How are vertical shear wave splitting measurements affected by variations in the orientation of azimuthal anisotropy with depth?

Rebecca L. Saltzer, James B. Gaherty* and Thomas. H. Jordan

Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
E-mail: rsaltzer@mit.edu

SUMMARY
Splitting measurements of teleseismic shear waves, such as *SKS*, have been used to estimate the amount and direction of upper mantle anisotropy worldwide. These measurements are usually made by approximating the anisotropic regions as a single, homogeneous layer and searching for an apparent fast direction ($\tilde{\phi}$) and an apparent splitting time ($\Delta \tilde{t}$) by minimizing the energy on the transverse component of the back-projected seismogram. In this paper, we examine the validity of this assumption. In particular, we use synthetic seismograms to explore how a vertically varying anisotropic medium affects shear wave splitting measurements. We find that weak heterogeneity causes observable effects, such as frequency dependence of the apparent splitting parameters. These variations can be used, in principle, to map out the vertical variations in anisotropy with depth through the use of Fréchet kernels, which we derive using perturbation theory. In addition, we find that measurements made in typical frequency bands produce an apparent orientation direction that is consistently different from the average of the medium and weighted towards the orientation of the anisotropy in the upper portions of the model. This tendency of the measurements to mimic the anisotropy at the top part of the medium may explain why shear wave splitting measurements tend to be correlated with surface geology.

When the heterogeneity becomes stronger, multiple scattering reduces the amplitude of the tangential-component seismogram and the associated splitting time, so that a null result may be obtained despite the fact that the waves have travelled through a strongly anisotropic medium. Regardless of the amount of vertical heterogeneity, we find that there is very little dependence on backazimuth for the measured fast-axis direction or splitting time if the top and bottom halves of the medium average to similar fast-axis directions. If, however, the average fast-axis direction in the top half of the model differs from that in the bottom half, splitting-time measurements will show a significant dependence on backazimuth, but fast-axis direction measurements will remain relatively constant.

Key words: Fréchet derivatives, layered media, seismic anisotropy, shear-wave splitting.

1 INTRODUCTION

Measurements of seismic anisotropy are used to infer mantle deformation and flow patterns. While several different methods for constraining upper mantle anisotropy have been developed, such as *Pn* refraction surveys (e.g. Raitt *et al.* 1969; Shearer & Orcutt 1986) and surface wave polarization analyses

* Now at: School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA.

(e.g. Forsyth 1975; Nataf *et al.* 1984; Tanimoto & Anderson 1985; Montagner & Tanimoto 1990), the last decade has seen an explosion in shear wave splitting studies using vertically propagating shear waves (see reviews by Silver 1996 and Savage 1999). Typically, these analyses are performed on waves such as *SKS* or *SKKS* because they have a known polarization direction (S_V) as a result of passing through the liquid outer core. The standard procedure is to find the inverse splitting operator Γ^{-1} which, when applied to the observed waveform, minimizes the energy on the tangential component (Silver &

Chan 1991). When other phases such as *S* and *ScS* are used, the splitting parameters are found either by assuming a rectilinear source mechanism (Ando & Ishikawa 1982; Ando 1984) or by explicitly diagonalizing the covariance matrix of surface-corrected horizontal particle motions (Vidale 1986; Fouch & Fischer 1996).

A basic assumption in interpreting measurements using these techniques is that the splitting operator Γ corresponds to a single homogeneous layer in which the anisotropy has a horizontal symmetry axis and a constant magnitude. The parameters used to describe this model (the splitting parameters) are the polarization azimuth of the fast eigenwave, ϕ, and the travel-time difference between the fast and slow eigenwaves, Δt. It is straightforward to construct the splitting operator for an arbitrary stack of layers with depth-dependent properties and more general forms of anisotropy using propagator matrices (e.g. Keith & Crampin 1977; Mallick & Frazer 1990), but it is less clear how one might use such constructions to make inferences about anisotropic structure. A potentially fruitful direction is to fit the waveform data by optimizing the homogenous-layer operator and then interpret the two recovered quantities, denoted here by $\tilde{\phi}$ and $\Delta\tilde{t}$, as apparent splitting parameters which are functionals of the vertical structure. This approach was adopted by Silver & Savage (1994), who showed how an approximation to the variation of $\tilde{\phi}$ and $\Delta\tilde{t}$ with the incident polarization angle could be inverted for a two-layer anisotropic model. They also discussed the generalization of their approximate functional relations, which are valid for forward scattering at low frequencies (wave periods $\gg \Delta t$), to an arbitrary layer stack. Rümpker & Silver (1998) have recently expanded this theoretical discussion of vertical heterogeneity to include expressions for the apparent splitting parameters valid at high frequencies, as well as some statistical properties of the parameters for random layer stacks, and they have tested various aspects of their theory with numerical calculations.

In this paper, we consider several additional aspects of this interpretation problem. Using a propagator-matrix method that includes both forward- (upgoing) and back-scattered (downgoing) waves, we compute synthetic seismograms for various types of depth dependence, including smooth models as well as those with discontinuous variations in the anisotropy axis. We investigate the behaviour of the apparent splitting parameters with increasing amounts of vertical heterogeneity in the azimuthal anisotropy, and use the results to define three wave-propagation regimes corresponding to weak, inter-mediate, and strong scattering. For weakly heterogeneous media, we employ perturbation theory to calculate the sensitivity (Fréchet) kernels for band-limited, apparent-splitting measure-ments, and show how these measurements sample the depth dependence as a function of frequency and incident polarization angle. In realistic situations, the centre frequencies of the observations are sufficiently small that the kernels are one-sided, and we can define an apparent depth of sampling that we demonstrate is biased towards the upper part of the structure. In principle, the Fréchet kernels can be used to set up the problem of inverting frequency-dependent splitting measurements for depth-dependent anisotropy. We show that in practice, however, strong scattering by vertical heterogeneity can invalidate the assumptions that underlie this linearized approach, especially at higher frequencies.

2 PROBLEM FORMULATION

Because our purpose is to investigate some elementary aspects of vertical shear wave propagation, we adopt a very simple model for the mantle comprising a heterogeneous, anisotropic layer of thickness *d* overlying a homogeneous, isotropic half-space (Fig. 1). The anisotropy is assumed to be hexagonally symmetric with a horizontal axis of symmetry, and the lateral heterogeneity is assumed to be sufficiently smooth that hori-zontal gradients in the wave velocities can be ignored. Vertically propagating shear waves can thus be represented as linear combinations of orthogonal eigenwaves with shear velocities v_1 and v_2 that depend on the depth coordinate *z*. To simplify the problem further, we assume that the mean velocity $\bar{v} = (v_1 + v_2)/2$ and the velocity difference $\Delta v = v_1 - v_2$ are con-stants, and we label the eigenwaves such that $\Delta v > 0$. The heterogeneity in the medium is specified by a single function of depth that we take to be the azimuth of the fast (v_1) axis, $\phi(z)$, measured clockwise from the *x*-axis.

For the calculations in this paper, we adopt a layer thick-ness of $d = 200$ km and a mean velocity of $\bar{v} = 4.54$ km s^{-1}, and we take the velocity of the isotropic half-space to equal this mean velocity. The maximum splitting time for shear waves propagating from the base of the anisotropic layer to the surface—we ignore the crust—is $\Delta t = d(v_1 - v_2)/v_1 v_2 \approx d\Delta v/\bar{v}^2$, which represents the 'splitting strength' of the model. We refer to Δt in some of our numerical experiments as the 'true' splitting time.

2.1 Forward problem: synthetic seismograms

We use a stack of thin, homogeneous layers to represent the medium, and a propagator-matrix method to propagate shear waves vertically through the layers (e.g. Kennett 1983). For all our calculations, these layers are less than 1 km thick, in order to ensure that seismic wavelengths do not approach layer thick-ness. Boundary conditions restrict displacements and tractions to be continuous at the interfaces between the layers, and tractions to be zero at the surface. The Fourier-transformed,
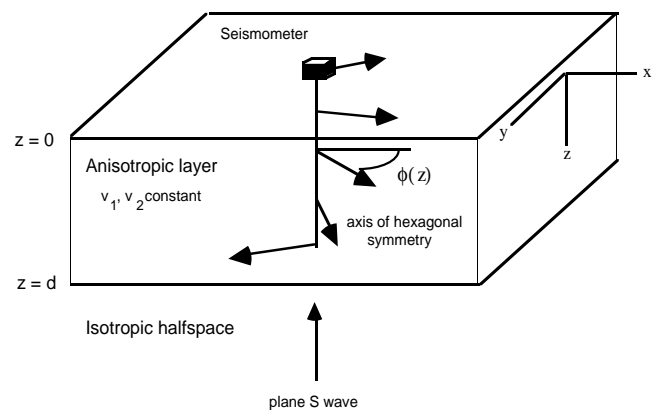

Figure 1. Model used in the calculations. A vertically travelling, rectilinearly polarized shear wave impinges at depth *d* on the base of a heterogeneous, anisotropic layer in which the fast-axis direction ϕ varies as a function of depth *z*. The two eigenvelocities, v_1 and v_2, are constant throughout the layer, and the velocity of the isotropic half-space is taken to be equal to their mean.

density-normalized stress vector $\tau(z, \omega) = \rho^{-1}[T_{xz}\, T_{yz}]^{\mathrm{T}}$ is related to the depth derivative of the displacement vector $\mathbf{u}(z, \omega) = [u_x u_y]^{\mathrm{T}}$ by the Christoffel matrix

$$\mathbf{C}(z) = \begin{bmatrix} v_1^2 \cos^2 \phi(z) + v_2^2 \sin^2 \phi(z) & (v_1^2 - v_2^2) \cos \phi(z) \sin \phi(z) \\ (v_1^2 - v_2^2) \cos \phi(z) \sin \phi(z) & v_1^2 \sin^2 \phi(z) + v_2^2 \cos^2 \phi(z) \end{bmatrix}. \tag{1}$$

The equations of motion are $\partial_z f = \mathscr{A} f$, where the displacement-stress vector and system matrix are given by

$$f(z, \omega) = \begin{bmatrix} \mathbf{u}(z, \omega) \\ \mathbf{t}(z, \omega) \end{bmatrix}, \tag{2}$$

$$\mathscr{A}(z, \omega) = \begin{bmatrix} \mathbf{0} & \mathbf{C}^{-1}(z) \\ -\omega^2 \mathbf{I} & \mathbf{0} \end{bmatrix}. \tag{3}$$

The rotation operator

$$\mathbf{U}(\phi(z)) = \begin{bmatrix} \cos \phi(z) & -\sin \phi(z) \\ \sin \phi(z) & \cos \phi(z) \end{bmatrix} \tag{4}$$

diagonalizes the Christoffel matrix: $\hat{\mathbf{C}} \equiv \mathrm{diag}[v_1^2, v_2^2] = \mathbf{U}\mathbf{C}\mathbf{U}^{\mathrm{T}}$.

