper is chosen to be near the lower left corner of the χ^2/N versus $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$ trade-off curve, where small reductions in the misfit to the data begin to require large increases in the roughness of the model. We compare the 'optimal' finite-frequency model (black stars in Fig. 4) with two different ray-theoretical models, one chosen to have the same roughness $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$ (white stars) and the other to have the same data misfit χ^2/N (grey stars). Map views of the velocity models obtained by independent inversions of the Love-wave and the Rayleigh-wave data are plotted in Figs 5–8.

The poor agreement between the models obtained by inverting Love and Rayleigh phase delays confirms that there is strong radial anisotropy in the upper mantle. A joint Love–Rayleigh inversion using isotropic kernels K_β and K_α results in a poor data fit,

with systematic Love–Rayleigh discrepancies, and, only the largest-scale common structures are well constrained. The inability of a single isotropic S-wave velocity model to provide a satisfactory fit to both Love and Rayleigh waves has been widely reported in previous surface-wave studies (e.g. Anderson 1961; Aki & Kaminuma 1963). In this paper we confine ourselves to independent inversions of Love and Rayleigh data, and focus upon the comparison between traditional surface-wave ray theory and finite-frequency sensitivity kernels. In both the Love-wave and Rayleigh-wave inversions, finite-frequency tomographic models exhibit a better fit to the data than ray-theoretical models, at the same level of model roughness $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$ (this is most evident in the magnified trade-off curves in Fig. 4).

Figs 5 and 6 show map views of $\delta\beta/\beta$ at depths of 120, 250 and 390 km, obtained by inverting Love-wave phase delays. The ray-theoretical model in Fig. 5 has the same roughness, $||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.016$, as the 'optimal' finite-frequency model, whereas the model in Fig. 6 has the same data misfit, $\chi^2/N \approx 1.51$. The long-wavelength features in the finite-frequency and ray-theoretical Love-wave models are in good agreement; in particular, all three models are characterized by slow anomalies beneath midocean ridges and fast anomalies beneath continental cratons. A number of small-scale velocity anomalies that are weak or absent in the ray-theoretical
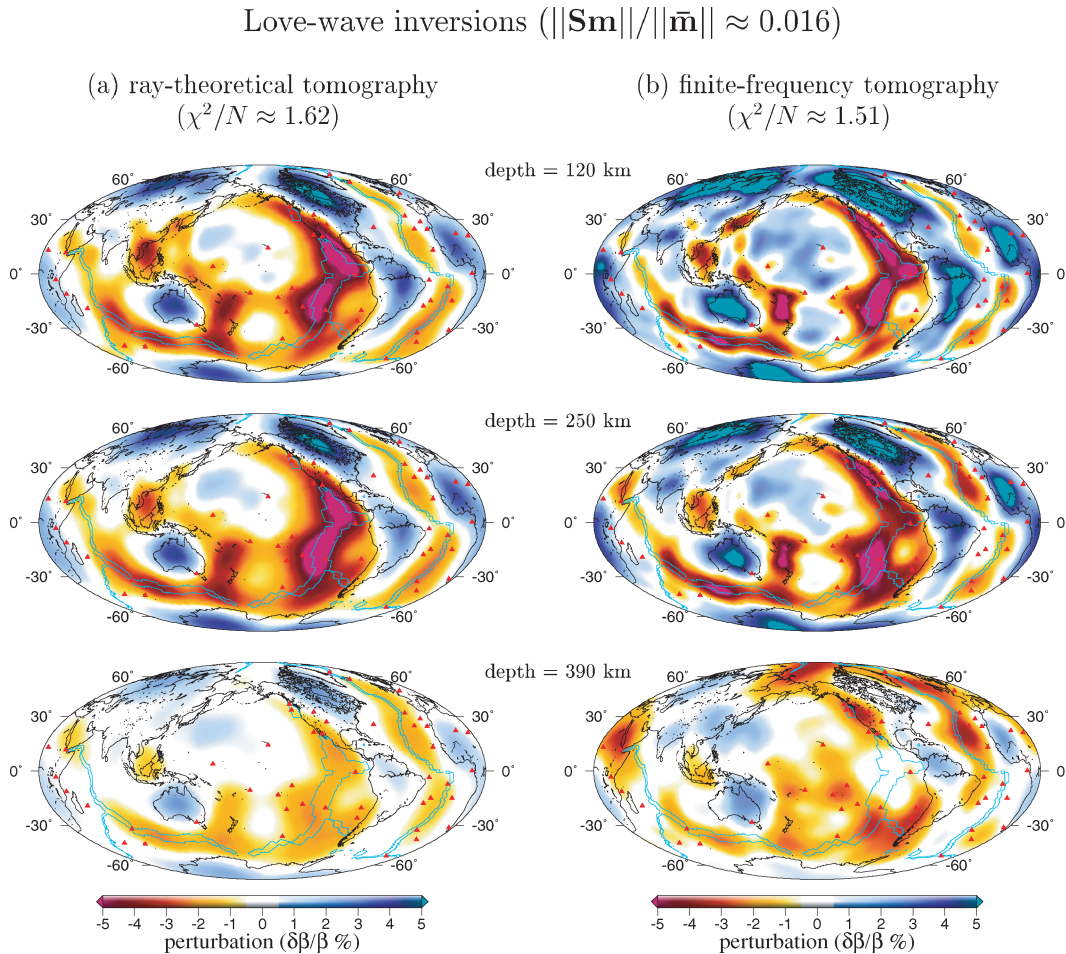
Love-wave inversions ($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.016$)

(a) ray-theoretical tomography
($\chi^2/N \approx 1.62$)

(b) finite-frequency tomography
($\chi^2/N \approx 1.51$)

Figure 5. *S*-wave velocity model obtained by (a) ray-theoretical tomography and (b) finite-frequency tomography. The two models have the same model roughness, $||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.016$, with different values of χ^2/N, as indicated. The velocity perturbations are with respect to model 1066A, with the spherical average in each map removed for better illustration. Models are inverted using Love-wave phase delays, and are plotted at depths of 120, 250 and 390 km. Hotspots are indicated by red triangles, and the age contours (blue) indicate 10 Myr old seafloor. The two models agree in their large-scale structure; however, anomalies in the finite-frequency tomographic model are stronger in amplitude, with small-scale features better resolved.
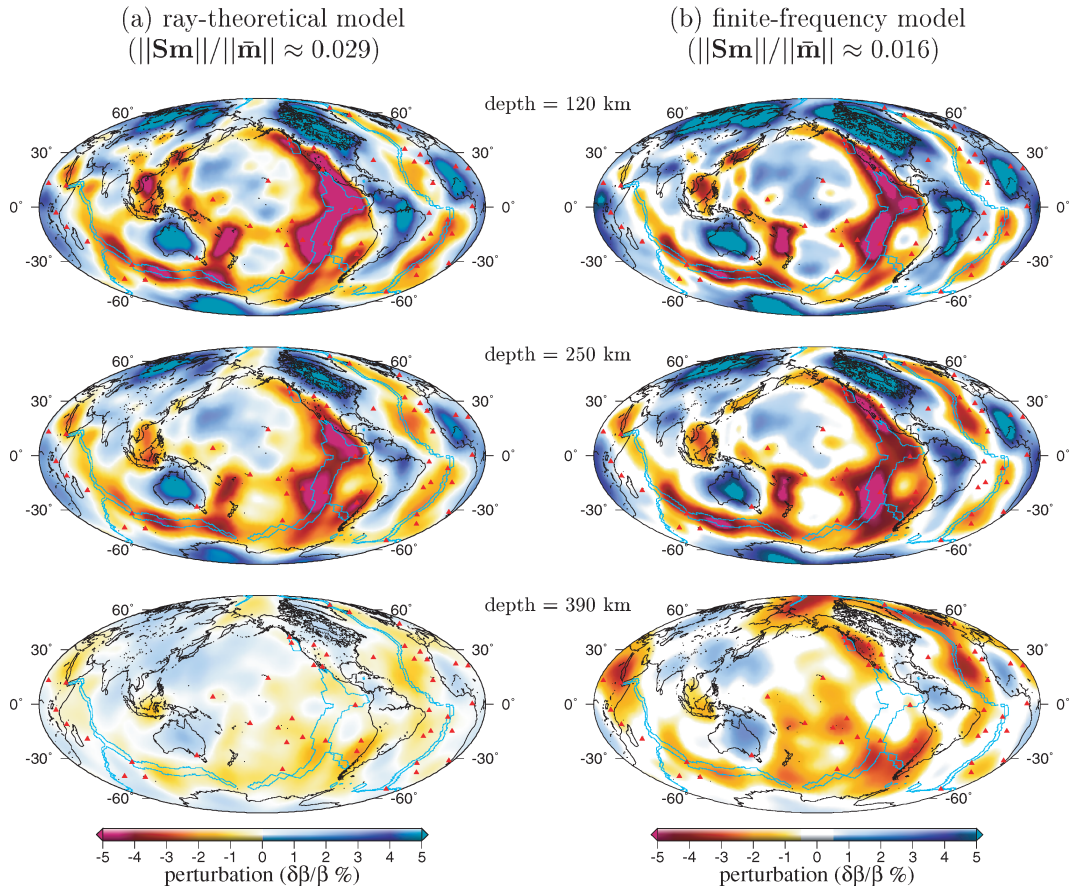
Love wave inversions ($\chi^2/N \approx 1.51$)

(a) ray-theoretical model
($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.029$)

(b) finite-frequency model
($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.016$)

Figure 6. Same as Fig. 5, expect that the two models now fit the data equally well, $\chi^2/N \approx 1.51$, with different values of the model roughness $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$, as indicated. The *S*-wave velocity anomalies are stronger in the finite-frequency model, especially in the deeper part of the upper mantle.

models can be more easily identified using finite-frequency tomography; notable examples are the slow anomalies at 390 km depth beneath the southern Pacific superswell, the Red Sea, the Basin and Range Province, and the North Mid-Atlantic Ocean. Figs 7–8 show map-view comparisons of the equal-roughness, $||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.022$, and equal-misfit, $\chi^2/N \approx 1.22$, models obtained by inverting Rayleigh-wave phase delays. The essential observations are similar to those in the Love-wave comparisons. In the central Pacific, the differences between the ray-theoretical and finite-frequency Rayleigh-wave models are significant in the middle portion of the upper mantle; the western Pacific subduction zones are better resolved in the finite-frequency model, particularly at 390 km depth. Geographically, there does not appear to be any correlation between the ray-path density (Fig. 3) and the differences between the ray-theoretical and finite-frequency tomographic velocity models. For example, the models differ significantly in the southern Pacific Ocean, where the ray-path coverage is relatively poor; on the other hand, there are also large differences beneath western North America, where the coverage is relatively dense.

In both the Love-wave and Rayleigh-wave comparisons, the finite-frequency model has stronger small-scale anomalies than either the equal-roughness or the equal-misfit ray-theoretical model. This is mainly because the 3-D sensitivity kernels applied in finite-

frequency tomography account for the effects of wave front healing, that is, a diverging or converging surface wave heals as it propagates beyond the causative fast or slow velocity anomaly. As a result, the phase delays observed at seismic stations are generally smaller than the original phase delay incurred by a wave upon passage through an anomaly. In ray theory, wave front healing effects are not accounted for, and as a result, the recovered anomalies are generally weaker in amplitude. Analogous finite-frequency effects in body-wave tomography have been noted by Montelli *et al.* (2004) and Hung *et al.* (2004). The main result is the same for surface waves as for body waves: finite-frequency tomography generally recovers stronger small-scale anomalies than ray theory.

The mean velocity perturbation at every depth has been subtracted, so that the lateral variations in all of the maps 5–8 average to zero. The updated, spherically averaged velocity profiles $\beta(r)$ obtained by finite-frequency tomography and ray-theoretical tomography are compared in Fig. 9. The degree-zero profiles $\beta(r)$ are seen to be almost identical, so maps in Figs 5–8 can be properly compared. The rms velocity perturbations $\langle(\delta\beta/\beta)^2\rangle^{1/2}$ are also plotted versus depth in Fig. 9. Regardless of whether the ray-theoretical model is constrained to have the same roughness, $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$, or the same data misfit, χ^2/N, the rms velocity anomalies in the 'optimal' finite-frequency model are larger at all depths, by as much as 30 per cent.
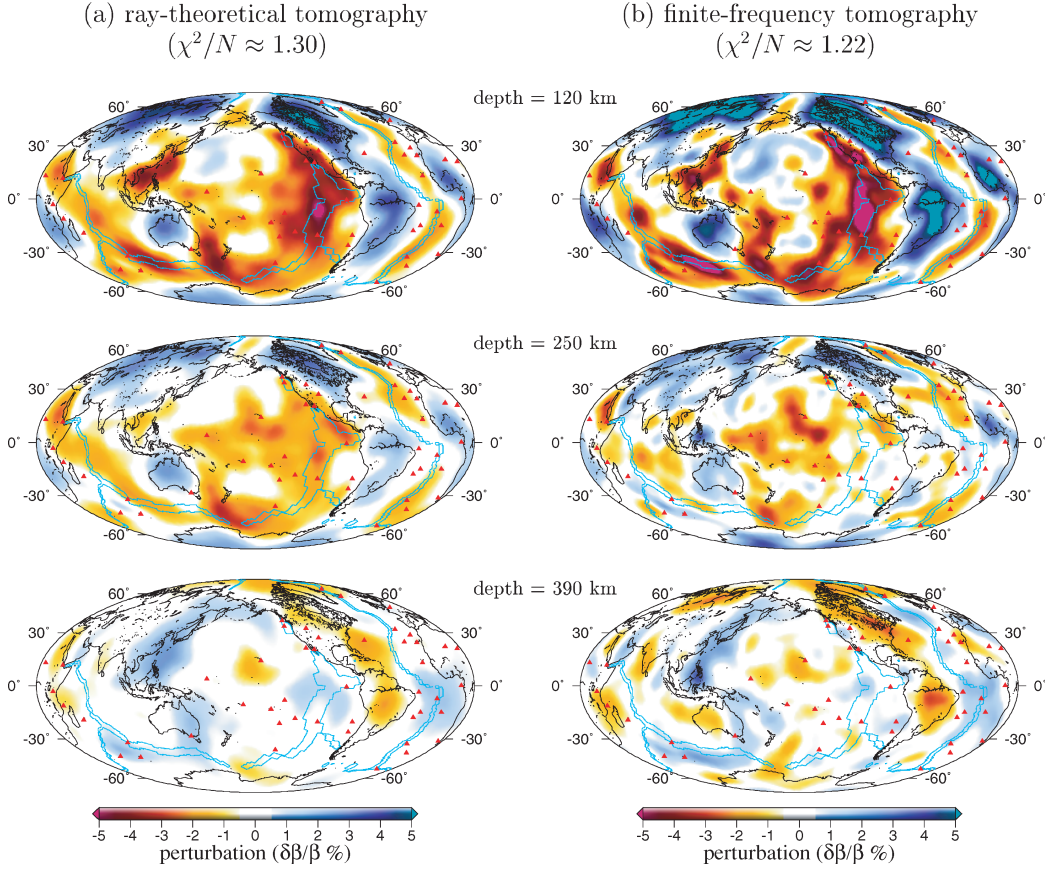
Rayleigh-wave inversions ($\|\mathbf{Sm}\|/\|\bar{\mathbf{m}}\| \approx 0.022$)

(a) ray-theoretical tomography
($\chi^2/N \approx 1.30$)

(b) finite-frequency tomography
($\chi^2/N \approx 1.22$)

depth = 120 km

depth = 250 km

depth = 390 km

perturbation ($\delta\beta/\beta$ %)

perturbation ($\delta\beta/\beta$ %)

Figure 7. Same as Fig. 5, except for Rayleigh waves rather than Love waves. The two models have the same roughness $\|\mathbf{Sm}\|/\|\bar{\mathbf{m}}\| \approx 0.022$.