The propagator matrix for this problem and some of its approximations are discussed in Appendix A. For an upgoing wave $\mathbf{u}_I(\omega)$ incident at the base of the anisotropic layer, the free-surface displacement vector can be written

$$\mathbf{u}(0, \omega) = [\mathbf{P}_{uu} + i\omega\bar{v}\mathbf{P}_{u\tau} + (\mathbf{P}_{uu} - i\omega\bar{v}\mathbf{P}_{u\tau})\mathbf{R}]\mathbf{u}_I(\omega), \tag{A16}$$

where $\mathbf{P}_{uu}$ and $\mathbf{P}_{u\tau}$ are 2×2 submatrices of the propagator matrix (A7), $\mathbf{R}$ is the 2×2 matrix of reflection coefficients (A15), and $\mathbf{U}_z = \mathbf{U}(\phi(z))$.

The pulse shape at the base of the anisotropic layer in all of our calculations is taken to be of the form $u_I(t) = \exp[-a/(t - t_0) - (t - t_0)/b] H(t - t_0)$ with a duration $a = 2$ s and a decay constant $b = 4$ s. The convolution of this initial pulse shape with the broad-band instrument response and pre-filter is given in Fig. 2(a). There is no energy on the tangential component because the initial pulse is radially polarized; however, propagation of the pulse through an anisotropic layer produces energy on both the radial and tangential components of the surface seismogram. An example of a synthetic seismogram calculated for a homogeneous anisotropic layer is shown in Fig. 2(b) for a velocity contrast of $\Delta v/\bar{v} = 4.54$ per cent. This corresponds to a splitting strength of $\Delta t = 2$ s, which lies towards the high end of the observations summarized by Silver (1996) and Savage (1999). The azimuth of the fast axis, measured clockwise from the radial ($\hat{\mathbf{x}}$) direction, is $45°$. The seismograms in this example are 'broad-band' with a corner at 50 mHz and at 300 mHz, and a centre frequency of ~ 140 mHz.

A simple model of a heterogeneous anisotropic layer, used extensively in our numerical illustrations, is one in which $\phi(z)$ varies linearly with depth:

$$\phi(z) = \phi_0 + \kappa z = \phi_d + \kappa(z - d). \tag{5}$$

The constant $\kappa = d\phi/dz$ is the vertical rotation rate. The orientation of the fast axis varies from ϕ_0 at the surface to ϕ_d at the base of the layer, and the total change in the orientation $\Delta\phi = \phi_d - \phi_0 = \kappa d$ measures the strength of the heterogeneity. Figs 2(c) to (e) show synthetic seismograms for the linear-rotation model with a splitting strength of $\Delta t = 2$ s and
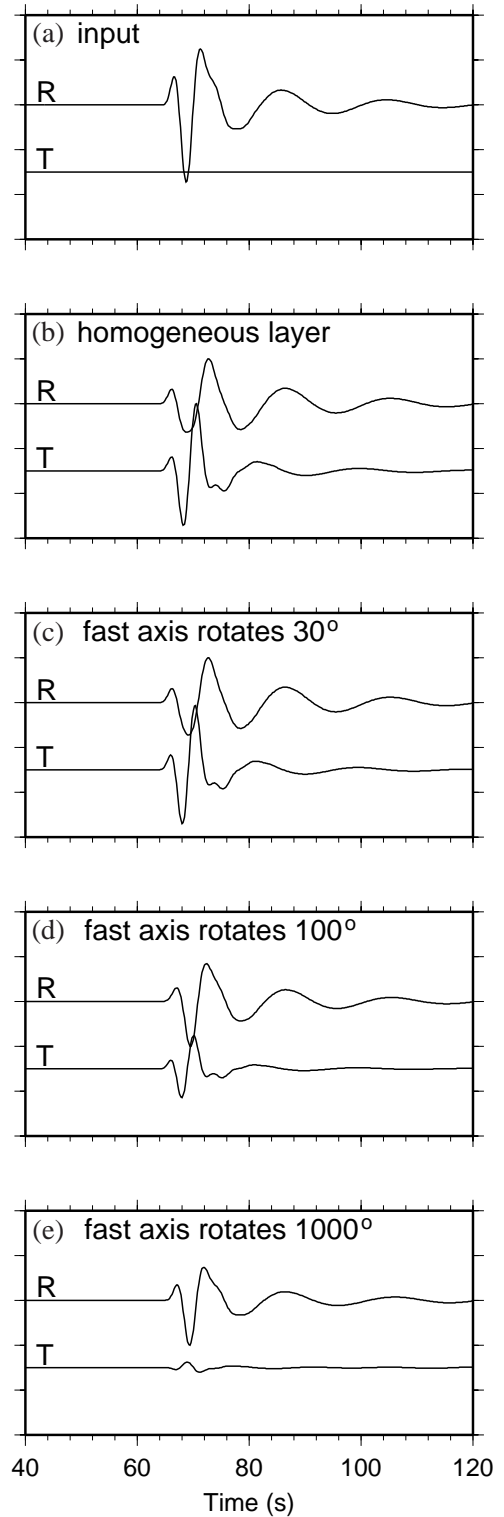

Figure 2. Radial- and tangential-component seismograms at (a) the base of the anisotropic layer, (b) the surface after passing through a homogeneous anisotropic layer, (c) the surface after passing through a weakly heterogeneous anisotropic layer in which the fast-axis direction ($\Delta\phi$) linearly rotates 30°, (d) intermediate heterogeneous anisotropic layer in which the fast-axis direction ($\Delta\phi$) linearly rotates 100°, and (e) strongly heterogeneous anisotropic layer in which the fast-axis direction ($\Delta\phi$) linearly rotates 1000°. Seismograms are band-pass filtered with a Butterworth filter to frequencies between 50 and 300 mHz.

increasing amounts of heterogeneity: $\Delta\phi = 30°$, $100°$, and $1000°$, respectively. In all three examples, the incident polarization was chosen such that the mean orientation of the fast axis was $45°$; that is,

$$\bar{\phi} \equiv \frac{1}{d} \int_0^d \phi(z)\,dz = \phi_0 + \Delta\phi/2 = \pi/4\,. \qquad (6)$$

In the case where $\Delta\phi = 30°$, the surface seismograms (Fig. 2c) are very similar to those produced in the homogeneous anisotropic case (Fig. 2b). Both the radial- and tangential-component seismograms have a comparable amount of energy. As $\Delta\phi$ increases to $100°$, the energy on the radial component becomes greater than that on the tangential component (Fig. 2d), and when $\Delta\phi$ is as large as $1000°$, the net effect of the anisotropy is to put very little energy onto the tangential component (Fig. 2e).

2.2 Inverse problem: apparent splitting parameters

If the incident pulse is known to be radially polarized, then the apparent splitting parameters $\tilde{\phi}$ and $\Delta\tilde{t}$ can be defined as the values that minimize the energy on the transverse ($\hat{\mathbf{y}}$) component of the displacement field back-projected to $z = d$ using a homogeneous-layer splitting operator (Silver & Chan 1991). Parseval's theorem allows the transverse-component energy to be written as a frequency-domain integral:

$$\varepsilon^2(\phi', \Delta t') = \int_{-\infty}^{\infty} |\hat{\mathbf{y}}^{\mathrm{T}} \Gamma_{\mathrm{h}}^{-1}(\phi', \Delta t')\mathbf{u}(0, \omega)|^2\,d\omega\,. \qquad (7)$$

Here Γ_{h}^{-1} is the inverse of the splitting operator given by eq. (A24). In practice, the determination of these so-called splitting parameters requires a search over a grid of fast-axis directions and delay times. For typical teleseismic observations, these parameters can be determined to within $\pm 10°$ and 0.15 s (Fouch & Fischer 1996; Silver & Chan 1991). Fig. 3 shows the 'energy map' contoured as a function of ϕ' ($0°$ to $180°$) and $\Delta t'$ (0–4 s) for the three linear-gradient examples shown in the previous section, and Fig. 4 shows the inferred seismograms at the base of the layer, calculated by back-projecting the splitting parameters and assuming propagation of waves through a homogeneous, anisotropic layer.

The energy map for the seismograms computed for a homogeneous layer (Fig. 3a) shows a well-defined minimum at the correct values of splitting parameters ($\tilde{\phi} = 45°$, $\Delta\tilde{t} = 2$ s). For weak heterogeneity ($\Delta\phi = 30°$), the energy minimum remains close to the layer mean ($\tilde{\phi} = 48°$, $\Delta\tilde{t} = 1.9$ s), and the bulk of the tangential-component energy has been removed (Fig. 4c). At intermediate values of the heterogeneity ($\Delta\phi = 30°$), the energy minimum is still well defined (Fig. 3c), but it is displaced away from the layer mean by $18°$ ($\tilde{\phi} = 63°$). Moreover, the scattering from the vertical gradients in the anisotropy is sufficient to reduce the apparent splitting time significantly below its true value ($\Delta\tilde{t} = 1.5$ s). When the heterogeneity gets to be very large ($\Delta\phi = 1000°$), the scattering is sufficiently strong as to cause destructive interference that nearly wipes out the arrivals on the transverse component (Fig. 2e). The resulting energy map (Fig. 3d) is characteristic of a 'null measurement', with the lowest values occurring near the horizontal axis where $\Delta t' = 0$ and along vertical ridges corresponding to the degenerate azimuths of $\phi = 0°$ and $90°$.

The behaviours illustrated in Fig. 3 are typical of three scattering regimes that can be qualitatively described as 'weak', 'strong coherent', and 'strong incoherent'.

3 WEAK-SCATTERING REGIME

When the scattering is weak, the effects of the heterogeneity on the apparent splitting parameters can be approximated with a linearized perturbation theory. In this section, we derive analytical expressions for perturbations from a homogeneous starting model, test their applicability with numerical calculations, and use them to gain insight into how the sensitivity of the apparent splitting parameters varies with depth.