4.2 Spectral power and model correlations

To further investigate the enhancement of small-scale features in finite-frequency tomography, we decompose the *S*-wave velocity maps into spherical harmonics. We use the real spherical harmonic convention of Dahlen & Tromp (1998, Section B.8), in which a real, zero-mean, function of geographical position $\psi(\theta, \phi)$ is expanded in the form

$$\psi(\theta, \phi) = \sum_{l=1}^{\infty} \left[a_{l0} X_{l0}(\theta) + \sqrt{2} \right.$$
$$\left. \sum_{m=1}^{l} X_{lm}(\theta)(a_{lm} \cos m\phi + b_{lm} \sin m\phi) \right], \quad (4.1)$$

where the orthonormality relation governing the real colatitudinal harmonics $X_{lm}(\theta)$ is

$$\int_0^{\pi} X_{lm}(\theta) X_{l''}(\theta) \sin\theta \, d\theta = \frac{1}{2\pi} \delta_{ll'}. \quad (4.2)$$

The power per degree *l* and per unit area is defined by

$$P_l = \frac{1}{2l+1} \left[a_{l0}^2 + \sum_{m=1}^{l} \left(a_{lm}^2 + b_{lm}^2 \right) \right], \quad (4.3)$$

and the correlation coefficient C_l is defined by

$$C_l = \frac{a_{l0} a_{l0}' + \sum_{m=1}^{l} (a_{lm} a'_{lm} + b_{lm} b'_{lm})}{\sqrt{\sum_{m=1}^{l} \left(a_{lm}^2 + b_{lm}^2 \right)} \sqrt{\sum_{m=1}^{l} \left(a'^2_{lm} + b'^2_{lm} \right)}}, \quad (4.4)$$

where a_{lm}, b_{lm} and a'_{lm}, b'_{lm} are the real spherical expansion harmonic coefficients of the two velocity maps being compared. The power spectral density P_l of the velocity perturbations $\delta\beta/\beta$, as well as the finite-frequency versus ray-theoretical correlation coefficients C_l are plotted versus spherical harmonic degree *l*, at the three depths 120, 250 and 390 km, in Figs 10 and 11. The spectral comparisons confirm that the imaged velocity anomalies are in general 'stronger' in finite-frequency tomography. The correlation C_l between the ray-theoretical and finite-frequency models decreases rapidly with increasing spherical harmonic degree above $l \approx$ 10; in addition, the correlation is markedly weaker at greater depth (390 km) in both the Love-wave and the Rayleigh-wave inversions. These are both expected results, inasmuch as finite-frequency effects become more significant with increasing wavelength, and deep-seated structures are mainly constrained by long-period surface waves. In Fig. 10, where the two models have the same roughness $\|\mathbf{Sm}\|/\|\bar{\mathbf{m}}\|$, there is a clear and growing excess of high-degree power in the finite-frequency model; this scale dependence of the difference in P_l versus *l* is not as pronounced in Fig. 11, where the two models fit the phase-delay data equally well but have different roughness.

4.3 Checkboard tests: scale-dependent resolution

In this subsection we investigate the resolution and fidelity of ray-theoretical and finite-frequency tomography using synthetic checkerboard tests. The input checkboards consist of five

Rayleigh wave inversions ($\chi^2/N \approx 1.22$)

(a) ray-theoretical model
($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.036$)

(b) finite-frequency model
($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.022$)
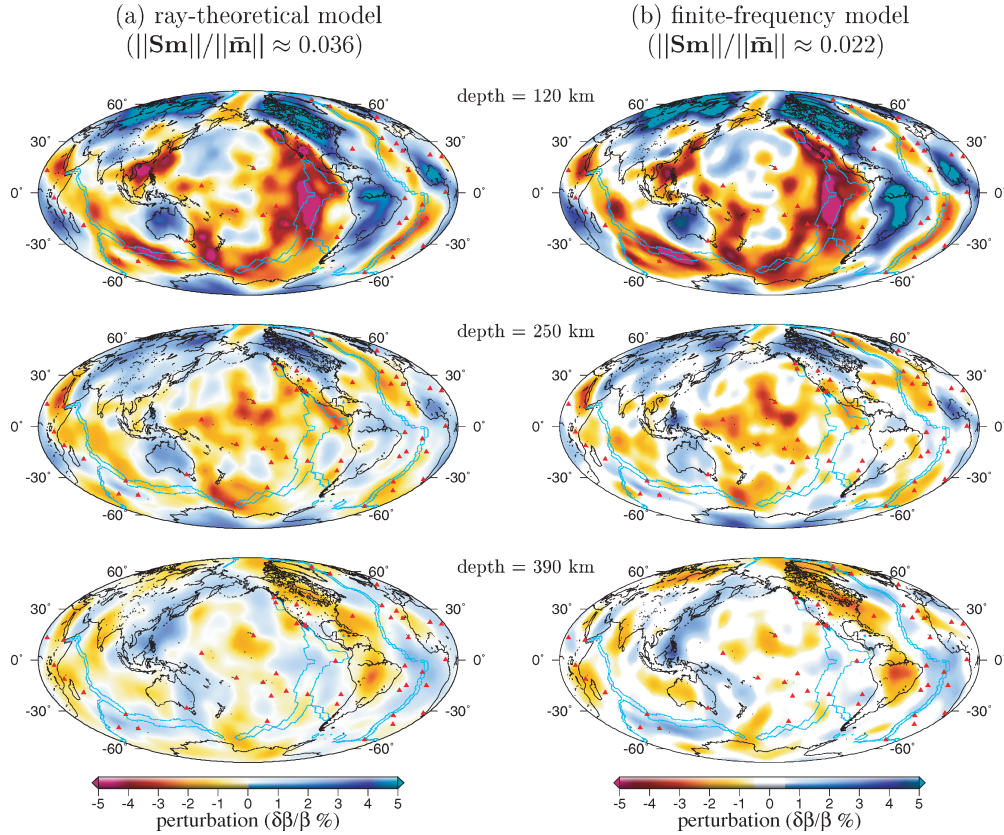

Figure 8. Same as Fig. 6, except for Rayleigh waves rather than Love waves. The finite-frequency and ray-theoretical models have the same data misfit, $\chi^2/N \approx 1.22$. Small-scale structures differ significantly between the two models throughout the upper mantle.
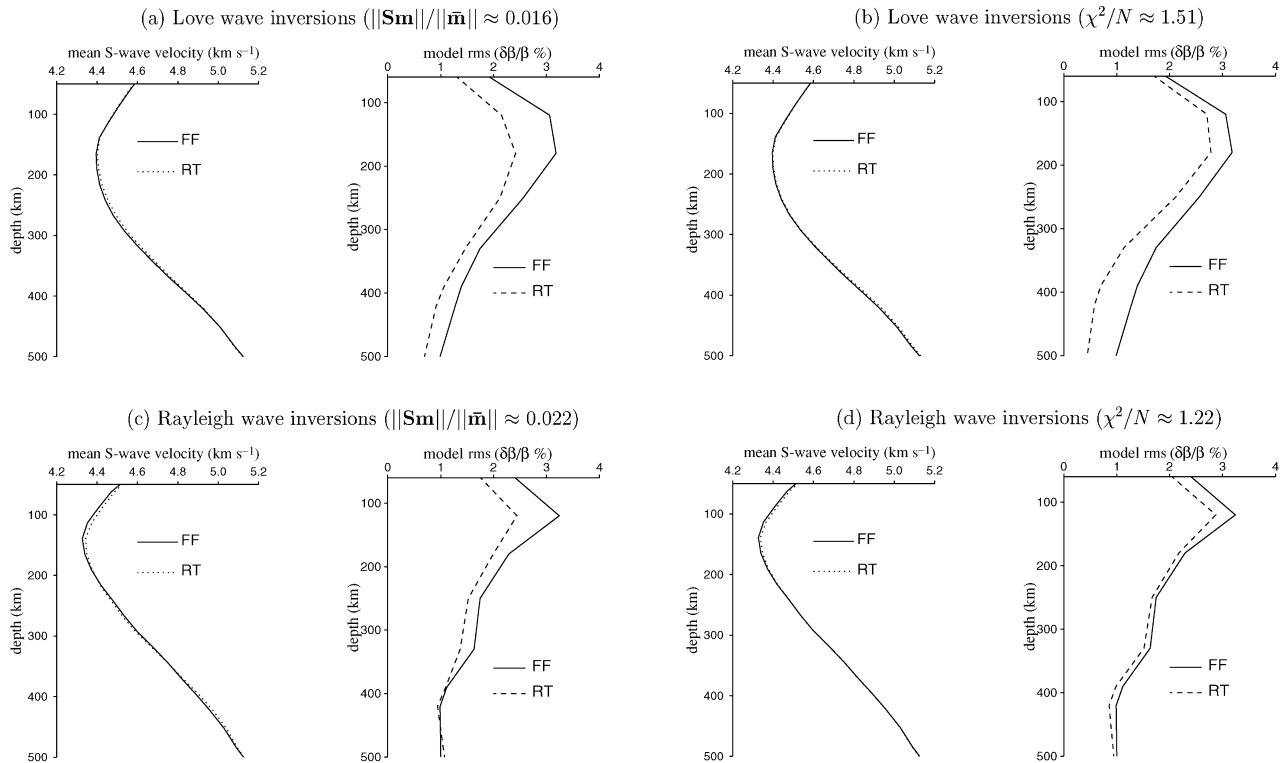
(a) Love wave inversions ($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.016$)

(b) Love wave inversions ($\chi^2/N \approx 1.51$)

(c) Rayleigh wave inversions ($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.022$)

(d) Rayleigh wave inversions ($\chi^2/N \approx 1.22$)

Figure 9. Depth variation of the spherically averaged *S*-wave velocity β and the rms variation of the 3-D perturbations, $\langle(\delta\beta/\beta)^2\rangle^{1/2}$, for both finite-frequency (FF) and ray-theoretical (RT) tomography.

(a) Love-wave inversions ($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.016$)

(b) correlation between ray-theoretical and finite-frequency models

(c) Rayleigh-wave inversions ($||\mathbf{Sm}||/||\bar{\mathbf{m}}|| \approx 0.022$)

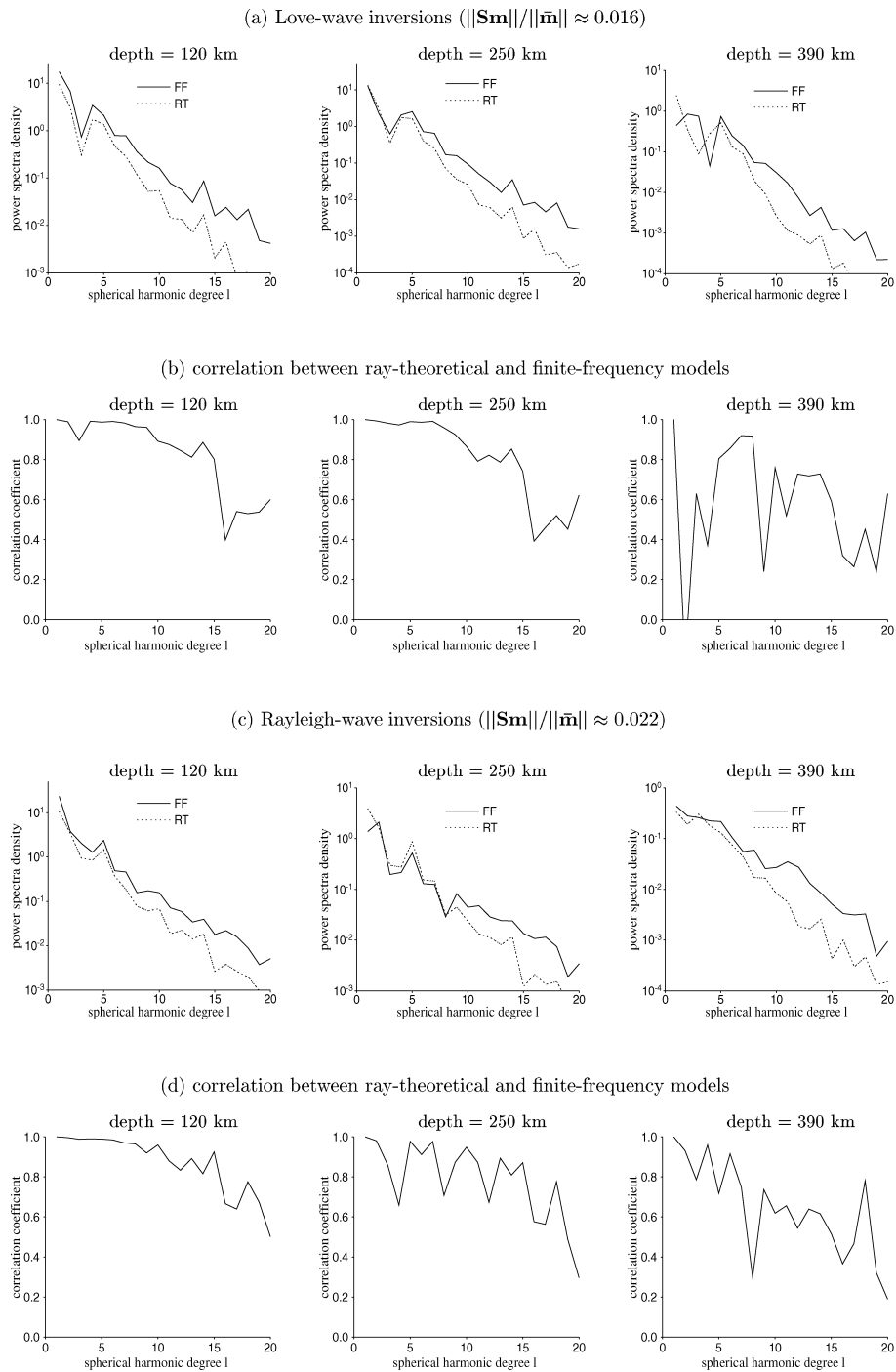(d) correlation between ray-theoretical and finite-frequency models