3.1 Fréchet kernels

The linearized equations that relate a structural perturbation $\delta\phi(z)$ to perturbations in the apparent splitting parameters, $\delta\tilde{\phi}$ and $\delta\tilde{t}$, can be written as integrals over the layer:

$$\delta\tilde{\phi} \approx \int_0^d G_\phi(z)\,\delta\phi(z)\,dz\,, \qquad (8)$$

$$\delta\tilde{t} \approx \int_0^d G_t(z)\,\delta\phi(z)\,dz\,. \qquad (9)$$

This approximation ignores terms of order $\delta\phi^2$. $G_\phi(z)$ and $G_t(z)$ are the sensitivity functions, or Fréchet kernels. They generally depend on the structural model that is being perturbed, as well as on the spectral properties of the waves being measured. For the structures considered here, a model is specified by the fast-axis orientation function $\{\phi(z): 0 \leq z \leq d\}$ and the two velocity constants v_1 and v_2.

The Fréchet kernels can be numerically approximated for an arbitrary starting model by computing the small change in the splitting parameters due to a small perturbation in the fast-axis orientation distributed over a thin layer. However, in the special case of a homogeneous starting model ($\phi(z) = \phi_0$), the kernels for narrow-band pulses can be derived analytically. The details are relegated to Appendix B. The approximate results for a pulse with centre frequency ω_0 and half-bandwidth σ are

$$G_\phi = \frac{g_4 + [g_5 \cos 4\phi_0 + g_6](\sigma^2/\omega_0^2)}{g_1 + (g_2 \cos 4\phi_0 + g_3)(\sigma^2/\omega_0^2)}\,\Delta k + O\left(\frac{\sigma^4}{\omega_0^4}\right), \qquad (10)$$

$$G_t = \frac{g_7 + [g_8 \cos 4\phi_0 + g_9](\sigma^2/\omega_0^2)}{g_1 + (g_2 \cos 4\phi_0 + g_3)(\sigma^2/\omega_0^2)}\,\Delta k \cot 2\phi_0 + O\left(\frac{\sigma^4}{\omega_0^4}\right), \qquad (11)$$

where $\Delta k = \omega_0(v_2^{-1} - v_1^{-1})$ is the differential wavenumber. The three parameters appearing in the denominator of these expressions are independent of depth z:

$$g_1 = (1 - \cos \Delta kd)^2\,, \qquad (12a)$$

$$g_2 = \Delta kd\,[\sin^2 \Delta kd + (\Delta kd - 2)\sin \Delta kd \cos \Delta kd]\,, \qquad (12b)$$

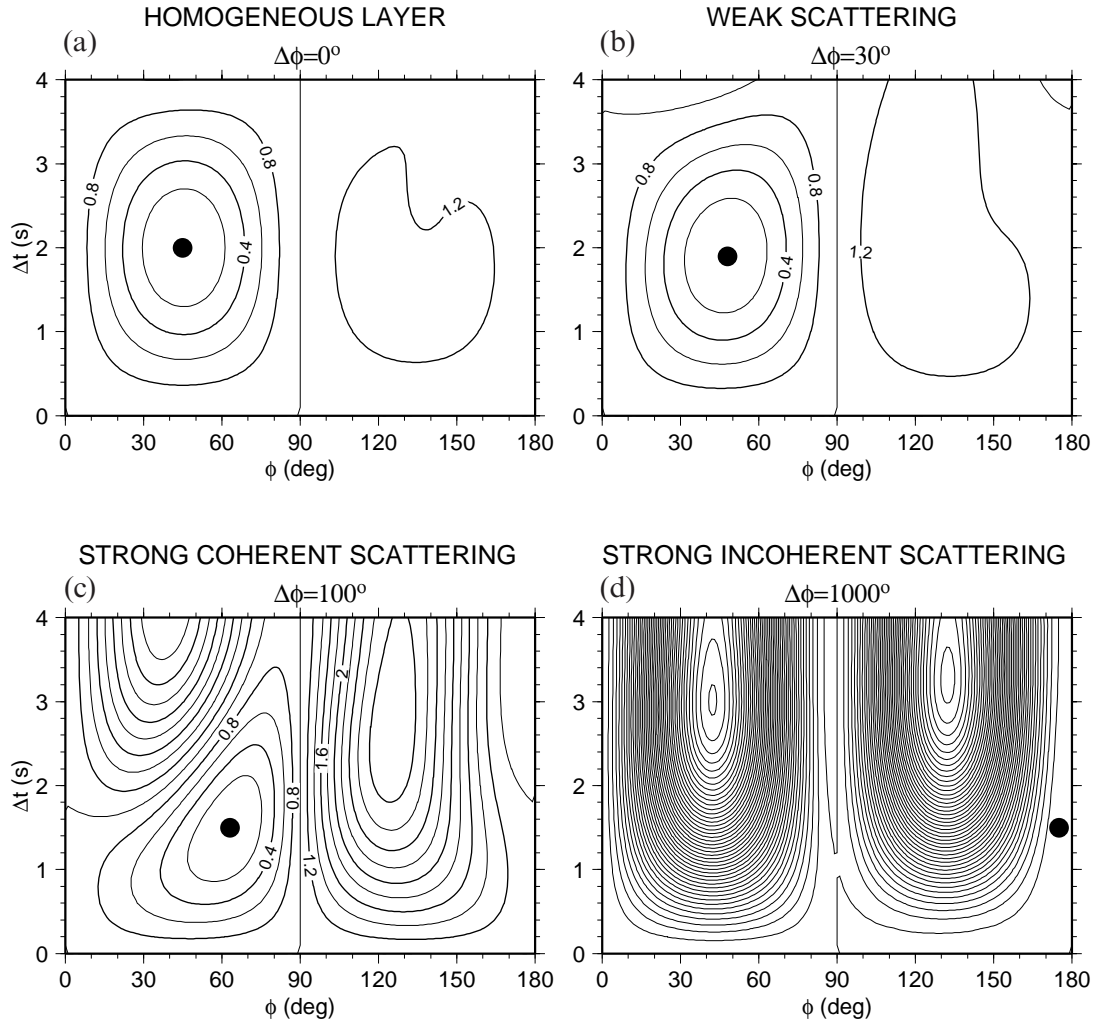$$g_3 = g_2 + 2(\Delta kd)^2[\sin^2 \Delta kd + (1 - \cos \Delta kd)\cos \Delta kd]\,. \qquad (12c)$$

Figure 3. Energy diagrams for (a) a homogeneous anisotropic layer, (b) a weakly heterogeneous anisotropic layer ($\Delta\phi = 30°$), (c) an intermediate heterogeneous anisotropic layer ($\Delta\phi = 100°$), and (d) a strongly heterogeneous anisotropic layer ($\Delta\phi = 1000°$).

The other six can be written in terms of trigonometric functions of the height variable $r = d - z$:

$$g_4(r) = (1 - \cos \Delta kd) \sin \Delta kr, \tag{13a}$$

$$g_5(r) = (\Delta kd - 1) \sin \Delta kd \cos \Delta kr + (\Delta kd - \Delta kd\Delta kr) \sin \Delta kd$$
$$\times \sin \Delta kr - \Delta kd \cos \Delta kd \cos \Delta kr, \tag{13b}$$

$$g_6(r) = g_5(r) + (\Delta kd)^2 \cos \Delta kd \sin \Delta kr + 2(\Delta kd)(\Delta kr) \sin \Delta kd$$
$$\times \cos \Delta kr - (\Delta kr)^2(1 - \cos \Delta kd) \sin \Delta kr, \tag{13c}$$

$$g_7(r) = (1 - \cos \Delta kd)^2 \cos \Delta kr - (1 - \cos \Delta kd) \sin \Delta kd \sin \Delta kr, \tag{14a}$$

$$g_8(r) = \sin \Delta kd(\sin \Delta kd - 2 \cos \Delta kd)$$
$$\times [\Delta kd \cos \Delta kr - (\Delta kd)(\Delta kr) \sin \Delta kr]$$
$$+ (\Delta kd)^2 \sin \Delta kd \cos \Delta kd \cos \Delta kr, \tag{14b}$$

$$g_9(r) = g_8(r) + 2(\Delta kd)^2$$
$$\times [\sin^2 \Delta kd + (1 - \cos \Delta kd) \cos \Delta kd] \cos \Delta kr$$
$$- (1 - \cos \Delta kd)^2[\Delta kr \sin \Delta kr + (\Delta kr)^2 \cos \Delta kr]. \tag{14c}$$

The relative bandwidth σ/ω_0, which is less than 0.5 in most seismological applications (e.g. Silver 1996; Fouch & Fischer 1996; Wolfe & Solomon 1998), is sufficiently small that it is safe to ignore the fourth-order terms in (10) and (11). Fig. 5 displays kernels computed under this approximation for a range of initial polarizations and centre frequencies. The sinuosity of the kernels increases with frequency, reflecting the first-order trigonometric dependence on $\Delta k\,z$. The kernel for the apparent splitting azimuth satisfies the lower boundary condition, $G_\phi(d) = 0$, which can be verified from the analytical expressions.

Integration of these expressions show that the kernel for $\delta\tilde{\phi}$ is unimodular and the kernel $\delta\tilde{t}$ averages to zero:

$$\int_0^d G_\phi(z)dz = 1, \tag{15}$$

$$\int_0^d G_t(z)dz = 0. \tag{16}$$

These properties must apply to the exact forms of the Fréchet kernels, not just to their narrow-band approximations given

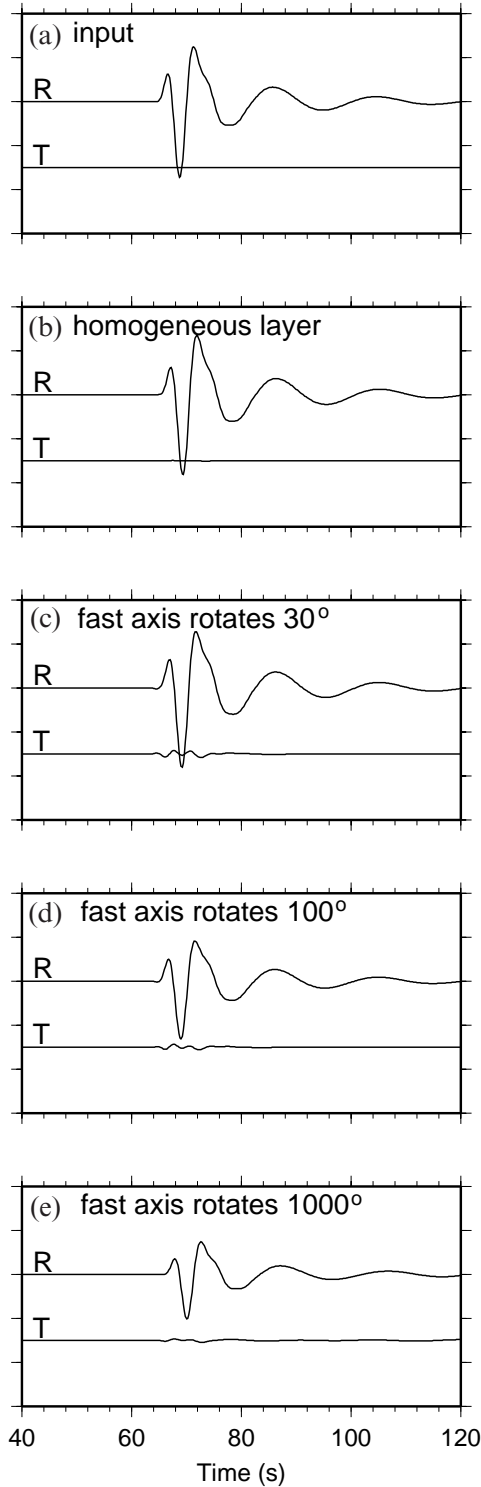© 2000 RAS, *GJI* **141,** 374–390

Figure 4. Radial- and tangential-component seismograms. After determining the apparent splitting parameters, the seismograms can be back-projected to the base of the layer again using those same parameters and assuming a homogeneous layer. (a) Input pulse at the base of the anisotropic layer to be compared with back-projected seismograms from (b) homogeneous anisotropic layer, (c) weakly heterogeneous anisotropic layer ($\Delta\phi$ linearly rotates 30°), (d) intermediate heterogeneous anisotropic layer ($\Delta\phi$ linearly rotates 100°), and (e) strongly heterogeneous anisotropic layer ($\Delta\phi$ linearly rotates 1000°). Seismograms are band-pass filtered with a Butterworth filter to frequencies between 50 and 300 mHz.

by (10) and (11), because a constant perturbation maintains the structural homogeneity of the starting model. In other words, setting the perturbation in (8) and (9) to a constant value of $\delta\phi$ must always yield $\delta\tilde{\phi} = \delta\phi$ and $\delta\tilde{t} = 0$.

3.2 Limiting forms of the kernels

By considering the zero-bandwidth limit, we gain additional insight into the nature of the sensitivity kernels:

$$G_\phi^0(z) \equiv \lim_{\sigma\to 0} G_\phi(z) = \frac{\sin \Delta k(d-z)}{1 - \cos \Delta kd}\Delta k, \qquad (17)$$

$$G_t^0(z) \equiv \lim_{\sigma\to 0} G_t(z)$$

$$= \left[\cos \Delta k(d-z) - \frac{\sin \Delta kd \sin \Delta k(d-z)}{1 - \cos \Delta kd} \right] \Delta k \cot \phi_0. \quad (18)$$

The splitting-time kernel (18) varies like the cotangent of $2\phi_0$ and is thus singular at $\phi_0 = n\pi/2$, which corresponds to incident polarizations aligned with an eigenwave orientation in the unperturbed model. In contrast to this singular behaviour, the splitting-azimuth kernel (17) is independent of the initial polarization, and remains well defined even at its degenerate values.

This important theoretical point deserves special emphasis. A shear wave with a polarization aligned with one of the eigenwave directions is not split by propagation through the reference model, and the tangential-component energy of the back-projected displacement field is thus identically zero at $\phi' = \phi_0 = n\pi/2$ for arbitrary values of $\Delta t' > 0$. Consequently, the energy map displays vertical nodal lines at these azimuths, as well as a horizontal nodal line at $\Delta t' = 0$, and the inversion of the seismograms for the apparent splitting parameters via the minimization of (7) becomes unstable (e.g. Silver & Chan 1991). Nevertheless, the apparent splitting azimuth remains formally defined in this limit by the orientation of the appropriate vertical node; that is, $\tilde{\phi}$ equals either ϕ_0 or $\phi_0 + \pi/2$. Moreover, $\tilde{\phi}$ is Fréchet differentiable, because a small perturbation to the model will result in a small, well-defined perturbation of the node in the ϕ' direction. The splitting-time functional $\Delta\tilde{t}$, on the other hand, is not Fréchet differentiable at the nodes, which is why its kernel is singular. This behaviour generalizes to pulse shapes with finite bandwidths, as is evident from eq. (11).

The cotangent dependence of the splitting-time kernel also implies that $G_t(z) = 0$ for $\phi_0 = n\pi/4$, which means that at polarization angles near 45° the apparent splitting time is only weakly dependent on perturbations to the local splitting orientation.

Eqs (17) and (18) show that, in the zero-bandwidth limit, both kernels become unbounded at $\Delta kd = 2\,n\pi$, where the total splitting time Δt is an integral multiple of the wave period $2\pi/\omega_0$. A finite bandwidth introduces a weak ($\sim\sigma^2/\omega_0^2$) dependence of G_ϕ and G_t on the initial azimuth through the $\cos 4\phi_0$ term, which suppresses this resonance singularity.

The expressions for the kernels presented thus far apply to splitting strengths that are arbitrarily large. At low frequencies, when the splitting strength is much less than the wave period, the maximum phase shift between the two eigenwaves is small, $\Delta kd = \omega_0\Delta t \ll 1$. If we expand the trigonometric functions and retain only the leading terms, the single-frequency kernels
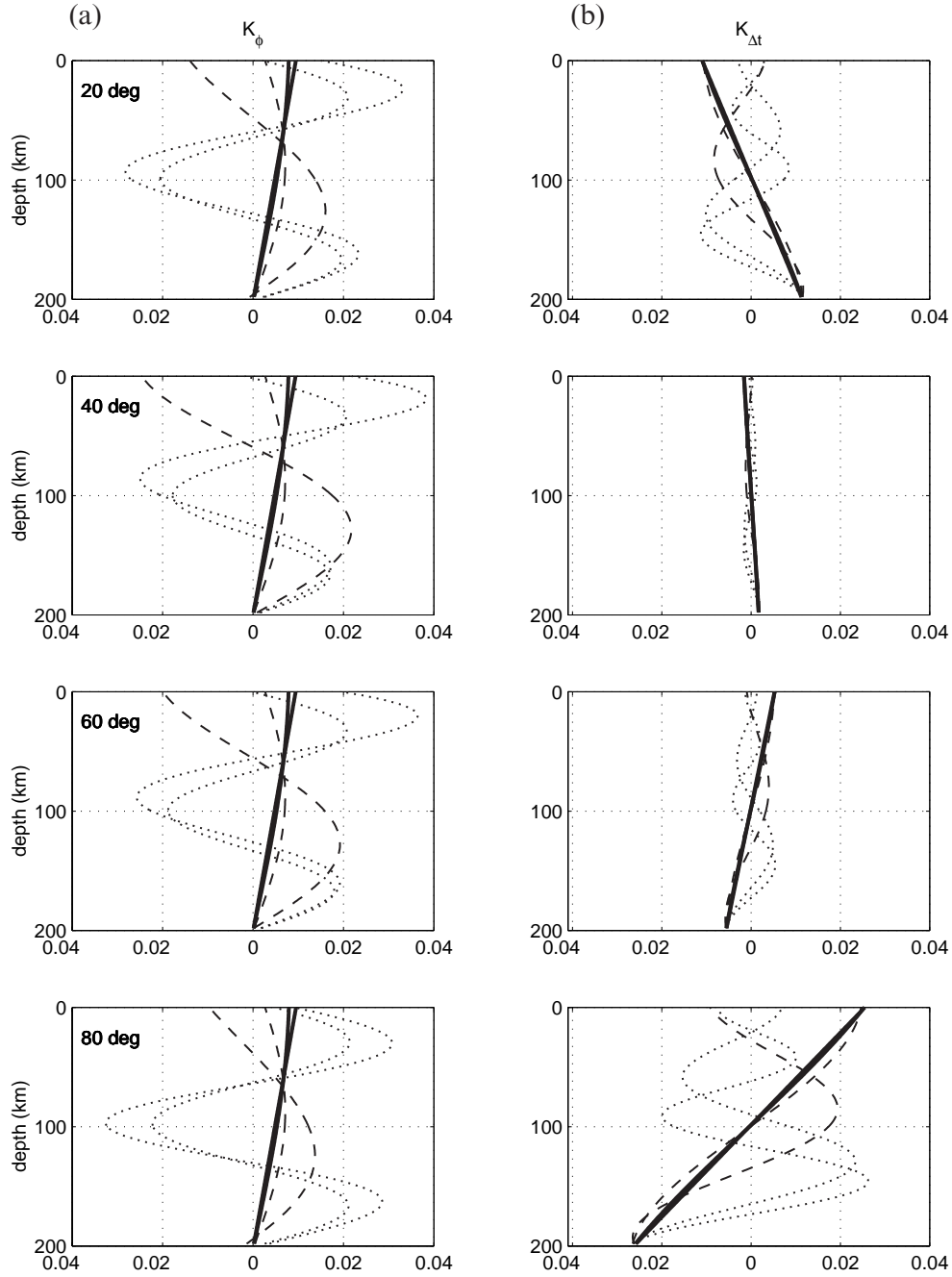
Figure 5. Fréchet kernels for the apparent splitting parameters computed from eqs (10) and (11) to second order in the relative bandwidth σ/ω_0. The reference model is a 200-km thick homogeneous layer with $\Delta t = 2$ s. Left panels show the apparent azimuth kernels $G_\phi(z)$, and right panels show the apparent splitting-time kernels $G_t(z)$. The initial polarization ϕ_0 increases downwards from 20° (top panel) to 80° (bottom panel). In each panel, the centre frequencies ω_0 range from 0.1 to 0.8 Hz (lowest frequencies are solid lines, intermediate are dashed, and highest frequencies are dotted), while the relative bandwidth is held constant at $\sigma/\omega_0 = 0.125$.

become linear functions of depth:

$$G_\phi^0(z) \approx \frac{2(d-z)}{d^2}, \tag{19}$$

$$G_t^0(z) \approx 2\Delta t \left(\frac{2z-d}{d^2} \right) \cot \phi_0. \tag{20}$$

Thus, in the low-frequency limit, the apparent splitting azimuth is insensitive to heterogeneity at the base of the layer and most sensitive to heterogeneity at the top of the layer. The sensitivity

of the apparent splitting time to azimuthal heterogeneity is zero in the middle of the layer, and it is of equal magnitude and opposite sign at the top and bottom of the layer.