Figure 10. Plots (a) and (c) show the spherical-harmonic power spectral density P_l of finite-frequency (FF) and ray-theoretical (RT) *S*-wave velocity models of the same roughness $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$; plots (b) and (d) show the correlation coefficient C_l between the same two models. Map views of these equal-roughness velocity models are plotted in Figs 5 and 7. In general, finite-frequency tomographic models (solid lines) have higher power than ray-theoretical tomographic models (dashed lines), especially at shorter length scales, with spherical harmonic degrees $l > 10$. The correlation between ray-theoretical and finite-frequency models decreases rapidly for small-scale structures with spherical harmonic degree $l > 10$; in addition, the correlation becomes significantly weaker at greater depth.

high-velocity and five low-velocity Gaussian *S*-wave anomalies; The anomaly amplitudes are maximum at a depth $z_0 = 180$ km, and fall off exponentially above and below, $A = A_0 e^{-f|z-z_0|}$, with a decay rate $f = 0.0125$ km^{-1}. Synthetic Love-wave phase delays are computed using the finite-frequency relation in eq. (2.4), with the same source–receiver and wave train configurations (G1, G2,

G3, G4) as in the Love-wave global data set used in the above study. To mimic the measurement errors in the actual phase delay data, we add Gaussian noise to the synthetic delays; the rms of the noise is chosen to be about 50 per cent of the rms of the signal. To understand the scale dependence of the resolution, we use synthetic checkerboards with anomalies of two different sizes:
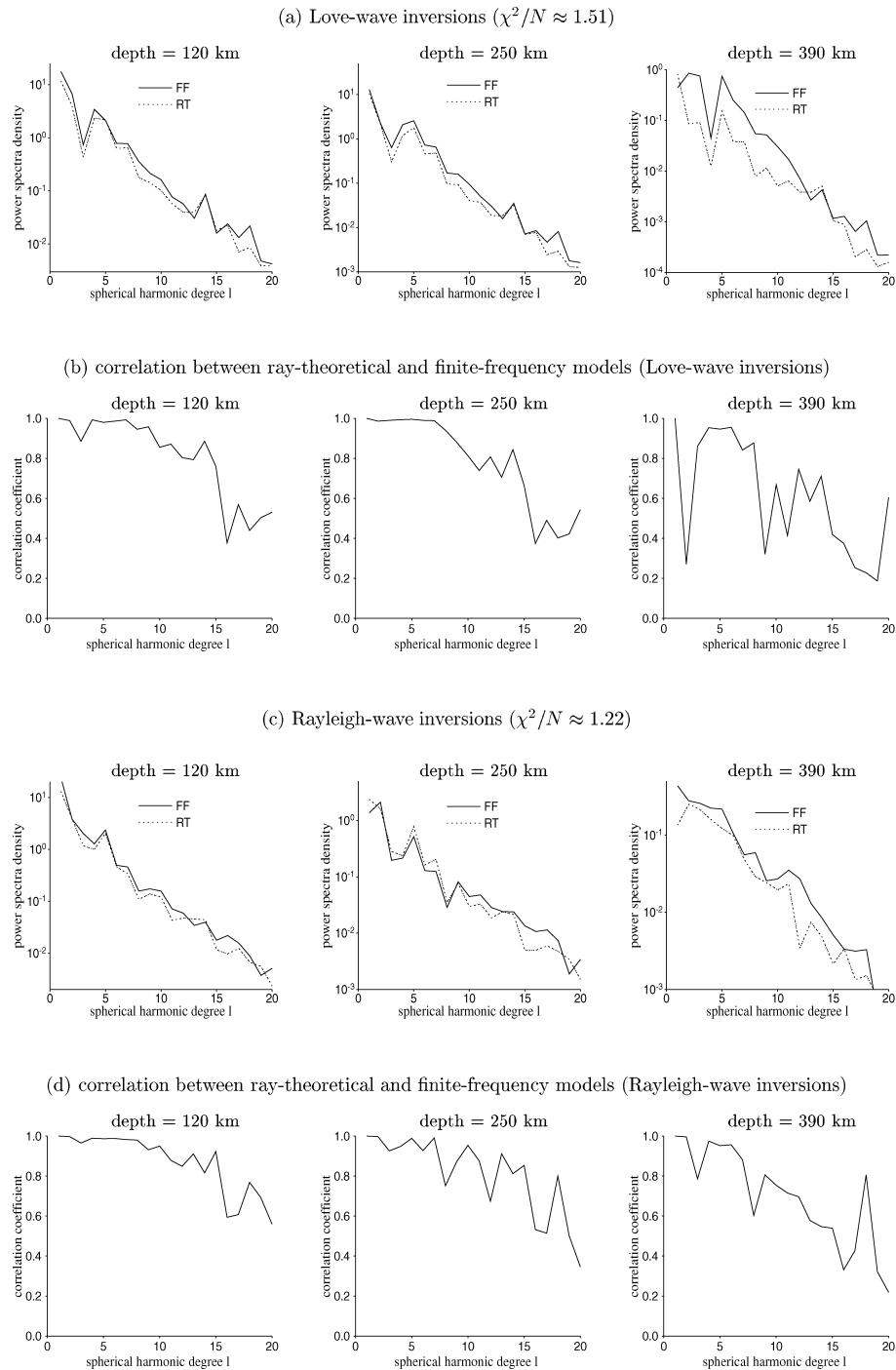
(a) Love-wave inversions ($\chi^2/N \approx 1.51$)

(b) correlation between ray-theoretical and finite-frequency models (Love-wave inversions)

(c) Rayleigh-wave inversions ($\chi^2/N \approx 1.22$)

(d) correlation between ray-theoretical and finite-frequency models (Rayleigh-wave inversions)

Figure 11. Same as Fig. 10, but for *S*-wave velocity models with the same data misfit χ^2/N. Map views of these equal-misfit models are plotted in Figs 6 and 8.

large anomalies with a surface radius ∼2400 km and a peak amplitude of $A_0 = \pm 7$ per cent, and small anomalies with a surface radius ∼600 km and a peak amplitude of $A_0 = \pm 5$ per cent. The misfit versus roughness trade-off curves are plotted in Fig. 12, and the output checkerboards recovered using finite-frequency kernels and ray theory are plotted in Fig. 13. At approximately the same level of model roughness $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$ in Fig. 12, the data misfit χ^2/N of the ray-theoretical models exceeds that of the corresponding finite-frequency models. The plotted finite-frequency models (black stars) fit the synthetics at the expected level, $\chi^2/N = 1$,

whereas the misfits of the plotted ray-theoretical models (white stars) are always higher than this, $\chi^2/N > 1$, due to the theoretical shortcomings of ray theory. The large-scale features of the input anomalies are well resolved in both finite-frequency and ray-theoretical tomography; however, there are significant small-scale artefacts in the ray-theoretical models, particularly at greater depths (250 km). The synthetic checkerboard tests both illustrate the limitations of ray theory in resolving small-scale lateral heterogeneities, and confirm that the spatial resolution of finite-frequency tomography is adequate for the path coverage of the global Love-wave
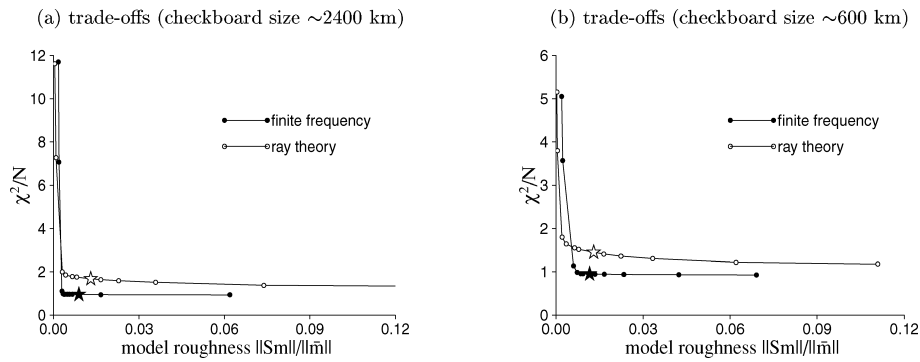
Figure 12. Trade-off curves of the checkboard tests for (a) large and (b) small anomalies. The 'optimal' models, shown as stars in the trade-off curves, are plotted in Fig. 13. Ray-theoretical inversions always show data misfit larger than $\chi^2/N = 1$, due to the theoretical errors inherent in ray theory. The finite-frequency 'no-smoothing' ($\epsilon = 0$) models have roughnesses $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$ equal to 0.06 and 0.07 for (a) large and (b) small anomalies, respectively; the ray-theoretical "no-smoothing" models have roughnesses $||\mathbf{Sm}||/||\bar{\mathbf{m}}||$ equal to 0.28 and 0.11 for large and small anomalies, respectively.
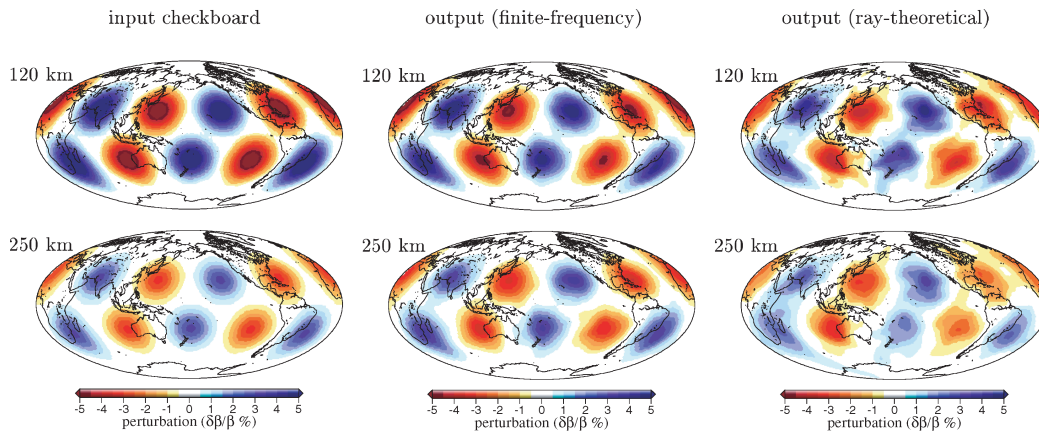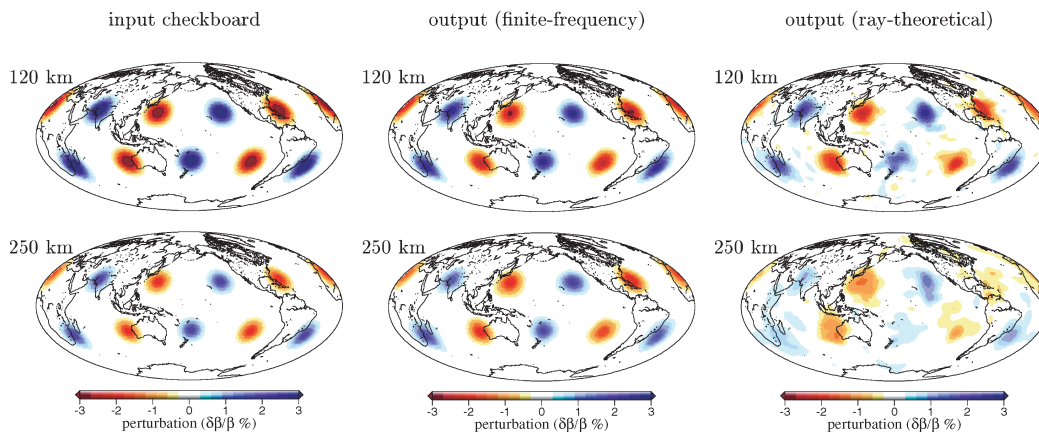

Figure 13. Checkboard resolution tests of finite-frequency and ray-theoretical surface-wave tomography. (a) The checkboard anomalies are smooth Gaussian anomalies, with a peak amplitude of ±7 per cent at 180 km, and a radius ∼2400 km. The amplitude of the anomalies decreases both upward and downward, with an exponential decay rate $f = 0.0125\,\mathrm{km}^{-1}$. (b) The peak amplitude of the Gaussian anomalies is ±5 per cent at 180 km, and the radius is ∼600 km. The synthetic Love-wave phase delays are computed using 3-D sensitivity kernels, with the same path and frequency configurations as in the global data set used in this paper. The output models are all inverted with 50 per cent synthetic Gaussian noise. Ray-theoretical tomography recovers the large-scale structure of the input model, but it poorly resolves small-scale anomalies, especially at deep depth (250 km). The trade-off curves are plotted in Fig. 12.
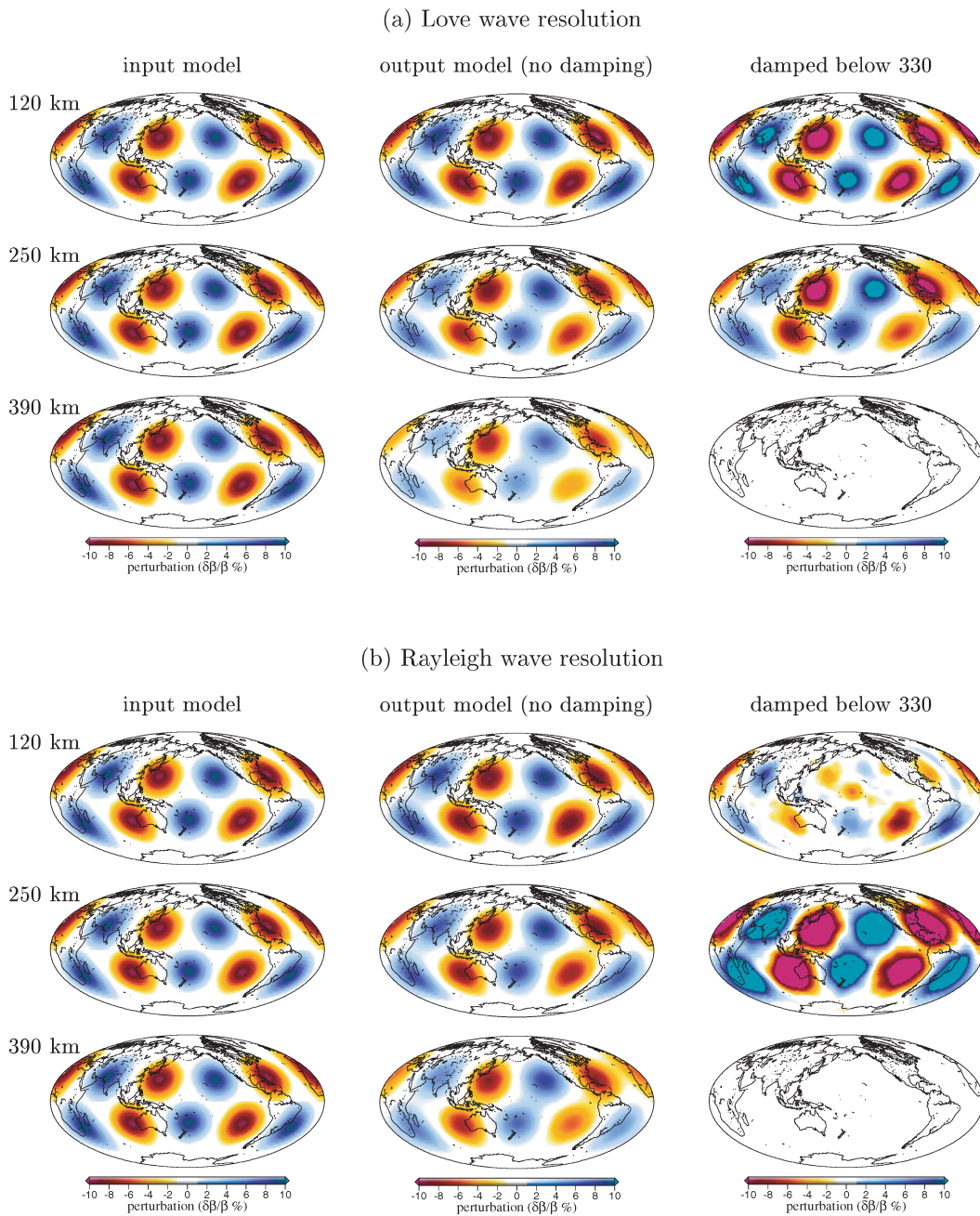
(a) Love wave resolution