3.3 Apparent depth of sampling

The previous discussion shows that the Fréchet kernel for the apparent splitting azimuth will be non-negative when the centre period of the wavegroup, $2\pi/\omega_0$, is greater than or equal to twice the splitting time Δt (i.e. $\Delta kd \leq \pi$). Under this condition,

which applies to most observations of teleseismic shear wave splitting, we can define an apparent depth of sampling by the centroid of the kernel:

$$z_{app} = \int_0^d G_\phi(z) z \, dz \, . \tag{21}$$

Using (15), we obtain for the zero-bandwidth limit

$$z_{app}^0 = \lim_{\sigma \to 0} z_{app} = \frac{\sin \Delta kd - \Delta kd \cos \Delta kd}{\Delta kd(1 - \cos \Delta kd)} \, . \tag{22}$$

In the low-frequency limit, where the $\tilde{\phi}$ kernel becomes a linear function of depth, the apparent depth of sampling goes to one-third of the layer thickness. This value increases to half of the layer thickness as the centre period approaches the splitting time. Fig. 6 shows three low-frequency $\tilde{\phi}$ kernels and their apparent depths of sampling for a 200-km thick, anisotropic layer. In the low-frequency limit, the apparent depth of sampling is just 66 km, whereas, for data low-pass filtered at 100 mHz, the depth of sampling increases to 71 km, and, for data low-pass filtered at 200 mHz, it increases to 88 km.

This bias in the sensitivity to near-surface structure explains why the recovered value for $\tilde{\phi}$ in the weak scattering case (Fig. 3b) is greater than the layer mean by about 3°. In this model, the fast axis rotates linearly from 40° at the base of the layer to 50° at the top. The kernels predict that the upper part of the model, where the fast axis ranges between 45° and 50°, will dominate the shear wave splitting measurement, and indeed the value found numerically ($\tilde{\phi} = 48°$) agrees with this

prediction, corresponding to the fast-axis direction approximately one-third of the way down the layer. This increase in sensitivity to the fast-axis direction near the surface may explain why shear wave splitting measurements tend to correlate with tectonic deformation observed at the surface (Silver 1996). Most shear wave splitting measurements are made on seismograms with relatively low centre frequencies (< 200 mHz), so that the apparent depth of sampling is less than the mean thickness of the layer. There are significant variations in the centre frequencies used by different researchers, however, so this apparent depth varies from study to study.

3.4 Numerical tests

We conducted a series of numerical experiments to test the perturbation theory derived for a homogeneous model. To investigate the sensitivity of the kernels to the reference structure, we have computed them by numerical perturbation to heterogeneous starting models. The results for starting models with a linear gradient and a step-wise discontinuity in $\phi(z)$ are compared with the homogeneous-layer case in Fig. 7. The average orientation was chosen to be the same for all three models, $\bar{\phi} \equiv d^{-1} \int_0^d \phi(z) dz = 45°$, while the total variation in $\phi(z)$ was taken to be 20° for the two heterogeneous models. The kernels are very similar, indicating only a weak dependence on the starting model when the heterogeneity is of this magnitude. In particular, the apparent depths of sampling for the layered and linear models are 91 km and 90 km, respectively, essentially the same as the value of 88 km calculated for the homogeneous model. The properties found for the homogeneous-layer kernels, such as their dependence on frequency, bandwidth, and incidence azimuth, should therefore pertain more generally in the weak-scattering regime.

We also compared the perturbations calculated from the Fréchet kernels using eqs (8) and (9) with the results of a direct numerical calculation that minimized the tangential-component energy on back-projected synthetic seismograms. Fig. 8 shows several examples of these comparisons as a function of the average polarization direction $\bar{\phi}$ for an initial pulse with a centre frequency at 60 mHz and corners at 45 mHz and 75 mHz. Fig. 9 shows the results for $\bar{\phi} = 45°$ for increasing values of the centre frequency ω_0. We note that care must be taken in the numerical calculations when evaluating the apparent splitting parameters near the azimuthal nodes at $\phi_0 = n\pi/2$, because the energy surfaces can be very flat in the $\Delta t'$ direction, and the location of the minimum is susceptible to numerical inaccuracies that can cause a $\pi/2$ ambiguity.

When the heterogeneity is small ($\Delta \phi = 10°$) in the linear-gradient models, the kernels do a good job of predicting $\tilde{\phi}$ and $\Delta \tilde{t}$ for all backazimuths (Fig. 8a). For heterogeneity with a total rotation angle as large as 60° (Fig. 8b), the kernels typically overestimate the apparent splitting time by about 0.3 s, and, near the nodes, $\tilde{\phi}$ can be off by as much as 20°. This failure of the kernels with greater heterogeneity reflects a breakdown in the small-angle approximations (e.g. eq. B5). Inclusion of the back-scattering terms in the calculation of synthetic seismograms (solid dots in Fig. 8) does not produce significantly different results from those obtained using just the forward-scattering terms.

In a second set of comparisons, we use what is perhaps a more geologically relevant model comprising 10 layers, each 20 km thick, of differing orientations constrained such that the
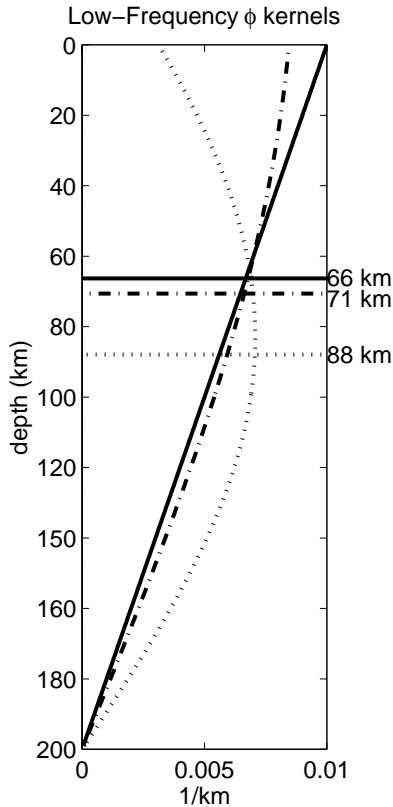

Figure 6. Low-frequency Fréchet kernels and the apparent depth of sampling computed with eq. (20). The solid line is the kernel in the low-frequency limit, the dashed line is the low-passed kernel for 100 mHz, and the dotted line is the low-passed kernel at 200 mHz.
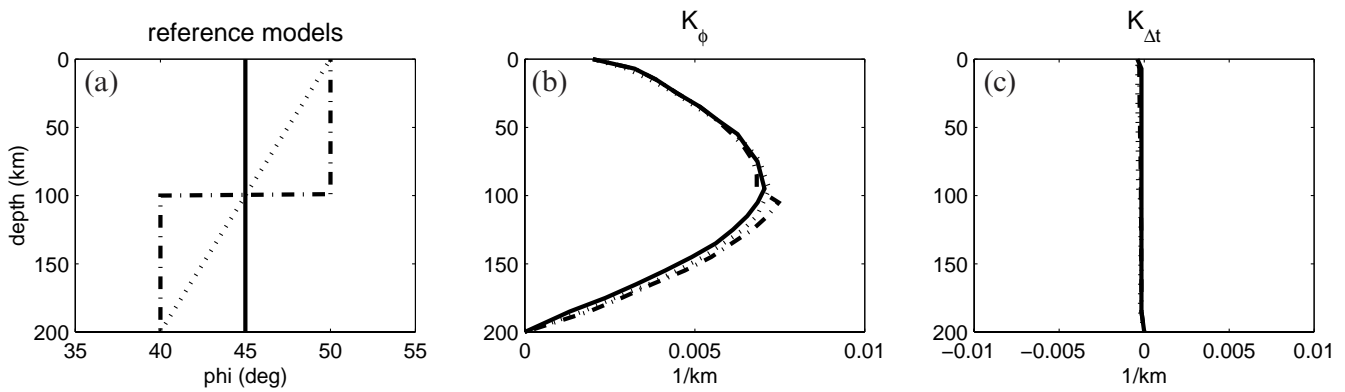
Figure 7. A comparison of Fréchet kernels for three starting models. Panel (a) is a plot of $\phi(z)$ for the homogeneous-layer (solid line), constant-gradient (dotted line), and two-layer (dashed line) models. Panels (b) and (c) show the corresponding kernels for these models, $G_\phi(z)$ and $G_t(z)$, respectively. The kernels were calculated by a numerical perturbation scheme for $\omega_0 = 0.14$ Hz, $\sigma = \omega_0/6$ Hz, which are similar to the values used in the processing of teleseismic shear waves. The models have the same average azimuth, $\bar{\phi} = 45°$. Eq. (10) shows that $G_t(z) = 0$ for a homogeneous layer with this initial azimuth. The agreement illustrates the weak dependence of the kernels on the starting model.

fast-axis directions varies between $0°$ and $30°$. When the layer orientations are distributed randomly, such that the average orientation in the top half of the model is similar to that in the bottom half, the agreement between the numerical results and the predictions of the analytical kernels is usually very good (Fig. 8c). In addition, both the exact values of the apparent splitting parameters and perturbation-theory predictions show very little dependence on the polarization angle, with the variations in the apparent splitting time associated with nodal singularities compressed into a narrow range of azimuths.

When the layer orientations are skewed, however, such that the average orientation in the top half of the model differs significantly from that in the bottom half (Fig. 8d), the $\pi/2$ periodicity in $\Delta\tilde{t}$ associated with the nodal singularities becomes more pronounced. This difference in the variation of $\Delta\tilde{t}$ with initial azimuth results from the fact that $G_t(z)$ is approximately a linear function of depth that averages to zero, as seen from its low-frequency form (18); that is, the perturbation to the apparent splitting time will be small and the $\pi/2$ periodicity will be suppressed when the first moment $\int_0^d \phi(z)z\,dz$ is small. For a specified level of heterogeneity, the constant-gradient case has the largest first moment of any model, which is why the initial-azimuth dependence in Figs 8(a) and (b) is so pronounced.

These results can be used to qualify Silver & Savage's (1994) argument that a $\pi/2$ periodicity in initial azimuth should be diagnostic of vertical heterogeneity. This periodicity will be relatively weak for heterogeneous structures where the azimuth of the anisotropy does not vary systematically with depth.

On increasing the heterogeneity in the 10 random layers so that the fast-axis direction ranges over $120°$, we find azimuthal discrepancies of up to $\pm5°$ and splitting-time discrepancies exceeding 1 s (Figs 8e and f). At this level of heterogeneity, back-scattering effects, given by the differences between the open and solid circles, begin to become important.

A comparison of the analytical and numerical results for different frequencies at a single backazimuth shows that the kernels do a nice job of predicting $\tilde{\phi}$ and $\Delta\tilde{t}$ up to 0.5 Hz when the heterogeneity is weak (Figs 9a and c), but that when the heterogeneity gets stronger, the kernels break down more quickly

at higher frequencies (Figs 9b and d) than at lower frequencies. In these examples, $\Delta t = 2.14$ s. At frequencies approaching $1/\Delta t$ (~ 0.46 Hz) the kernels predict highly oscillatory behaviour in $\tilde{\phi}$. This also corresponds to the frequency at which the kernels become unbounded in the single-frequency limit.

4 STRONG-SCATTERING REGIME

The numerical experiments demonstrate that the first-order perturbation theory expressed in eqs (8) and (9) provides an accurate description of the apparent splitting parameters in situations where the magnitude of the vertical heterogeneity is small. As this magnitude increases, the perturbation theory fails because the small-angle approximations employed in obtaining the scattering matrix (B5) and the linearized minimization condition (B12) become inaccurate, owing to the accumulating effects of multiple forward-scattering. The combination of these strong-scattering effects causes the behaviour of the apparent splitting parameters to deviate from the weak-scattering results.

Aspects of this behaviour were noted in the previous discussion of Fig. 3, which displays the numerical results for a constant-gradient model. In these calculations, $\Delta t = 2$ s, $\bar{\phi} = 45°$, and all forward- and back-scattering terms were retained. For a $45°$ average polarization, the general form of the kernel (11) shows that the splitting-time perturbation should be zero to first order; that is, the apparent splitting time $\Delta\tilde{t}$ should equal the total splitting strength Δt. Strong scattering acts to reduce $\Delta\tilde{t}$ below this theoretical limit, so that the ratio $(\Delta t - \Delta\tilde{t})/\Delta t$, to the extent that it can be accurately estimated, measures the higher-order effects. For the $30°$ rotation in Fig. 2(c), this reduction is only about 5 per cent, consistent with the weak-scattering approximations. The $120°$ rotation in Fig. 2(d) gives a much more substantial effect (~ 35 per cent), indicating that these approximations are not accurate for heterogeneity of this magnitude. For the $1000°$ rotation, the scattering is sufficiently large that the tangential-component arrivals are incoherent, so that the long-period amplitude is nearly zero, and the energy diagram looks nodal. In this case, there is no well-defined energy minimum, and it is difficult to measure the apparent splitting parameters.
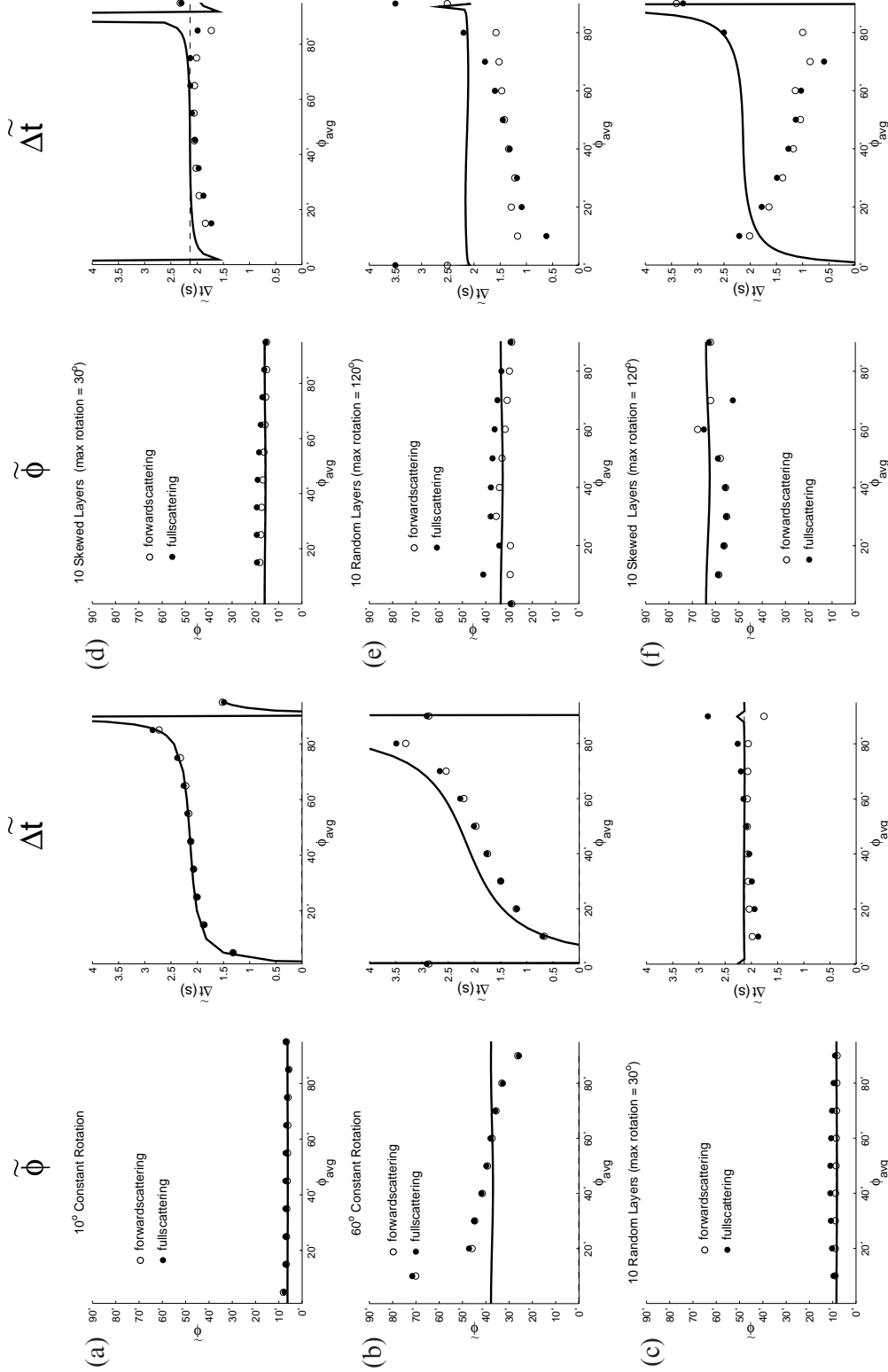
Figure 8. Comparison of numerical and analytical results. Left panels show $\tilde{\phi}$, and right panels show $\Delta \tilde{t}$, plotted as a function of backazimuth. Solid lines are analytical predictions from the kernels. Solid circles are the numerical results including back-scattering, and open circles include just forward-scattering. (a) Weakly heterogeneous ($\Delta \phi = 10°$) linear rotation model. (b) Strongly heterogeneous ($\Delta \phi = 60°$) linear rotation model. (c) Weakly heterogeneous 10-random-layer model (maximum $\Delta \phi = 30°$) in which the average fast-axis directions in the top and bottom halves are similar. (d) Weakly heterogeneous 10-random-layer model (maximum $\Delta \phi = 30°$) in which the layer orientations in the top half are different from those in the bottom half. (e) Strongly heterogeneous 10-random-layer model (maximum $\Delta \phi = 120°$) in which the average fast-axis directions in the top and bottom halves are similar. (f) Strongly heterogeneous 10-random-layer model (maximum $\Delta \phi = 120°$) in which the layer orientations in the top half are different from those in the bottom half.
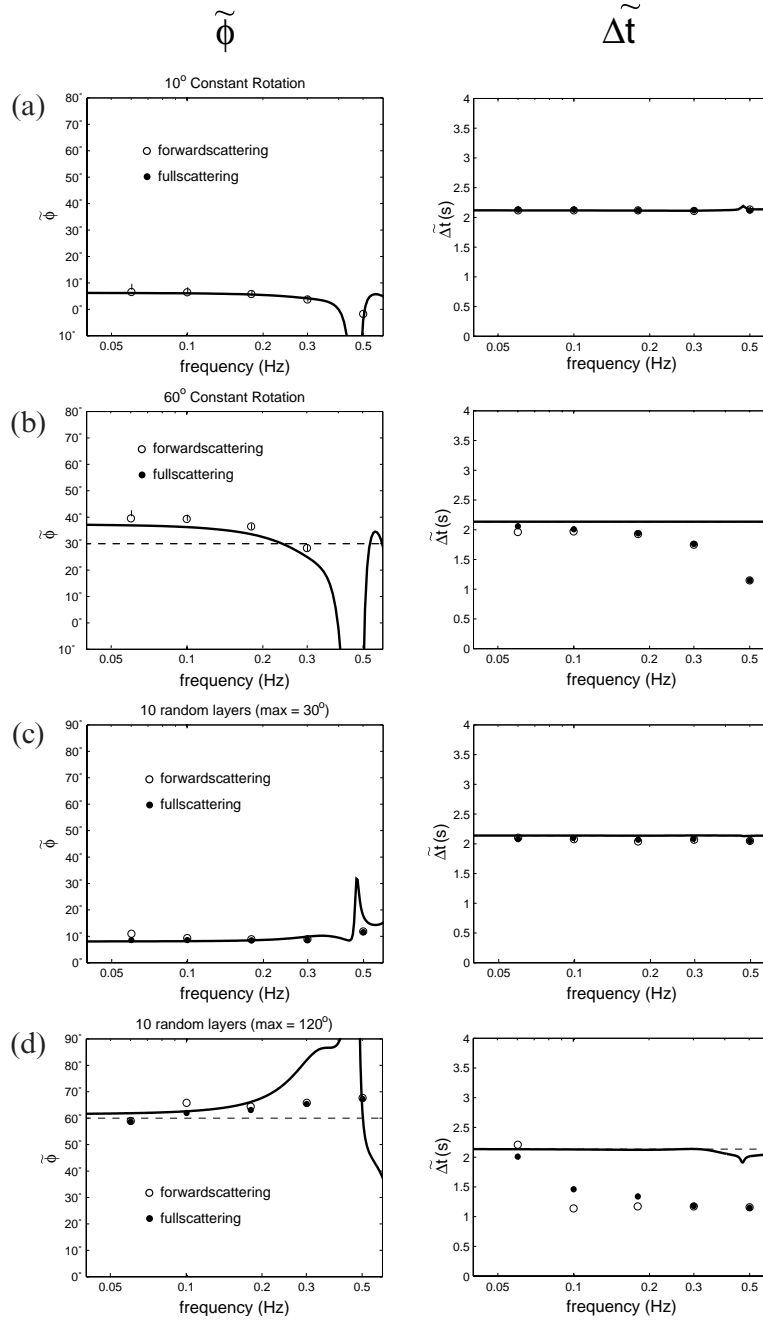
Figure 9. Comparison of numerical and analytical results. Left panels show $\tilde{\phi}$ and right panels show $\Delta\tilde{t}$, plotted as a function of frequency. Solid lines are analytical predictions from the kernels. Solid circles are the numerical results with both forward- and back-scattering, and open circles include just forward-scattering. (a) Weakly heterogeneous ($\Delta\phi = 10°$) linear rotation model. (b) Strongly heterogeneous ($\Delta\phi = 60°$) linear rotation model. (c) Weakly heterogeneous 10-random-layer model (maximum $\Delta\phi = 30°$). (d) Strongly heterogeneous 10-random-layer model (maximum $\Delta\phi = 120°$).