(b) Rayleigh wave resolution

Figure 14. Synthetic checkerboard tests to determine the depth resolution of (a) Love waves and (b) Rayleigh waves. The input model consists of 10 Gaussian anomalies, with a peak amplitude of 10 per cent. The radius of the anomaly is about $20°$, the size and amplitude of the anomalies do not vary with depth down to 580 km. Synthetic phase delays are inverted with 50 per cent (rms) Gaussian noise. Models in the right panels are inverted with strong damping below 330 km depth. The rms variation $\langle (\delta\beta/\beta)^2 \rangle^{1/2}$ of the input and output models are plotted versus depth in Fig. 15.

data-set used in this study. It is noteworthy that the same relation between the trade-off curves (i.e. the fact that finite-frequency tomography fits the data better than ray theory) is evident in our inversions of the actual phase-delay measurements (Fig. 4). This similarity in the pattern of the trade-off curves confirms the significance of the finite-frequency effects in our global phase-delay inversions.

4.4 Checkboard tests: depth resolution of finite-frequency tomography

The differences between the ray-theoretical and finite-frequency checkerboards in Fig. 13 become more significant with increasing depth; this is expected, since the lateral extent of finite-frequency, off-path sensitivity is stronger for longer-period surface waves. However, it is possible that these differences in deep structure may be introduced by a lack of depth resolution, as a result of limited data and uneven path coverage. Because noise in the data can be accommodated differently in ray-theoretical and finite-frequency tomography, it is important to determine whether the model resolution is significant at the depths of comparison. We investigate the depth resolution of finite-frequency tomography using the checkboard test illustrated in Fig. 14. The input model consists of five high-velocity and five low-velocity Gaussian cylinders with a peak amplitude of ± 10 per cent; in this case, the radius and

(a) no-damping (Love)

model rms ($\delta\beta/\beta$ %)

(b) damped below 330 (Love)

model rms ($\delta\beta/\beta$ %)

(c) no-damping (Rayleigh)

model rms ($\delta\beta/\beta$ %)

(d) damped below 330 (Rayleigh)

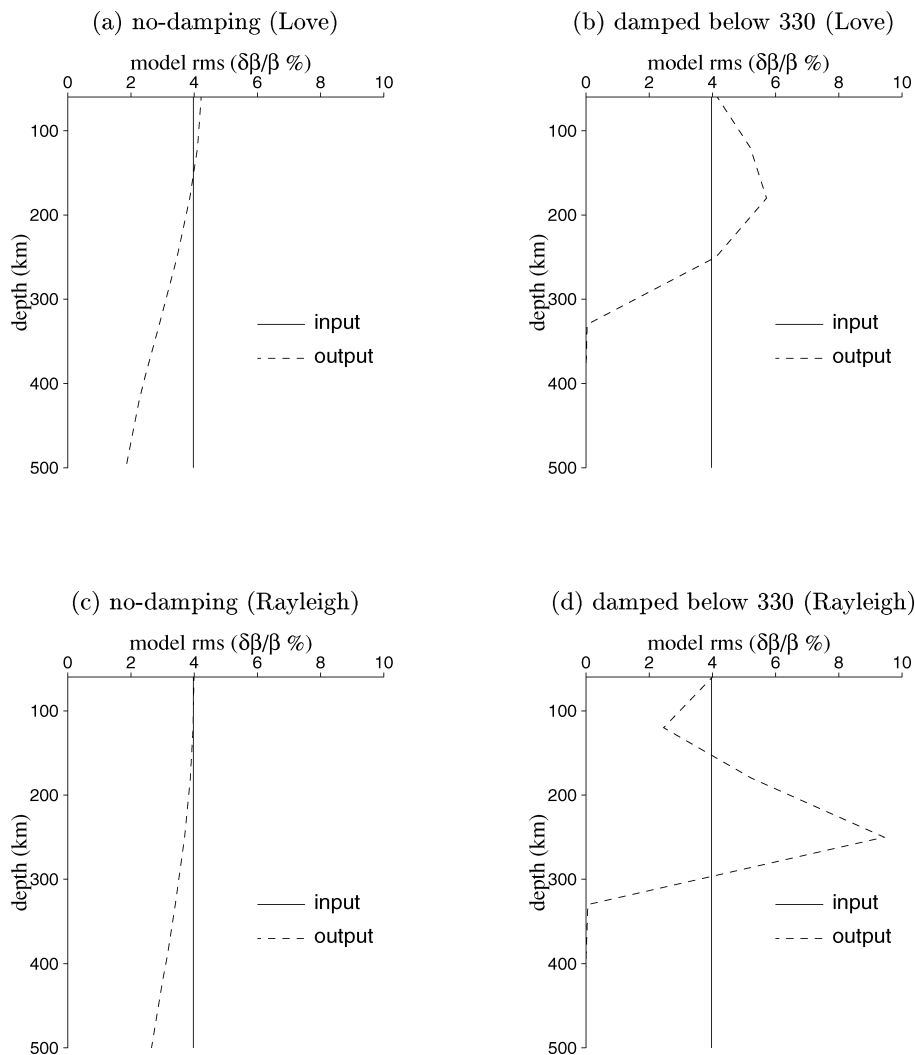model rms ($\delta\beta/\beta$ %)

Figure 15. Depth resolution of the fundamental-mode Love waves and Rayleigh waves used in this study. Plots (a) and (b) show the rms depth profiles $\langle(\delta\beta/\beta)^2\rangle^{1/2}$ of the Gaussian checkerboard models in Fig. 14(a). Plots (c) and (d) show the corresponding depth profiles of the models in Fig. 14(b). In (b) and (d) strong damping is applied to structures deeper than 330 km in the inversion; this results in large errors in the output model at shallow depths. The results of these resolution tests indicate both that the Love-wave and Rayleigh-wave data used in the paper have significant sensitivity to structures below 330 km, and that Rayleigh waves have stronger sensitivity at below 330 km depth than Love waves.

amplitude of the anomalies are constant down to a depth of 580 km. As before, to make the inversion more realistic, we invert synthetic Love-wave and Rayleigh-wave phase delays contaminated by 50 per cent (rms) random noise. The output models exhibit roughly Gaussian cylinders reasonably well-recovered geographically at all depths; as expected, the recovered anomalies are weaker at greater depths because of the reduced data sensitivity (Fig. 14). To test the model resolution in the lowermost upper mantle, we perform additional inversions, in which strong norm damping is applied to anomalies deeper than 330 km. The resulting velocity models show significant artefacts at shallow depths, especially in the Rayleigh-wave inversion. All of the output models in Fig. 14 fit the synthetic Love-wave phase-delay data equally well ($\chi^2/N = 1$). The rms variation of the recovered models is plotted versus depth in Fig. 15. This final series of checkboard tests indicates that model resolution below 330 km is less than that at shallower depths; nevertheless, the sensitivity in the lowermost upper mantle is sufficient that reliable model comparisons may be made even at those depths.

5 FINITE-FREQUENCY EFFECTS IN PHASE-VELOCITY MAPS

In traditional, ray-theoretical, surface-wave tomography, dispersion data are often used to make 2-D maps of the fractional perturbation in phase velocity, $\delta c/c$, at discrete frequencies (e.g. Trampert & Woodhouse 1995; Laske & Masters 1996; Ekström *et al.* 1997). Phase-velocity perturbations at high frequencies are associated with *S*-wave velocity perturbations at shallow depths, whereas perturbations at low frequencies are influenced by velocity anomalies down to greater depth. The intermediate 2-D phase velocity maps are then used as input 'data' to constrain 3-D *S*-wave velocity anomalies, using depth-dependent Fréchet kernels $\partial c(\omega)/\partial\beta(r)$. It has been pointed out that in the presence of lateral heterogeneities, surface-wave phase-velocity measurements can be contaminated by interference from scattered arrivals (Wielandt 1993). Whenever the length scale of lateral heterogeneities is comparable to the wavelength of the surface waves, finite-frequency effects must be taken into account. 2-D sensitivity kernels for the local phase-velocity
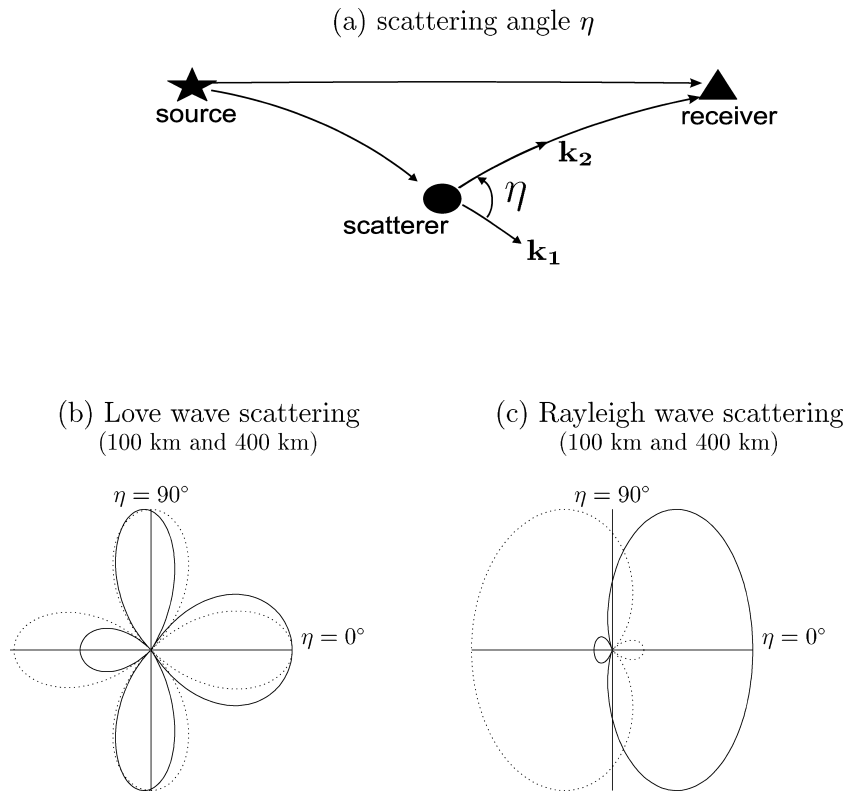
(a) scattering angle η

(b) Love wave scattering
(100 km and 400 km)

(c) Rayleigh wave scattering
(100 km and 400 km)

$\eta = 90°$ $\eta = 0°$ $\eta = 90°$ $\eta = 0°$

Figure 16. Depth and directional dependence of surface-wave scattering. (a) Cartoon illustrating the single scattering of a surface wave. The quantities $\mathbf{k}_1$ and $\mathbf{k}_2$ are the wavevectors of the incoming and scattered waves, respectively; the scattering angle η is measured counter-clockwise from $\mathbf{k}_1$ to $\mathbf{k}_2$. (b) Love wave scattering pattern, that is, the scattering strength $|\Omega^\beta|$ (Zhou *et al.* 2004) as a function of η, for *S*-wave velocity heterogeneities at depths of 400 km (solid line) and 100 km (dashed line). (c) Rayleigh wave scattering pattern for *S*-wave velocity heterogeneities at 400 km (solid line) and 100 km (dashed line). For both Love and Rayleigh waves, the directionality of the scattering depends upon depth.

perturbation, based upon the single-scattering approximation, have been developed (Spetzler *et al.* 2002; Zhou *et al.* 2004), and simplified 2-D sensitivity kernels have been applied in regional and global tomographic studies (Ritzwoller *et al.* 2002; Yoshizawa & Kennett 2004).

2-D phase-velocity sensitivity kernels are based upon a forward-scattering approximation, in which the directional dependence of seismic scattering is ignored; as a result, phase-velocity kernels have large oscillatory sidebands beyond the first Fresnel zone. These sidebands will be partly cancelled out by frequency averaging, if the time windows applied in making the phase-delay measurements are short; however, cancellation due to frequency averaging is very limited for long-period dispersion measurements, because the time windows required to make these measurements need to be long enough to include the dominant wave packets (Zhou *et al.* 2004). The off-ray sidebands are much less prominent in the 3-D sensitivity kernels depicted in Fig. 1, because the directional dependence of surface-wave scattering is properly taken into account. The directional scattering patterns for 5-mHz Love and Rayleigh waves are plotted in Fig. 16, for scatterers at depths of 100 km (dashed line) and 400 km (solid line). The strong depth dependence of the scattering makes surface-wave propagation in the presence of small-scale lateral heterogeneities a fully 3-D process.

Application of the forward-scattering approximation ($\eta = 0$ in Fig. 16) reduces the 3-D integral in eqs (2.4)–(2.5) to a 2-D integral over the local fractional phase-velocity perturbation:

$$\delta\phi(\omega) = \iint_\Omega K_c(\omega, \hat{\mathbf{r}}) \frac{\delta c}{c}(\omega, \hat{\mathbf{r}}) \, d\Omega. \qquad (5.1)$$

The quantity K_c is the phase-velocity sensitivity kernel, and the integration is over all geographical positions $\hat{\mathbf{r}}$ on the surface Ω of the unit sphere. In the remainder of this section, we compare tomographic results obtained using four different phase-velocity sensitivity kernels K_c that are based on different simplifications, and we point out that small-scale structures in tomographic models depend not only upon the geometry of the sensitivity kernels in the first Fresnel zone, but also upon the kernel sidebands beyond the first Fresnel zone. We conclude from this comparison that 2-D phase-velocity sensitivity kernels based upon the forward-scattering approximation or other simplifications do not reliably take into account surface-wave finite-frequency effects.