In Fig. 10, we extend these calculations to constant-gradient models with $\bar{\phi} = 45°$ and a range of splitting and heterogeneity strengths. The ordinate is taken to be $1/\kappa$, a quantity proportional to the inverse of the heterogeneity gradient, which defines a vertical correlation length. The contours of apparent splitting time on this plot can be used to delineate the three scattering regimes. The region with nearly vertical contours at large values of the correlation length corresponds to weak scattering, where $\Delta\tilde{t} \approx \Delta t$ and the perturbation theory of

Section 3 is valid. The deviation of these contours towards the horizontal defines a region where the scattering is too strong for perturbation theory to apply but not so strong as to prohibit the estimation of the apparent splitting parameters. The boundary between these two scattering regimes, indicated by the dashed line in Fig. 10, is given by a correlation length that increases exponentially with the anisotropy strength. As the correlation length of the vertical heterogeneity decreases at constant Δt, the apparent splitting time decreases, at first
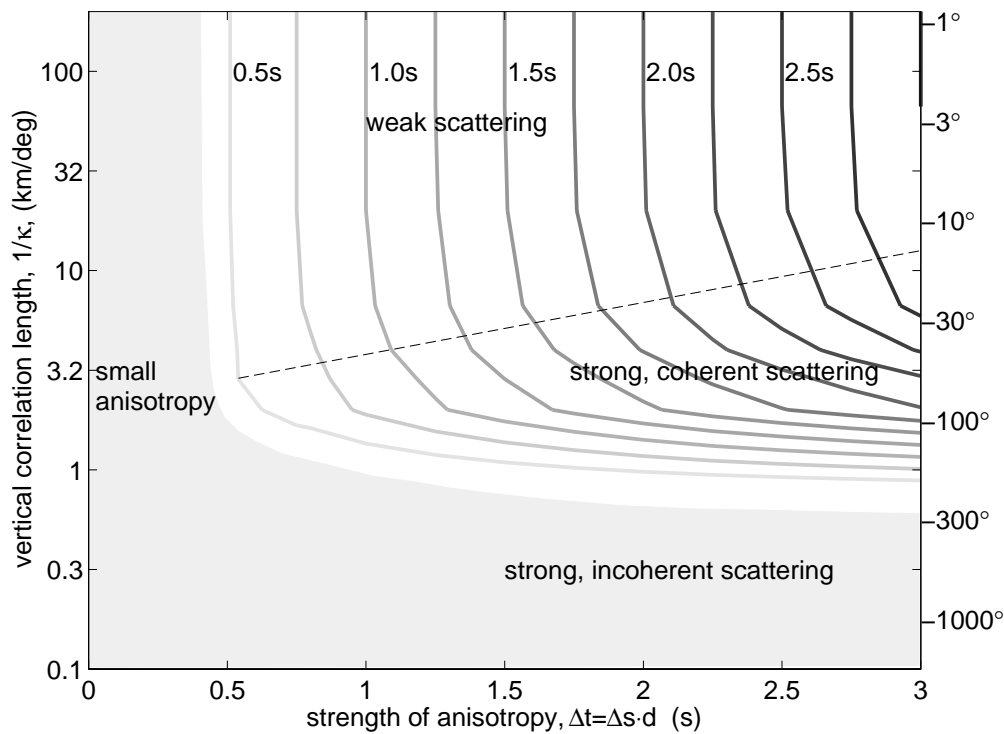
Figure 10. Scattering diagram contour of splitting-time measurements $\Delta \tilde{t}$ for smoothly varying models. The horizontal axis shows the range of anisotropy in the models ($\Delta t'$), and the vertical axis shows the range of vertical heterogeneity in the models (κ is the rotation rate). Splitting measurements cannot be made when $\Delta t' < 0.5$ s or $1/\kappa < 1$ km deg^{-1}.

slowly then rapidly. Below some critical value of the correlation length (~ 50 km in this example, corresponding to $\kappa = 1^\circ$ km^{-1}), the scattering becomes so strong that $\Delta \tilde{t}$ cannot be defined.

As this diagram makes clear, it is not possible to distinguish on the basis of low-frequency splitting observations the difference between highly heterogeneous anisotropy (Δt and κ large) and weak anisotropy (Δt small). Analysis of horizontally propagating surfaces waves (e.g. Jordan & Gaherty 1996) is one way to distinguish between these two cases.

5 DISCUSSION

For the case of weak scattering, differences in the sensitivity kernels as a function of frequency can, in principle, be exploited to invert frequency-dependent shear wave splitting measurements for a picture of anisotropic variation with depth. To do so, one applies standard splitting analysis to broad-band recordings in order to obtain an apparent fast direction to be used as a starting estimate for $\bar{\phi}$. Frequency-dependent apparent splitting parameters are then extracted by applying the back-projection procedure to narrow-band filtered seismograms, and their kernels constructed from (10) and (11). Adherence to the weak-scattering regime can be checked by confirming that minimal signal remains on the tangential-component seismogram after back-projection via the apparent splitting parameters (Fig. 4). The frequency-dependent splitting parameters can then be inverted for $\phi(z)$. (In all of our calculations, the difference between the speeds of the two eigenwaves remained constant throughout the model. In the real world, this parameter, like the anisotropy orientation, probably varies with depth. It is a simple matter, however, to extend the theory to depth-dependent wave speeds.)

The applicability of this procedure is likely to be limited by difficulties in extracting apparent splitting parameters at higher frequencies. Above 0.1 Hz, observed shear waveforms become increasingly complex due to microseisms, crustal scattering, and other sources of 'noise'. In addition, split shear waves have distinct 'holes' in their amplitude spectra at frequencies with integer multiples of $1/\Delta t$ (Silver & Chan 1991), which complicate analysis at higher frequencies. As a result, most shear wave splitting analyses in the literature utilize centre frequencies that fall within a relatively narrow frequency band of approximately 0.05–0.2 Hz (e.g. Silver 1996; Fouch & Fischer 1996; Wolfe & Solomon 1998). Our numerical experiments in the weak-scattering regime indicate that, across this bandwidth, variations in $\Delta \tilde{t}$ and $\tilde{\phi}$ are generally less than 0.1 s and 5°, respectively (Figs 8a and 9a). These variations are smaller than the typical error estimates in observational studies, and thus they cannot resolve changes in anisotropy with depth.

In the case of strong scattering, the approximations made in deriving the kernels are no longer valid, so that the kernels cannot be used to solve the inverse problem. We can, however, utilize the numerical results in the interpretation of splitting observations. Marson-Pidgeon & Savage (1997) report frequency-dependent shear wave splitting results from New Zealand (between 50 and 200 mHz) that are consistent with our numerical results for the strong coherent scattering regime (Figs 9b and d), implying significant vertical heterogeneity with depth. In addition, splitting results from cratons in South Africa (Gao *et al.* 1998), Australia (Clitheroe & van der Hilst 1998; Özalaybey & Chen 1999), India (Chen & Özalaybey 1998), and Tanzania (Hill *et al.* 1996) all find splitting times that are smaller (generally < 0.6 s) than for many other continental environments (e.g. Silver 1996). Such observations are

typically interpreted as evidence for little or no anisotropy. Our calculations provide an alternative explanation for these null results in terms of strong incoherent scattering in an upper mantle that is anisotropic, but has a high degree of vertical heterogeneity. Other data, such as horizontally propagating surface waves, are necessary to distinguish between these two possibilities.

In at least two cratonic regions where null or near-null results are reported (Australia and Southern Africa), analyses using surface waves find that the upper mantle is anisotropic between the Moho and 200–250 km depth, with $\Delta v_s/v_s$ of approximately 3–4 per cent (Gaherty & Jordan 1995; Saltzer *et al.* 1998). If perfectly aligned, such anisotropy would produce over 2 s of splitting. We interpret the apparent discrepancy between the splitting and surface wave results as evidence for strongly heterogeneous anisotropy in these regions. In particular, both sets of observations can be explained by a model in which the local anisotropy axis remains, on average, close to horizontal, but varies in azimuth as a function of position, with a vertical correlation length of the order of 50 km (Jordan *et al.* 1999).

6 CONCLUSIONS

In both weakly and strongly heterogeneous media we find that shear wave splitting measurements made in the low-frequency bands typically used to make observations are more sensitive to the upper portions of the model than to the lower portions. Consequently, if the orientation of the anisotropy is heterogeneous, the measured splitting direction will be more reflective of the fast-axis direction near the top of the upper mantle, and the measured splitting time will vary as a function of backazimuth. This effect may explain why global shear wave splitting measurements tend to correspond to the local tectonic fabric in crust beneath the observing station (Silver 1996).