5.1 Phase-velocity sensitivity kernels

Plots of the simplified phase-velocity sensitivity kernels are shown in Fig. 17. The so-called 'wide-band' and 'narrow-band' kernels are the phase-velocity sensitivity kernels of Zhou *et al.* (2004), with the oscillatory sidebands arbitrarily truncated at the fifth and third zero crossings, respectively. The 'Colorado' kernel and the 'Australia' kernel are simplified 'fat' rays, in which ray-theoretical sensitivities are uniformly spread over a selected width across the great-circle path. The sensitivity in the Colorado kernel is confined to a region $\Delta_1 + \Delta_2 - \Delta \leq 3\lambda/8$, where λ is the wavelength, Δ_1 and Δ_2 are the source-to-scatterer and scatterer-to-receiver great-circle distances, and Δ is the source–receiver epicentral distance. Such a uniform-sensitivity 'fat ray' was applied by the seismic group at the University of Colorado in global group-velocity tomography
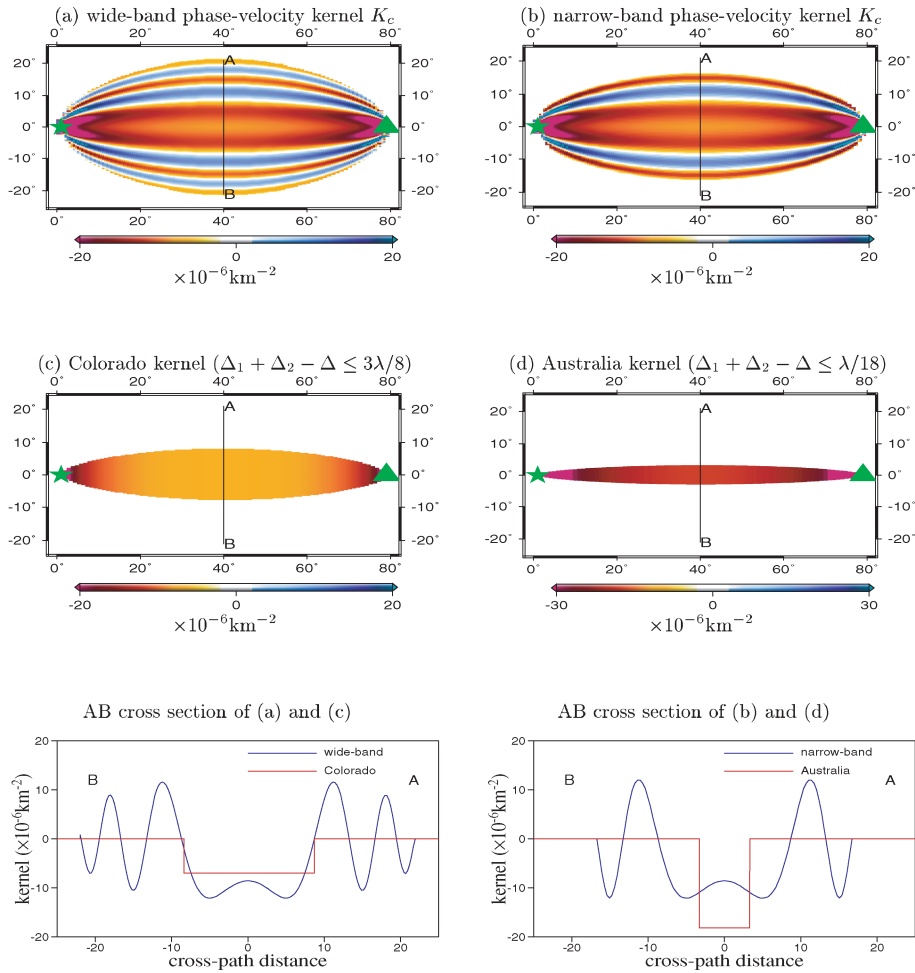
Figure 17. Examples of 2-D phase-velocity sensitivity kernel K_c. Green star and triangle denote the source and receiver, respectively. The wide-band and narrow-band sensitivity kernels are computed using the expressions given by Zhou *et al.* (2004), with the sidebands truncated at the fifth and third zero-crossing, respectively. The Colorado kernel, in which the cross-path sensitivity is constant within the region $\Delta_1 + \Delta_2 - \Delta \leq 3\lambda/8$ is adapted from Ritzwoller *et al.* (2002). The Australia kernel is that proposed to Yoshizawa & Kennett (2002), with a cross-path width $\Delta_1 + \Delta_2 - \Delta \leq \lambda/18$.

(Ritzwoller *et al.* 2002). The cross-path width of the Australia kernel is one-third of the first Fresnel zone, that is, $\Delta_1 + \Delta_2 - \Delta \leq \lambda/18$ (Yoshizawa & Kennett 2002). Such a narrower 'fat ray' has been applied in Australian regional phase-velocity tomography by Yoshizawa & Kennett (2004).

5.2 Phase-velocity kernel comparisons

Love-wave phase-velocity maps at 5 and 15 mHz obtained using traditional ray theory, eq. (2.1), as well as the four different phase-velocity sensitivity kernels in eq. (5.1), are plotted in Fig. 18. All of the recovered maps fit the data equally well ($\chi^2/N \approx 2$). We attribute our inability to fit the data better than this to be due to the severity of the approximations used in the computation of the phase-velocity kernels. Regardless of the approximate inversion procedure, the results agree regarding the distribution and amplitude of large-scale anomalies. However, the smaller-scale heterogeneities vary significantly among the maps obtained using different sensitivity kernels, especially at low frequency (5 mHz). The maps obtained using the Colorado finite-frequency kernel exhibit the strongest amplitudes, whereas those obtained using the Australia kernel most closely resemble the ray-theoretical maps, due to the assumed narrow region of sensitivity. The differences between the maps obtained using the

wide-band and narrow-band kernels are far from negligible, particularly at at 5 mHz, where differences exist in the Pacific and Indian Oceans. This indicates that the sensitivity outside of the first Fresnel zone can have significant effects on the resolution of small-scale anomalies; in fact, these differences can be as large as those between ray-theoretical tomography and finite-frequency tomography using 2-D phase-velocity kernels.

5.3 Phase-velocity checkboard tests

The poor resolution of small-scale anomalies in low-frequency phase-velocity maps can be further demonstrated by synthetic checkboard tests (Fig. 19). The synthetic 3-D model is the same as in the top panel in Fig. 13; the radius of the Gaussian anomalies is ~2400 km. The synthetic phase delays at 5 mHz are generated following the same procedure as in Fig. 13, and the rms of the added random noise is again about 50 per cent of the rms of the structural signal. The output models in Fig. 19, all fit the phase-delay data approximately equally well ($\chi^2/N \approx 2$). The checkboard tests confirm that large-scale structure can be resolved in phase-velocity tomography, whereas small-scale structures are distorted in all phase-velocity maps regardless of the inversion method. The trade-off curves of the checkboard phase-velocity inversions are plotted
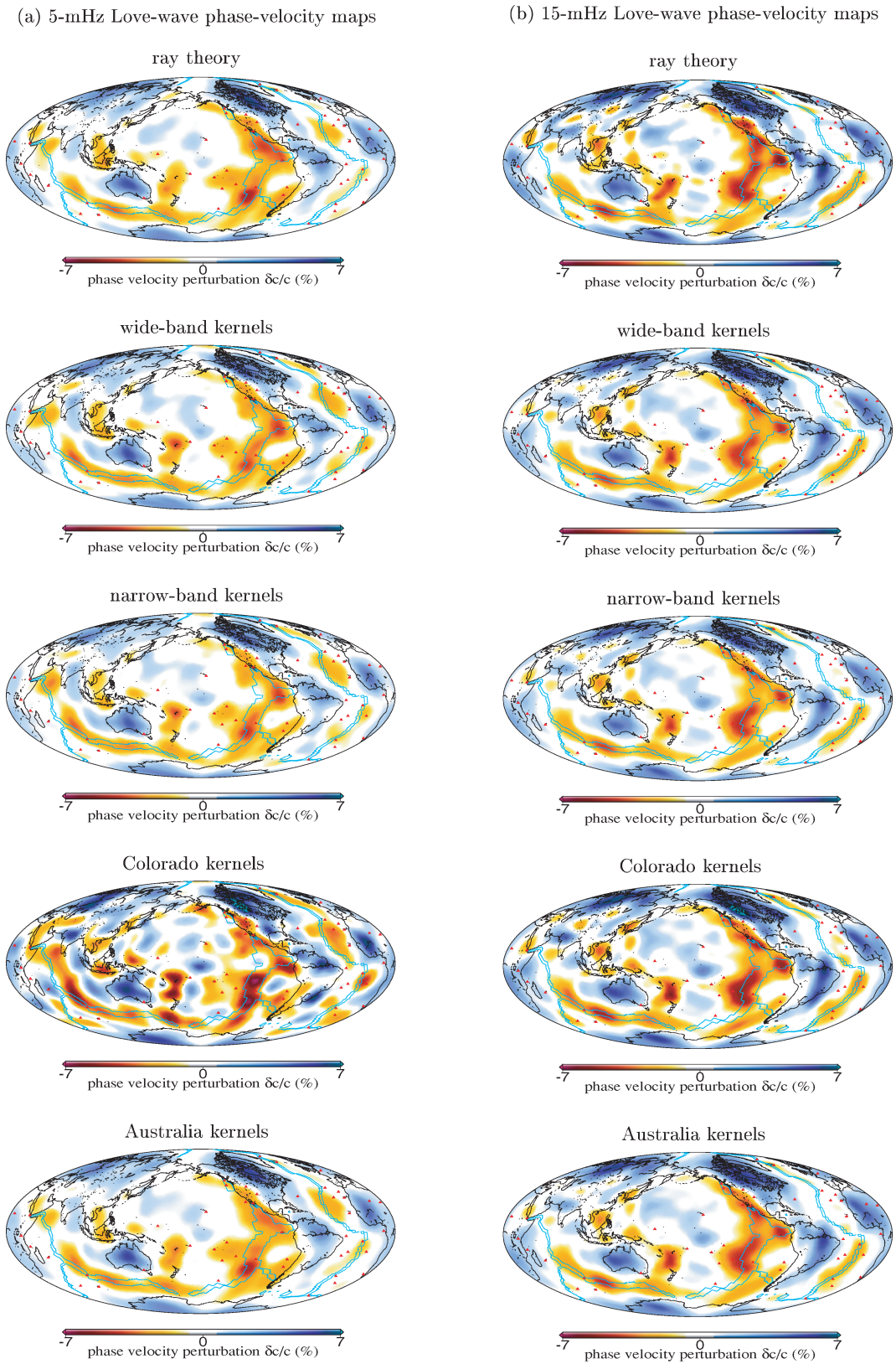
(a) 5-mHz Love-wave phase-velocity maps

(b) 15-mHz Love-wave phase-velocity maps

Figure 18. Love-wave phase-velocity maps at (a) 5 mHz and (b) 15 mHz, obtained using the four phase-velocity sensitivity kernels in Fig. 17, compared with the phase-velocity maps obtained using ray theory (top row). All maps fit the data equally well, $\chi^2/N \approx 2$.
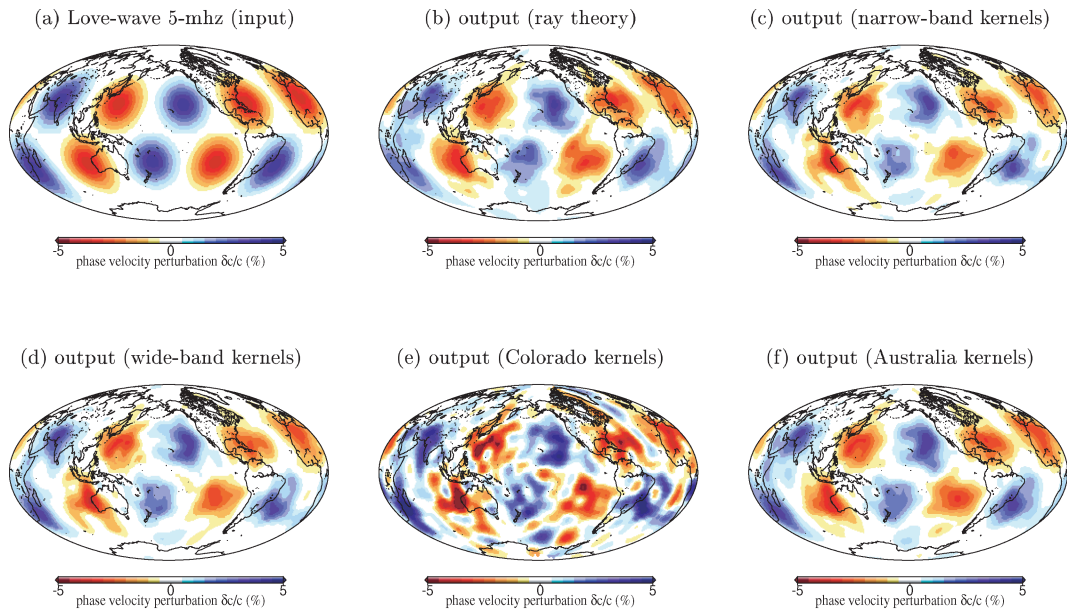
Figure 19. Checkboard resolution tests of the four phase-velocity sensitivity kernels in Fig. 17. (a) Input Love-wave phase-velocity map at 5 mHz, computed from the 3-D checkboard input model in Fig. 13(a). (b) Output model obtained using ray theory. (c)–(f) Output models obtained using the four 2-D, finite-frequency, phase-velocity kernels. All of the output models fit the synthetic data equally well, $\chi^2/N \approx 2$. The misfit versus roughness trade-off curves are plotted in Fig. 20. Inversions using phase-velocity kernels are unable to recover the input model any more faithfully than ray theory.
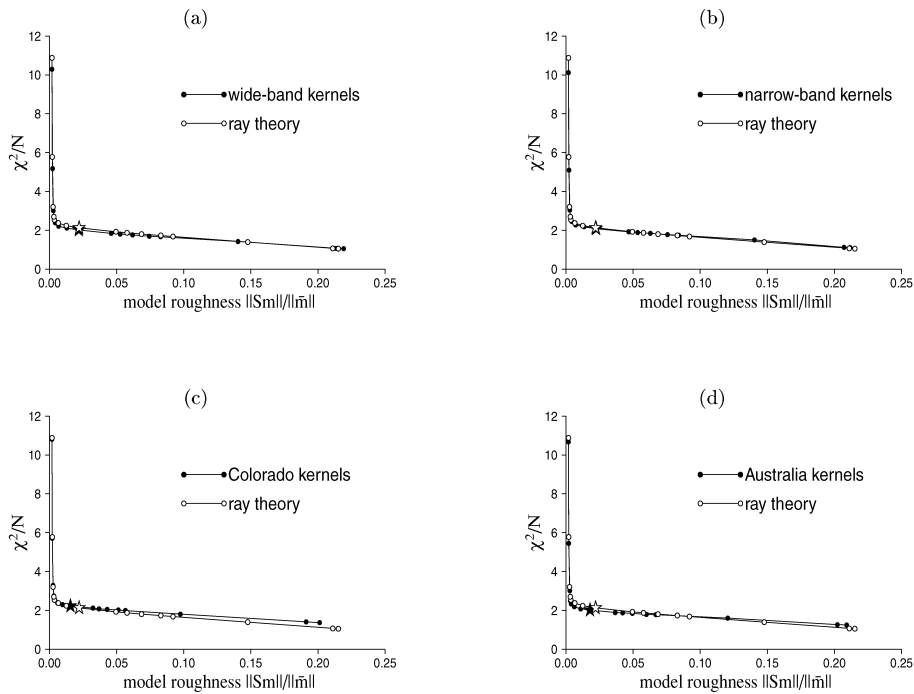

Figure 20. Trade-off curves of 5-mhz Love-wave phase-velocity checkboard tests. Models with data misfit $\chi^2/N \approx 2$ (stars) are plotted in Fig. 19. The trade-off curves are all very close, indicating that inversions using 2-D, finite-frequency, phase-velocity kernels do not fit the synthetic phase-delay data any better than ray theory.

in Fig. 20. The similarity between the ray-theoretical and the various finite-frequency trade-off curves indicates that 2-D phase-velocity sensitivity kernels are unable to fit the synthetic data any better than ray theory. Models that fit the data to within the errors ($\chi^2/N = 1$) are at the far right end of the trade-off curves, where the resulting tomographic models have been adjusted well beyond the bounds of physical reasonability to fit the noise in the data, and the model

roughness increases rapidly in return for very small reductions in data misfit.

5.4 Unreliability of phase-velocity kernels

We conclude from the above comparisons that large-scale features in ray-theoretical phase-velocity maps can be trusted, but that 2-D

phase-velocity sensitivity kernels are unable to recover small-scale anomalies significantly better than ray theory. The directionality of surface-wave scattering must be neglected in order to derive a self-consistent, 2-D, phase-velocity sensitivity kernel K_c, and this is not a very satisfactory approximation, particularly at the longest periods. Accurate imaging of small-scale heterogeneities requires the use of 3-D, finite-frequency sensitivity kernels, with the full directional dependence of the surface-wave scattering properly taken into account.

6 FINITE-FREQUENCY CRUSTAL CORRECTIONS

Lateral variations in the thickness and structure of the earth's crust exert significant effects upon fundamental-mode surface-wave phase delays, even at long period (e.g. Dziewonski 1971). In long-period, mantle surface-wave tomography, crustal contributions to the phase delay are usually removed by applying crustal corrections. It has been noticed that corrections based upon linear perturbation theory may not be adequate in global surface-wave tomography because of the large variation in thickness between oceanic and continental crust (Montagner & Jobert 1988). On the other hand, crustal corrections based upon ray theory may be inadequate due to finite-frequency wave front healing, scattering and diffraction effects. In this section, we formulate finite-frequency sensitivity kernels for variations in the depth of the Moho, and investigate the effects of crustal thickness variations upon long-period surface-wave phase delays, based upon the first-order single-scattering approximation. For brevity, we suppress the dependence upon angular frequency ω.

6.1 Moho boundary sensitivity kernels

Consider a spherically symmetric reference earth model, in which an internal solid–solid boundary (the Moho), denoted by Σ, is subject to a topographic perturbation δd, considered to be positive if the Moho is elevated. The surface-wave Green tensor $\mathbf{G}$ and associated stress tensor $\mathbf{T}$ are perturbed as a result of this topographic displacement:

$$\mathbf{G} \to \mathbf{G} + \delta\mathbf{G}, \quad \mathbf{T} \to \mathbf{T} + \delta\mathbf{T}. \tag{6.1}$$

The perturbations $\delta\mathbf{G}$ and $\delta\mathbf{T}$ satisfy the elastic momentum equation

$$-\rho\omega^2\delta\mathbf{G} + \nabla\cdot\delta\mathbf{T} = \mathbf{0} \quad \text{in } \oplus, \tag{6.2}$$

and the perturbed kinematic and dynamic boundary conditions

$$[\delta\mathbf{G}]_-^+ = -\delta d\,[\partial_r\mathbf{G}]_-^+ \quad \text{on } \Sigma, \tag{6.3}$$

$$[\hat{\mathbf{r}}\cdot\delta\mathbf{T}]_-^+ = -\delta d\,[\hat{\mathbf{r}}\cdot\partial_r\mathbf{T}]_-^+ + \nabla^\Sigma\delta d\,[\mathbf{T}]_-^+ \quad \text{on }\Sigma, \tag{6.4}$$

where $\hat{\mathbf{r}}$ is the outward pointing radial unit vector, $\nabla^\Sigma = \nabla - \hat{\mathbf{r}}\partial_r$ is the tangential gradient operator, and the symbol $[\,\cdot\,]_-^+$ denotes the jump discontinuity in the enclosed quantity in going from above to below the unperturbed spherical boundary. Upon utilizing the seismic representation theorem (Dahlen & Tromp 1998; Aki & Richards 2002), the Green tensor perturbation $\delta\mathbf{G}_{rs}$ can be expressed as

$$\delta\mathbf{G}_{rs} = \iint_\Sigma \left[(\hat{\mathbf{r}}\cdot\mathbf{T}_{xr})^T\cdot\delta\mathbf{G}_{xs} - \mathbf{G}_{xr}^T\cdot(\hat{\mathbf{r}}\cdot\delta\mathbf{T}_{xs})\right]_-^+ d\Sigma. \tag{6.5}$$

We have adopted the same notation as in Zhou *et al.* (2004), in which the roman subscripts s, r and x designate the source, receiver and the scatterer. The quantities $\delta\mathbf{G}_{xs}$ and $\delta\mathbf{T}_{xs}$ represent perturbations in the Green tensor and associated stress tensor at the scatterer $\mathbf{x}$ due to a point-source at the source $\mathbf{s}$. Because the unperturbed traction

is continuous across the solid–solid boundary, $[\hat{\mathbf{r}}\cdot\mathbf{T}_{xs}]_-^+ = 0$, these perturbations satisfy the relation (Dahlen & Tromp 1998, eq. 13.64)

$$[\hat{\mathbf{r}}\cdot\delta\mathbf{T}_{xs}\cdot\mathbf{G}_{rx} + \hat{\mathbf{r}}\cdot\mathbf{T}_{xs}\cdot\delta\mathbf{G}_{rx}]_-^+ = -\delta d\,[\hat{\mathbf{r}}\cdot\partial_r\mathbf{T}_{xs}\cdot\mathbf{G}_{rx}$$
$$+ \hat{\mathbf{r}}\cdot\mathbf{T}_{xs}\cdot\partial_r\mathbf{G}_{rx}]_-^+$$
$$+ \nabla^\Sigma(\delta d)\cdot[\mathbf{T}_{xs}\cdot\mathbf{G}_{rx}]_-^+. \tag{6.6}$$

Upon substituting eq. (6.6) into eq. (6.5) and utilizing the surface version of Gauss's theorem (Dahlen & Tromp 1998, eq. A.79), the Green tensor perturbation $\delta\mathbf{G}_{rs}$ due to the Moho depth perturbation δd becomes

$$\delta\mathbf{G}_{rs} = \iint_\Sigma \delta d\,[-\rho\,\omega^2\mathbf{G}_{rx}\mathbf{G}_{xs} + \varepsilon_{xs}:\mathbf{C}:\varepsilon_{rx}]_-^+ d\Sigma$$
$$- \iint_\Sigma \delta d\,[\hat{\mathbf{r}}\cdot(\mathbf{C}:\varepsilon_{xs})\cdot\partial_r\mathbf{G}_{rx} + \hat{\mathbf{r}}\cdot(\mathbf{C}:\varepsilon_{rx})\cdot\partial_r\mathbf{G}_{xs}]_-^+ d\Sigma. \tag{6.7}$$

The quantity $\mathbf{C}$ is the fourth-order elastic tensor, and $\varepsilon = (1/2)[\nabla\mathbf{G} + (\nabla\mathbf{G})^T]$ is the third-order strain tensor. The 2-D integration is over the unperturbed spherical Moho Σ.

The far-field surface-wave Green tensor $\mathbf{G}_{rs}$ in a spherically symmetric earth model can be written as a summation over all surface-wave modes σ (Snieder & Nolet 1987; Dahlen & Tromp 1998, Section 11.3):

$$\mathbf{G}_{rs} = \sum_\sigma \frac{\mathbf{p}_s^*\mathbf{p}_r\,e^{-i(k\Delta-n\pi/2+\pi/4)}}{\sqrt{8\pi k|\sin\Delta|}}, \tag{6.8}$$

where $\mathbf{p} = \hat{\mathbf{r}}U - i\hat{\mathbf{k}}V + i(\hat{\mathbf{r}}\times\hat{\mathbf{k}})W$ is the surface-wave polarization vector, with $U(r)$, $V(r)$ and $W(r)$ being the surface-wave displacement eigenfunctions. The quantity k is the wavenumber, $\hat{\mathbf{k}}$ is the unit wavevector in the direction of propagation, and n is the polar passage index, or the number of times that the wave train has passed through either the source or its antipode. The quantity Δ is the source–receiver epicentral distance; the asterisk denotes the complex conjugate. Upon substituting the Green tensor, eq. (6.8), into eq. (6.7), the perturbation $\delta\mathbf{G}_{rs}$ becomes a double summation over all surface-wave modes:

$$\delta\mathbf{G}_{rs} = \sum_{\sigma'}\sum_{\sigma''}\iint_\Sigma \delta d$$
$$\times\left(\mathbf{p}_s'^*\mathbf{p}_r''\frac{e^{-i\left[k'\Delta'+k''\Delta''-(n'+n''-1)\pi/2\right]}}{8\pi\sqrt{k'k''|\sin\Delta'||\sin\Delta''|}}\left[\Omega^{(1)}+\Omega^{(2)}\right]_-^+\right)d\Sigma, \tag{6.9}$$

where the single and double primes refer to the surface-wave mode σ' along the source-to-scatterer leg and the surface-wave mode σ'' along the scatterer-to-receiver leg, respectively. The quantities Δ' and Δ'' are the source-to-scatterer and scatterer-to-receiver great-circle distances, respectively. Finally, $k' = k_{\sigma'}$ and $k'' = k_{\sigma''}$ are the wavenumbers of surface-wave modes σ' and σ'', whereas $n' = n_{\sigma'}$ and $n'' = n_{\sigma''}$ are the associated polar-passage indices. The boundary scattering coefficients $\Omega^{(1)}$ and $\Omega^{(2)}$ for an isotropic reference earth model are given in Appendix A. We have adopted the surface-wave normalization of Tromp & Dahlen (1992).

The unperturbed displacement of surface-wave mode σ generated by a moment-tensor source in the symmetrically symmetric reference earth model is (Snieder & Nolet 1987; Dahlen & Tromp 1998, Section 11.4)

$$s = (i\omega)^{-1}\left[\mathbf{M}:\nabla_s\mathbf{G}_{rs}^T\right]\cdot\hat{\boldsymbol{\nu}} = \mathcal{S}\left(\frac{e^{-i(k\Delta-n\pi/2+\pi/4)}}{\sqrt{8\pi k|\sin\Delta|}}\right)\mathcal{R}, \tag{6.10}$$

where $\hat{\boldsymbol{\nu}}$ is the unit vector describing the polarization of the seismometer at the receiver, and $\mathbf{M}$ is the source moment tensor. The

source and receiver terms $\mathcal{S} = (i\omega)^{-1}(\mathbf{M} : \mathbf{E}_s^*)$ and $\mathcal{R} = \mathbf{p}_r \cdot \hat{\boldsymbol{\nu}}$, where $\mathbf{E} = (1/2)[\nabla\mathbf{p} + (\nabla\mathbf{p})^T]$ is the surface-wave strain tensor, are given in Zhou *et al.* (2004).

The perturbation in surface displacement produced by a moment-tensor source is (Snieder & Nolet 1987; Dahlen & Tromp 1998, Section 11.4):

$$\delta s = (i\omega)^{-1}\mathbf{M} : \nabla_s[\delta\mathbf{G}_{rs}^T] \cdot \hat{\boldsymbol{\nu}}. \tag{6.11}$$

Upon substituting eq. (6.9) the displacement of the scattered wave, eq. (6.11), becomes

$$\delta s = \sum_{\sigma'}\sum_{\sigma''}\iint_\Sigma \delta d$$
$$\times \left(\mathcal{S}'\left[\Omega^{(1)} + \Omega^{(2)}\right]_-^+ \mathcal{R}'' \frac{e^{-i[k'\Delta'+k''\Delta''-(n'+n''-1)\pi/2]}}{8\pi\sqrt{k'k''\sin\Delta'||\sin\Delta''|}}\right) d\Sigma, \tag{6.12}$$

where $\mathcal{S}' = (i\omega)^{-1}\mathbf{M} : \mathbf{E}_s'^*$ and $\mathcal{R}'' = \mathbf{p}_r'' \cdot \hat{\boldsymbol{\nu}}$ are the source and receiver term of the scattered wave, respectively.

Correct to first order in the small perturbations, phase-delay perturbations $\delta\phi$ can be related to displacement perturbations δs by $\delta\phi = -\text{Im}[\delta s/s]$. Upon utilizing eqs (6.10) and (6.12), the phase-delay perturbation $\delta\phi$ can be written as 2-D integral over the boundary Σ:

$$\delta\phi = \iint_\Sigma K_d\, \delta d\, d\Sigma, \tag{6.13}$$

where the boundary sensitivity kernel is given by

$$K_d =$$
$$-\text{Im}\left(\sum_{\sigma'}\sum_{\sigma''}\frac{\mathcal{S}'}{\mathcal{S}}\frac{\mathcal{R}''}{\mathcal{R}}\frac{e^{-i[k'\Delta'+k''\Delta''-k\Delta-(n'+n''-n)\pi/2+\pi/4]}}{\sqrt{8\pi(k'k''/k)(|\sin\Delta'||\sin\Delta''|/|\sin\Delta|)}}\right.$$
$$\left.\times\left[\Omega^{(1)} + \Omega^{(2)}\right]_-^+\right). \tag{6.14}$$

In the rest of this section, we ignore the effects of mode coupling, and consider only fundamental-mode surface waves, that is, $\sigma' = \sigma'' = \sigma = 0$. The boundary sensitivity kernel in the absence of mode coupling can be written as:

$$K_d = -\text{Im}\left(\frac{\mathcal{S}'}{\mathcal{S}}\frac{\mathcal{R}''}{\mathcal{R}}\frac{e^{-i[k\Delta'+k\Delta''-k\Delta-(n'+n''-n)\pi/2+\pi/4]}}{\sqrt{8\pi k(|\sin\Delta'||\sin\Delta''|/|\sin\Delta|)}}\right.$$
$$\left.\times\left[(\Omega^{(1)} + \Omega^{(2)})\right]_-^+\right). \tag{6.15}$$

It can be shown, upon applying the forward-scattering and paraxial approximations as in Zhou *et al.* (2004, Sections 6 and 7), that if the cross-path length scale of the boundary-depth variations is much larger than the seismic wavelength, eqs (6.13) and (6.15) reduce to ray theory,

$$\delta\phi = \int_0^\Delta \delta d\left[\left(-\frac{k}{c}\right)\left(\frac{\partial c}{\partial d}\right)\right] dl, \tag{6.16}$$

where $(\partial c/\partial d)$ is the boundary Fréchet derivative for spherically symmetric (1-D) boundary depth perturbations (Dahlen & Tromp 1998, Sections 9.3 and 11.8).

Examples of the Moho boundary sensitivity kernel K_d for 10-mHz fundamental-mode minor-arc Love waves are plotted in Fig. 21. The geometry of the boundary kernels is similar to 3-D velocity kernels at a specific depth, that is, the width of the kernel becomes narrower near the source (or receiver) where the sensitivities are relatively strong. In the example shown in Fig. 21(a), we may expect ray theory to be a good approximation over the Pacific Ocean, where variations in crustal thickness are smooth and the length scale of the variations is larger than the wavelength of the 10-mHz Love wave. However, ray theory may not be adequate whenever the source–receiver path is along a continent–ocean boundary, as in Fig. 21(b), so that the crustal thickness varies significantly across the great-circle ray path.

6.2 Crustal corrections

The magnitude of the finite-frequency effects due to lateral variations in crustal thickness is shown in Fig. 22, in which ray-theoretical crustal corrections are plotted versus corrections made using finite-frequency boundary sensitivity kernels, for the suite of path configurations used in this study. To compute both corrections, we use the global crustal model, CRUST2.0 (Laske *et al.* 2001), which specifies a seven-layer crust in each 2° by 2° global cell. For simplicity, only variations in the crustal thickness (i.e. Moho depth) are considered in this comparison. Ray-theoretical crustal corrections are path integrals of the local phase-velocity perturbations, computed using
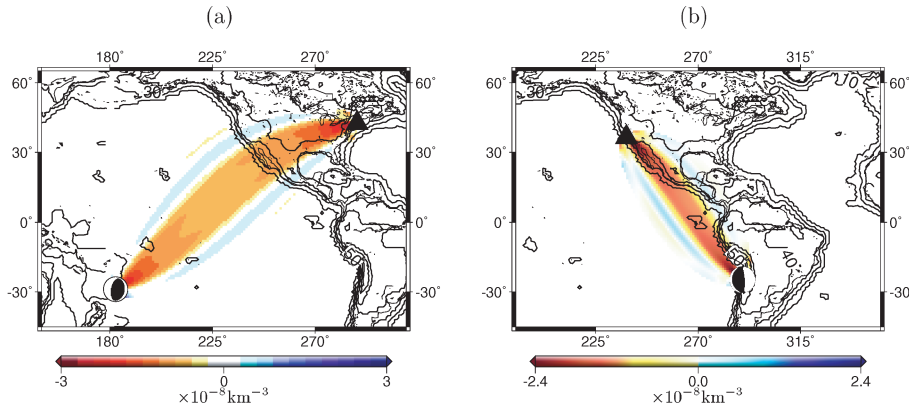

(a)   (b)

Figure 21. 2-D Fréchet kernels K_d, expressing the sensitivity to Moho depth variations, for 10-mHz fundamental-mode Love-wave phase delays. The Moho depth in model CRUST2.0 (Laske *et al.* 2001) is contoured with a 10 km interval. In example (a) the cross-path crustal thickness variations are small and ray theory may be a good approximation. In example (b), the ray path runs along a continent–ocean boundary, and the variations in crustal thickness are significant within the sensitive region. The triangles indicate seismic stations; earthquake focal mechanisms are indicated by the beachballs.
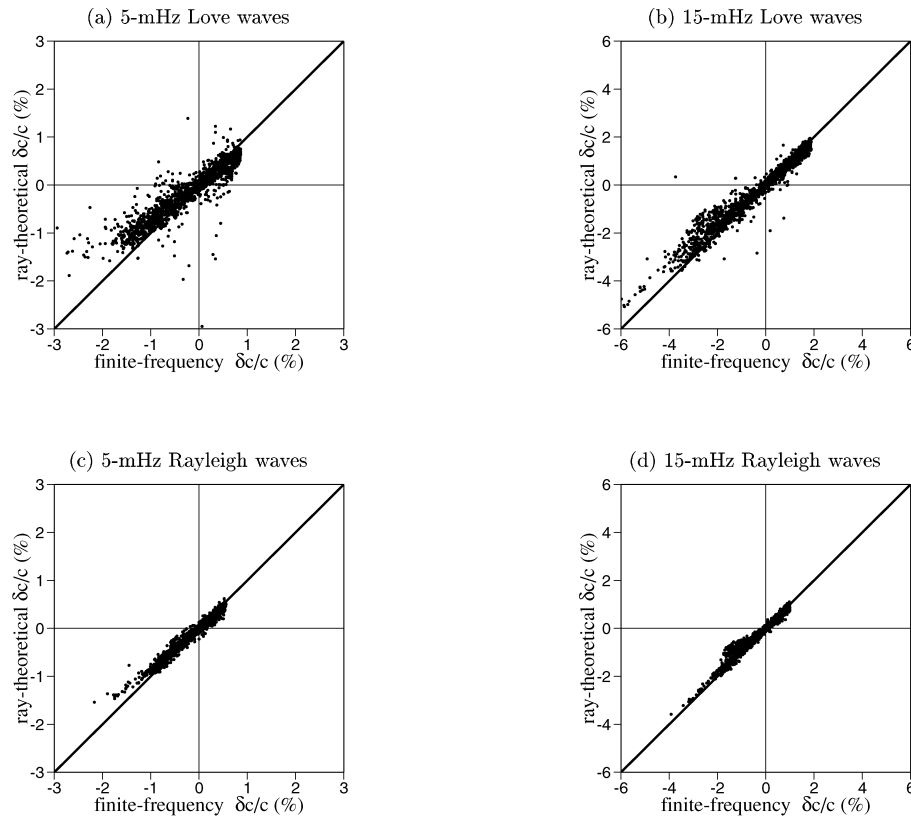
Figure 22. Finite-frequency effects of Moho depth perturbations upon Love and Rayleigh waves. Ray-theoretical crustal corrections are path integrals of the local crustal phase delay computed using linear perturbation theory; finite-frequency predictions are computed using the 2-D boundary sensitivity kernels derived in this paper. The scatter plots include corrections for all Love wave trains G1, G2, G3 and G4, and all Rayleigh wave trains R1, R2, R3 and R4 in our global data set. The reference earth model has a globally averaged seven-layer crust on top of the 1066A mantle. For many paths, finite-frequency effects are not negligible, more so for Love waves and especially at low frequencies (5 mHz).

eq. (6.16). The finite-frequency crustal corrections are computed using eqs (6.13) and (6.15), with mode-coupling effects ignored. The scatterplots in Fig. 22 show that finite-frequency effects in crustal corrections are significant, and that ray-theoretical corrections may not be sufficient in making crustal corrections. In general, finite-frequency effects are more significant for long-period waves than for short-period waves, and they are more significant for Love waves than for Rayleigh waves.

The finite-frequency sensitivity kernels derived in this paper are based upon first-order perturbation theory, which is valid whenever the perturbations with respect to the reference earth model are small. To determine whether the boundary kernels are reliable for making crustal corrections, we examine the validity of linear perturbation theory in global crustal corrections, by comparing exact crustal phase delays with predictions based upon linear perturbation theory. The exact phase delays in each 2° by 2° global cell are computed by solving the radial equations for spherically symmetric earth models with and without the perturbation in the thickness of crust. The corrections based on linear perturbation theory are computed using the boundary depth Fréchet derivative for a spherically symmetric perturbation, eq. (6.16). The crustal thickness in the reference earth model is 19.2 km, the globally averaged crustal thickness in CRUST2.0. Maps of crustal corrections are plotted for Love waves and Rayleigh waves in Figs 23 and 24, respectively. The comparisons show that linear perturbation theory overpredicts crustal corrections by about 20 per cent in regions of thick crust, such as Tibet, the Andes and old continental cratons. The discrepancy between the maps is more significant at high frequency

(15 mHz). The magnitude of the errors introduced by first-order perturbation theory (Figs 23 and 24) are comparable to the magnitude of uncorrected finite-frequency effects in ray-theoretical crustal corrections (Fig. 22).

Linear perturbation theory has been successfully applied in regional crustal studies, where variations in crustal thickness are relatively small (e.g. Das & Nolet 1998). However, in global applications, the crustal thickness varies from 7 km beneath the oceanic basins to 75 km beneath elevated continent regions, and first-order perturbation theory is inadequate in this case. Therefore, the finite-frequency boundary sensitivity kernels derived in this paper do not reliably account for finite-frequency effects in global studies. The crustal corrections applied in Section 3.2 are ray-path integrals of the exact local crustal phase delays; we have opted to use ray theory rather than first-order, finite-frequency perturbation theory as a 'lesser' of two evils' expedient. In any surface-wave tomographic study that employs traditional ray-based crustal corrections, there may be biases in upper-mantle *S*-wave velocity maps, due to neglected finite-frequency crustal thickness effects. Incorporation of finite-frequency effects beyond the first-order Born approximation requires a more comprehensive numerical treatment, such as that of Komatitsch & Tromp (1999).

7 CONCLUSIONS

We consider three different aspects of the problem of inverting long-period, fundamental-mode, surface-wave dispersion measurements. First, we show that finite-frequency tomography with 3-D Born sensitivity kernels offers a significant improvement over
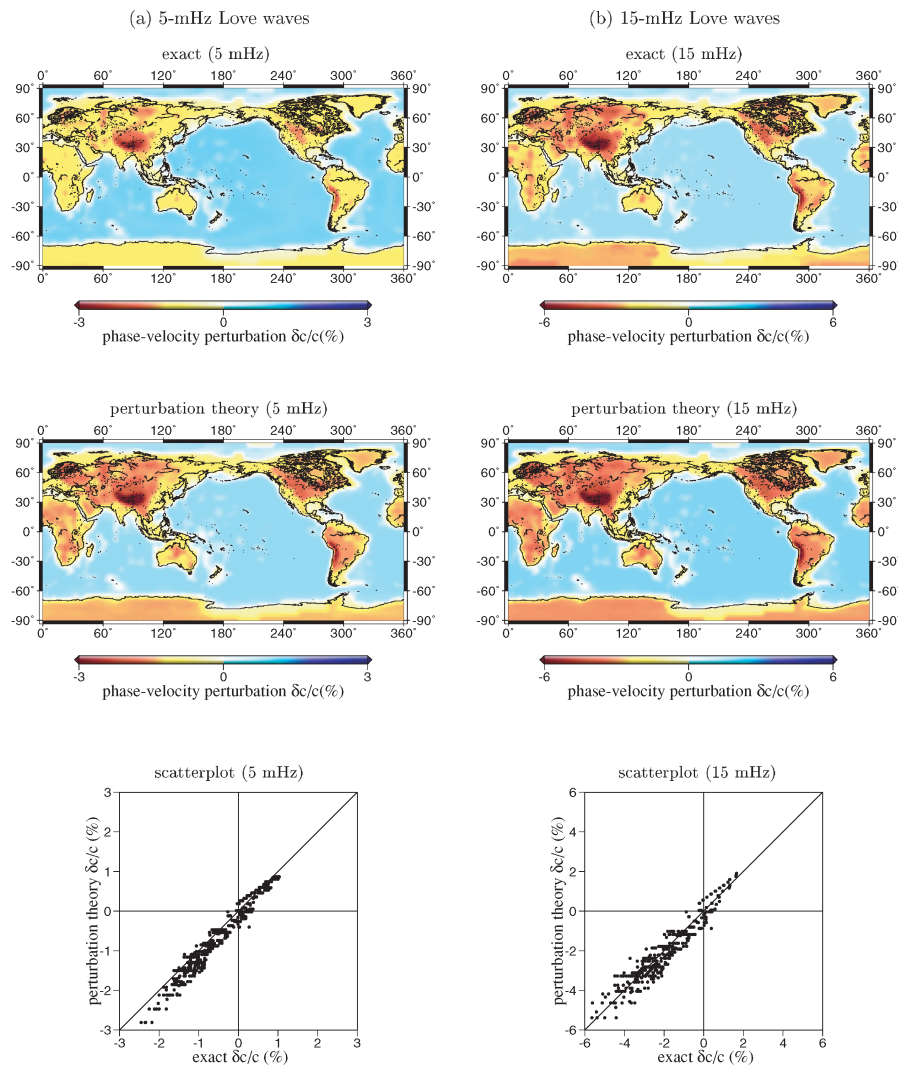
Figure 23. Comparisons between exact local crustal corrections and corrections made using first-order perturbation theory. The exact phase delays in each 2° by 2° global cell are computed by solving the radial equations for spherically symmetric earth models with and without the local crust. The corrections based on linear perturbation theory are computed using the boundary depth Fréchet kernel $\partial c/\partial d$ for a spherically symmetric perturbation. The maps are computed for a global crustal model, CRUST2.0 (Laske *et al.* 2001). Corrections based upon linear perturbation theory exhibit large errors, as indicated both by the differences between the maps, particularly in regions of thick crust, and the first-order versus exact scatter plots.

traditional ray-theoretical, surface-wave tomography, enabling a better fit to a global data set of phase-delay measurements at the same level of model roughness. Finite-frequency and ray-theoretical tomographic images are in good agreement at the largest lateral length scales; however, finite-frequency tomography enhances the amplitude of upper-mantle heterogeneities, because wave front healing effects are properly taken into account. Small-scale anomalies are better resolved using 3-D sensitivity kernels, and this improvement in resolution becomes more significant with increasing depth. Second, we show that 2-D, phase-velocity sensitivity kernels that are based upon a forward-scattering approximation or other simplifications do not reliably account for the finite-frequency effects of surface waves. As a result, the small-scale anomalies in phase-velocity maps obtained using 2-D kernels may not be any more reliable than ray theory. Third and finally, we develop finite-frequency sensitivity kernels for Moho depth perturbations, based upon the single-scattering approximation. Finite-frequency, off-path effects are not negligible in a global, long-period dispersion data set, especially for surface waves that travel along continent–ocean boundaries. How-

ever, global variations in crustal thickness are so large that first-order perturbation theory, which is the basis for computing the Moho depth sensitivity kernels, breaks down. Although it may be feasible to use first-order, finite-frequency sensitivity kernels to account for crustal structure effects in a well-controlled regional dispersion study, crustal corrections based upon Born sensitivity kernels are not reliable for a global, long-period data set. A simple, accurate method for making global surface-wave crustal corrections remains to be found.

ACKNOWLEDGMENTS

(a) 5-mHz Rayleigh waves
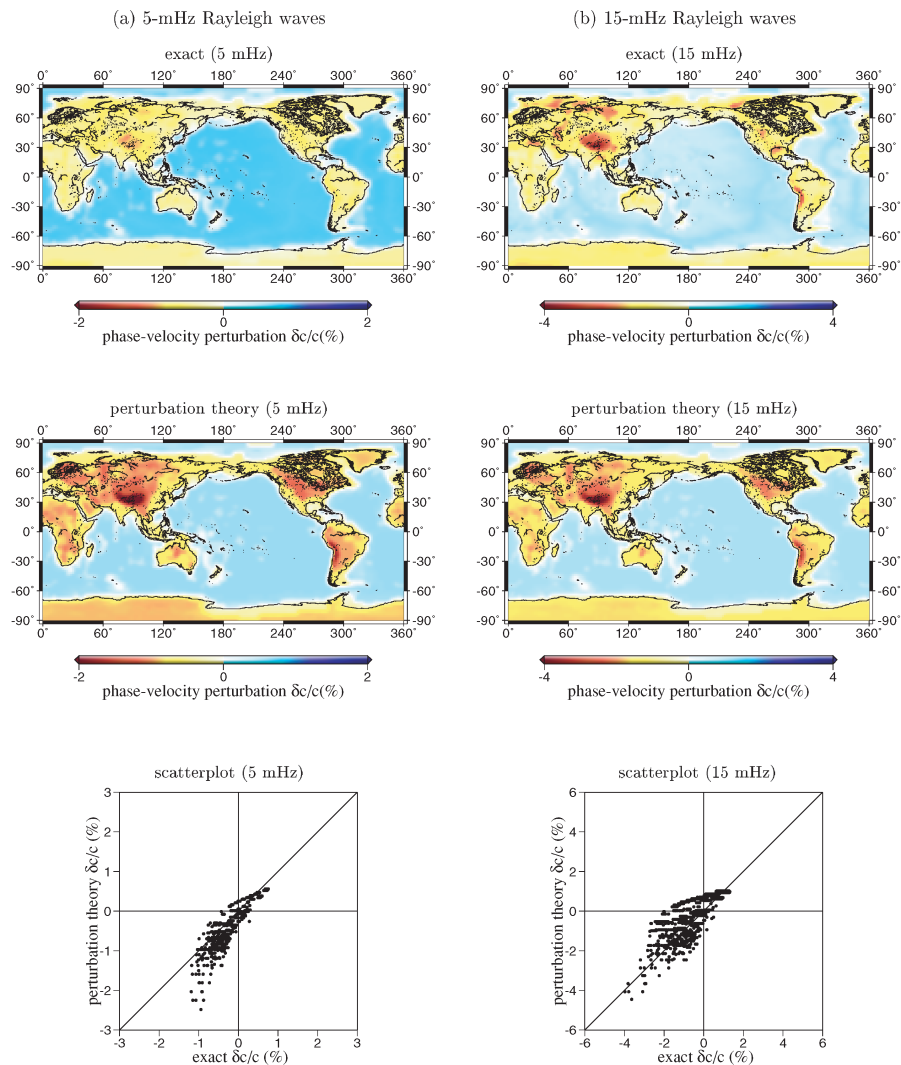(b) 15-mHz Rayleigh waves

Figure 24. Same as Fig. 23, except for Rayleigh waves.

REFERENCES

Aki, K. & Kaminuma, K., 1963. Phase velocity of Love waves in Japan (part 1): Love waves from the Aleutian shock of March 1957, *Bull. Earthq. Res. Inst.,* **41,** 243–259.

Aki, K. & Richards, P.G., 2002. *Quantitative Seismology, Second Edition*, University Science Books, Sausalito, California, USA.

Anderson, D.L., 1961. Elastic wave propagation in layered anisotropic media, *J. geophys. Res.,* **66,** 2953–2963.

Baumgardner, J.R. & Frederickson, P.O., 1985. Icosahedral discretization of the two-sphere, *SIAM Journal of Numerical Analysis,* **22,** 1107–1115.

Das, T. & Nolet, G., 1998. Crustal thickness map of the western United States by partitioned waveform inversion *J. geophys. Res.,* **103,** 30 021–30 038.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton University Press, Princeton, New Jersey, USA.

Dziewonski, A.M., 1971. Upper mantle models from 'pure path' dispersion data, *J. geophys. Res.,* **76,** 2587–2601.

Ekström, G., Tromp, J. & Larson, E.W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.,* **102,** 8137–8157.

Gilbert, F. & Dziewonski, A.M., 1975. An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, *Phil. Trans. R. Soc. Lond.*, A, **278,** 187–269.

Hung, S.-H., Shen, Y. & Chiao, L.-Y., 2004. Imaging seismic velocity structure beneath the Iceland hot spot: a finite frequency approach, *J. geophys. Res.,* **109,** doi: 10.1029/2003JB002889.

Friederich, W., 1999. Propagation of seismic shear and surface waves in a laterally heterogeneous mantle by multiple forward scattering, *Geophys. J. Int.,* **136,** 180–204.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation *Geophys. J. Int.,* **139,** 806–822.

Laske, G. & Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data, *J. geophys. Res.,* **101,** 16 059–16 075.

Laske, G., Masters, G. & Reif, C., 2001. CRUST2.0—a new global crustal model at 2 × 2 degrees, http://mahi.ucsd.edu/Gabi/crust2.html.

Li, X.-D. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.,* **121,** 695–709.

Meier, T., Lebedev, S., Nolet, G. & Dahlen, F.A., 1997. Diffraction tomography using multimode surface waves, *J. geophys. Res.,* **102,** 8255–8267.

Montagner, J.P. & Anderson, D.L., 1989. Petrological constraints on seismic anisotropy, *Phys. Earth planet. Inter.,* **54,** 82–105.

Montagner, J.P. & Jobert, N., 1988. Vectorial tomography—II. Application to the Indian Ocean, *Geophys. J.,* **94,** 309–344.

Montelli, R., Nolet, G., Masters, G., Dahlen, F.A. & Hung, S.-H., 2004. Global P and PP traveltime tomography: rays versus waves, *Geophys. J. Int.,* **158,** 637–654.

Paige, C.C. & Saunders, M.A., 1982. LSQR: An algorithm for sparse linear equations and sparse lease squares, *ACM TOMS,* **8,** 43–71.

Ritzwoller, M.H., Shapiro, N.M., Barmin, M.P. & Levshin, A.L., 2002. Global surface wave diffraction tomography, *J. geophys. Res.,* **107,** B12, 10.1029/2002JB001777.

Spetzler, J., Trampert, J. & Snieder, R., 2002. The effects of scattering in surface wave tomography, *Geophys. J. Int.,* **149,** 755–767.

Snieder, R. & Nolet, G., 1987. Linearized scattering of surface waves on a spherical Earth, *J. geophys.* **61,** 55–63.

Trampert, J. & Woodhouse, J.H., 1995. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 s, *Geophys. J. Int.,* **122,** 675–690.

Tromp, J. & Dahlen, F.A., 1992. Variational principles for surface wave propagation on a laterally heterogeneous Earth—II. Frequency-domain JWKB theory, *Geophys. J. Int.,* **109,** 599–619.

Wessel, P. & Smith, W.H.F., 1995. New version of the Generic Mapping Tools released, *EOS, Trans. Am. geophys. Un.,* **76,** 329.

Wielandt, E., 1993. Propagation and structural interpretation of non-plane waves, *Geophys. J. Int.,* **113,** 45–53.

Yomogida, K. & Aki, K., 1987. Amplitude and phase data inversions for phase velocity anomalies in the Pacific Ocean basin, *Geophys. J. R. astr. Soc.,* **88,** 161–204.

Yoshizawa, K. & Kennett, B.L.N., 2002. Determination of the influence zone for surface wave paths, *Geophys. J. Int.,* **149,** 440–453.

Yoshizawa, K. & Kennett, B.L.N., 2004. Multimode surface wave tomography for the Australian region using a three-stage approach incorporating finite frequency effects, *J. geophys. Res.,* **109,** doi: 10.1029/2002JB002254.

Zhou, Y., Dahlen, F.A. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.,* **158,** 142–168.

Zhou, Y., Nolet, G., Dahlen, F.A. & Laske, G., 2005. Global upper-mantle structure from finite-frequency surface-wave tomography, *J. geophys. Res.*, in press.

APPENDIX A: BOUNDARY SCATTERING COEFFICIENTS

In this appendix, we derive explicit expressions for the boundary scattering coefficients $\Omega^{(1)}$ and $\Omega^{(2)}$ in eqs (6.9), (6.12) and (6.14)–(6.15). The notation used in this appendix is the same as in the appendix of Zhou *et al.* (2004). The first scattering interaction coefficient $\Omega^{(1)}$ can be written as:

$$\Omega^{(1)} = \rho\,\omega^2 \mathbf{p}''^* \cdot \mathbf{p}' - \lambda\,(\mathrm{tr}\,\mathbf{E}''^*)(\mathrm{tr}\,\mathbf{E}') - \mu\,(\mathbf{E}''^* : \mathbf{E}'), \tag{A1}$$

where $\mathbf{p}' = \hat{\mathbf{r}} U' - i\,\hat{\mathbf{k}}' V' + i\,(\hat{\mathbf{r}} \times \hat{\mathbf{k}}') W'$ is the polarization vector of surface-wave mode σ' along the source-to-scatterer leg, and $\mathbf{p}'' = \hat{\mathbf{r}} U'' - i\,\hat{\mathbf{k}}'' V'' + i\,(\hat{\mathbf{r}} \times \hat{\mathbf{k}}'') W''$ is the polarization vector of surface-wave mode σ'' along the scatterer-to-receiver leg. The quantities $\mathbf{E}' = (1/2)[\nabla \mathbf{p}' + (\nabla \mathbf{p}')^{\mathrm{T}}]$ and $\mathbf{E}'' = (1/2)[\nabla \mathbf{p}'' + (\nabla \mathbf{p}'')^{\mathrm{T}}]$ are the associated surface-wave strain tensors, and an asterisk denotes the complex conjugate. All quantities in eq. (A1) are evaluated at the scatterer. The detailed expressions for the scattering coefficient $\Omega^{(1)}$, dependent upon the scattering interaction, are:

Love-to-Love scattering,

$$\begin{aligned}
\Omega^{(1)}_{L'L''} = {} & -\rho\,\omega^2 W'' W' \cos\eta \\
& + \mu\,(\dot{W}'' - r^{-1} W'')(\dot{W}' - r^{-1} W') \cos\eta \\
& + \mu\,k'' k' r^{-2} W'' W' \cos 2\eta;
\end{aligned}$$

Rayleigh-to-Rayleigh scattering,

$$\begin{aligned}
\Omega^{(1)}_{R'R''} = {} & -\rho\,(U'' U' + V'' V' \cos\eta) \\
& + \lambda\,(\dot{U}' + 2r^{-1} U' - k' r^{-1} V')(\dot{U}'' + 2r^{-1} U'' - k' r^{-1} V'') \\
& + \mu\,[2\dot{U}'' \dot{U}' + r^{-2}(2U'' - k'' V'')(2U' - k' V')] \\
& + \mu\,(\dot{V}'' - r^{-1} V'' + k'' r^{-1} U'')(\dot{V}' - r^{-1} V' + k' r^{-1} U') \cos\eta \\
& + \mu\,k'' k' r^{-2} V'' V' \cos 2\eta;
\end{aligned}$$

Love-to-Rayleigh scattering,

$$\begin{aligned}
\Omega^{(1)}_{L'R''} = {} & \rho\,\omega^2 V'' W' \sin\eta \\
& - \mu\,(\dot{V}'' - r^{-1} V'' + k'' r^{-1} U'')(\dot{W}' - r^{-1} W') \sin\eta \\
& - \mu\,k'' k' r^{-2} V'' W' \sin 2\eta;
\end{aligned}$$

Rayleigh-to-Love scattering,

$$\begin{aligned}
\Omega^{(1)}_{R'L''} = {} & -\rho\,\omega^2 W'' V' \sin\eta \\
& + \mu\,(\dot{W}'' - r^{-1} W'')(\dot{V}' - r^{-1} V' + k' r^{-1} U') \sin\eta \\
& + \mu\,k'' k' r^{-2} W'' V' \sin 2\eta.
\end{aligned}$$

The second scattering interaction coefficient $\Omega^{(2)}$ can be written as:

$$_{\sigma'}\Omega^{(2)}_{\sigma''} = [\hat{\mathbf{r}} \cdot \nabla_{\mathrm{x}} \mathbf{p}''^*] \cdot \{\hat{\mathbf{r}} \cdot [\lambda(\mathrm{tr}\,\mathbf{E}')\,\mathbf{I} + 2\mu \mathbf{E}']\} + \{\hat{\mathbf{r}} \cdot [\lambda(\mathrm{tr}\,\mathbf{E}''^*)\mathbf{I} + 2\mu \mathbf{E}''^*]\} \cdot [\hat{\mathbf{r}} \cdot \nabla_{\mathrm{x}} \mathbf{p}']. \tag{A2}$$

In this case the detailed expressions are:

Love-to-Love scattering,

$$\begin{aligned}
\Omega^{(2)}_{L'L''} = {} & -\mu\,\dot{W}''\,[\dot{W}' - r^{-1} W'] \cos\eta \\
& - \mu\,\dot{W}'\,[\dot{W}'' - r^{-1} W''] \cos\eta;
\end{aligned}$$

Rayleigh-to-Rayleigh scattering,

$$\Omega^{(2)}_{R'R''} = -\lambda\,\dot{U}''\,[\dot{U}' + 2\,r^{-1}U' - k'r^{-1}V']$$
$$-\lambda\,\dot{U}'\,[\dot{U}'' + 2\,r^{-1}U'' - k''r^{-1}V'']$$
$$-4\mu\,\dot{U}'\dot{U}''$$
$$-\mu\,\dot{V}''[\dot{V}' + r^{-1}k'U' - r^{-1}V']\cos\eta$$
$$-\mu\,\dot{V}'\,[\dot{V}'' + r^{-1}k''U'' - r^{-1}V'']\cos\eta;$$

Love-to-Rayleigh scattering,

$$\Omega^{(2)}_{L'R''} = \mu\,\dot{V}''\,[\dot{W}' - r^{-1}W']\sin\eta$$
$$+\mu\,\dot{W}'\,[\dot{V}'' + r^{-1}k''U'' - r^{-1}V'']\sin\eta;$$

Rayleigh-to-Love scattering,

$$\Omega^{(2)}_{R'L''} = -\mu\,\dot{V}'\,[\dot{W}'' - r^{-1}W'']\sin\eta$$
$$-\mu\,\dot{W}''\,[\dot{V}' + r^{-1}k'U' - r^{-1}V']\sin\eta.$$

The quantity $\eta = \arccos(\hat{\mathbf{k}}' \cdot \hat{\mathbf{k}}'')$ is the scattering angle, measured counter-clockwise from the incoming wavevector $\hat{\mathbf{k}}'$ to the outgoing wavevector $\hat{\mathbf{k}}''$. A dot denotes differentiation of U', V', W' or U'', V'', W'' with respect to radius r.