We have derived analytic expressions for sensitivity kernels that relate a perturbation in the measured splitting parameters to a perturbation in the anisotropy of the model. For weakly heterogeneous media, frequency-dependent shear wave splitting measurements can be inverted using these kernels to determine how the anisotropy varies as a function of depth. Variations in the strength of the anisotropy as a function of depth can be accounted for with a weighting function. Practically speaking, however, a sufficient number of measurements may not be available at high enough frequencies or with enough precision (errors are typically $\pm 10°$ or more) for such an inversion to be feasible.

Strongly heterogeneous media (i.e. where the fast axis direction varies anywhere between $0°$ and $180°$) have the additional property that they cause strong scattering. This scattering will cause the tangential-component seismograms to have very little energy, and a null-like energy measurement will be made despite the fact that the medium may be highly anisotropic. Our numerical experiments show that the greatest diagnostic of strong vertical heterogeneity is the drop-off in $\Delta \tilde{t}$ relative to Δt.

REFERENCES

Ando, M., 1984. *ScS* polarization anisotropy around the Pacific Ocean, *J. Phys. Earth,* **32,** 179–195.

Ando, M. & Ishikawa, Y., 1982. Observations of shear-wave velocity polarization anisotropy beneath Honshu, Japan: two masses with different polarizations in the upper mantle, *J. Phys. Earth,* **30,** 191–199.

Chen, W.-P. & Özalaybey, S., 1998. Correlation between seismic anisotropy and Bouguer gravity anomalies in Tibet and its implications for lithospheric structures, *Geophys. J. Int.,* **135,** 93–101.

Clitheroe, G. & van der Hilst, R., 1998. Complex anisotropy in the Australian lithosphere from shear-wave splitting in broad-band *SKS* records, in *Structure and Evolution of the Australian Continent,* eds Braun, J., Dooley, J., Goleby, R., van der Hilst, R. & Klootwijk, C., Am. geophys. Un. Geodynamic Series, **26,** 39–57.

Forsyth, D.W., 1975. The early structural evolution and anisotropy of the oceanic upper-mantle, *Geophys. J. R. astr. Soc.,* **43,** 103–162.

Fouch, M.J. & Fischer, K.M., 1996. Mantle anisotropy beneath northwest Pacific subduction zones, *J. geophys. Res.,* **101,** 15 987–16 002.

Gaherty, J.B. & Jordan, T.H., 1995. The Lehman discontinuity as the base of an anistropic layer beneath the continents, *Science,* **268,** 1468–1471.

Gao, S.S., Silver, P.G. & James, D.E. & the Kaapvaal Working Group, 1998. Seismic structure and tectonics of southern Africa—progress report, *EOS, Trans. Am. geophys. Un.,* **79,** 45.

Hill, A.A., Owens, T.J., Ritsema, J., Nyblade, A.A. & Langston, C.A., 1996. Constraints on the upper mantle structure beneath the Tanzanian Craton from teleseismic body waves, *EOS, Trans., Am. geophys. Un.,* **77,** F477.

Jordan, T.H. & Gaherty, J.B., 1996. Stochastic modeling of small-scale, anisotropic structure in the continental upper mantle, in *Proc. 17th Ann. seismic Res. Symp.,* pp. 433–451, eds Lewkowicz, J.F., McPhetre, J.M. & Reide, D.T., Phillips Labs, Hanscom Air Force Base.

Jordan, T.H., Saltzer, R.L. & Gaherty, J.B., 1999. Small-scale anisotropic heterogeneity in the continental upper mantle, *Seism. Res. Lett.,* **70,** 259.

Keith, C.M. & Crampin, S., 1977. Seismic body waves in anisotropic media: synthetic seismograms, *Geophys. J. R. astr. Soc.,* **49,** 225–243.

Kennett, B.L.N., 1983. *Seismic Wave Propagation in Stratified Media,* Cambridge University Press, Cambridge.

Mallick, S. & Frazer, N.L., 1990. Computation of synthetic seismograms for stratified, azimuthally anisotropic media, *J. geophys. Res.,* **95,** 8513–8526.

Marson-Pidgeon, K. & Savage, M.K., 1997. Frequency-dependent anisotropy in Wellington, New Zealand, *Geophys. Res. Lett.,* **24,** 3297–3300.

Montagner, J.P. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. geophys. Res.,* **95,** 4797–4819.

Nataf, H.C., Nakanishi, I. & Anderson, D.L., 1984. Anisotropy and shear-velocity heterogeneities in the upper mantle, *Geophys. Res. Lett.,* **11,** 109–112.

Özalaybey, S. & Chen, W.-P., 1999. Frequency-dependent analysis of SKS/SKKS waveforms observed in Australia: Evidence for null birefringence, *Phys. Earth planet. Inter.,* **114,** 197–210.

Raitt, R.W., Shor, G.G., Francis, T.J.G. & Morris, G.B., 1969. Anisotropy of the Pacific upper mantle, *J. geophys. Res.,* **74,** 3095–3109.

Rümpker, G. & Silver, P.G., 1998. Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy, *Geophys. J. Int.,* **135,** 790–800.

Saltzer, R.L., Jordan, T.H., Gaherty, J.B., Zhao, L. & the Kaapvaal Working Group, 1998. Anisotropic structure of the Kaapvaal craton from surface wave analysis, *EOS, Trans. Am. geophys. Un.,* **79,** 45.

Savage, M.K., 1999. Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting?, *Rev. Geophys.,* **37,** 65–106.

Shearer, P.M. & Orcutt, J.A., 1986. Compressional and shear wave anisotropy in the oceanic lithosphere—the Ngendei seismic refraction experiment, *Geophys. J. R. astr. Soc.,* **87,** 967–1003.

Silver, P.G., 1996. Seismic anisotropy beneath the continents: probing the depths of geology, *Ann. Rev. Earth. planet. Sci.,* **24,** 385–432.

Silver, P.G. & Chan, W.W., 1991. Shear wave splitting and mantle deformation, *J. geophys. Res.,* **96,** 16 429–16 454.

Silver, P.G. & Savage, M.K., 1994. The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers, *Geophys. J. Int.,* **119,** 949–963.

Tanimoto, T. & Anderson, D.L., 1985. Lateral heterogeneity and azimuthal anisotropy of the upper mantle: Love and Rayleigh waves 100–250 S, *J. geophys. Res.,* **90,** 1842–1858.

Varshalovich, D.A. & Moskalev, A.N., 1988. *Quantum Theory of Angular Momentum,* ed. Kheronski, V.K., World Scientific.

Vidale, J.E., 1986. Complex plarzation analysis of particle motin, *Bull. seism. Soc. Am.,* **76,** 1393–1405.

Wessel, P. & Smith, W.H.F., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.,* **72,** 441–446.

Wolfe, C.J. & Solomon, S.C., 1998. Shear-wave splitting and implications for mantle flow beneath the MELT region of the East Pacific Rise, *Science,* **280,** 1230–1232.

APPENDIX A: PROPAGATOR MATRICES AND SPLITTING OPERATORS

The system matrix $\mathscr{A}(z)$ is rotated into the eigenwave coordinate system by the 4×4 block-diagonal matrix $\mathscr{U} = \mathrm{diag}[\mathbf{U}, \mathbf{U}]$,

$$\hat{\mathscr{A}} = \mathscr{U}\mathscr{A}\mathscr{U}^{\mathrm{T}} = \begin{bmatrix} \mathbf{0} & \mathbf{C}^{-1} \\ -\omega^2 \mathbf{I} & \mathbf{0} \end{bmatrix}, \tag{A1}$$

and is diagonalized by a matrix whose columns represent the downgoing and upgoing eigenwaves: $\mathscr{D}^{-1}\hat{\mathscr{A}}\mathscr{D} = \mathrm{i}\,\mathrm{diag}[k_1, k_2, -k_1, -k_2]$. Here, $k_j = \omega/v_j$ $(j = 1, 2)$ are the eigenwavenumbers, and

$$\mathscr{D} = \begin{bmatrix} \mathbf{D}_u & \mathbf{D}_u \\ \mathbf{D}_\tau & \mathbf{D}_\tau^* \end{bmatrix}, \qquad \mathscr{D}^{-1} = \begin{bmatrix} \mathbf{D}_u^{-1} & \mathbf{D}_\tau^{-1} \\ \mathbf{D}_u^{-1} & \mathbf{D}_\tau^{*-1} \end{bmatrix}, \tag{A2}$$

where the 2×2 blocks are

$$\mathbf{D}_u = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix}, \qquad \mathbf{D}_\tau = \begin{bmatrix} \mathrm{i}\omega v_1 \varepsilon_1 & 0 \\ 0 & \mathrm{i}\omega v_2 \varepsilon_2 \end{bmatrix}. \tag{A3}$$

Kennett's (1983, eq. 2.63) energy normalization procedure yields $\varepsilon_j = (2v_j)^{-1/2}$.

In a homogeneous layer $[\phi(z) = \mathrm{constant}]$, the eigenwave propagator from z_0 to z is

$$\mathscr{E}(z, z_0) = \begin{bmatrix} \mathbf{E}(z, z_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^*(z, z_0) \end{bmatrix}, \tag{A4}$$

where * indicates complex conjugation and

$$\mathbf{E}(z, z_0) = \begin{bmatrix} \exp[\mathrm{i}k_1(z - z_0)] & 0 \\ 0 & \exp[\mathrm{i}k_2(z - z_0)] \end{bmatrix}. \tag{A5}$$

Therefore, the displacement–stress propagator in the eigenwave coordinate system is $\hat{\mathscr{P}}^{\mathrm{h}}(z, z_0) = \mathscr{D}\mathscr{E}(z, z_0)\mathscr{D}^{-1}$, and the general homogeneous-layer propagator matrix can be written as

$$\mathscr{P}^{\mathrm{h}}(z, z_0) \equiv \begin{bmatrix} \mathbf{P}_{uu}^{\mathrm{h}}(z, z_0) & \mathbf{P}_{u\tau}^{\mathrm{h}}(z, z_0) \\ \mathbf{P}_{\tau u}^{\mathrm{h}}(z, z_0) & \mathbf{P}_{\tau\tau}^{\mathrm{h}}(z, z_0) \end{bmatrix} = \mathscr{U}(\phi)\hat{\mathscr{P}}^{\mathrm{h}}(z, z_0)\mathscr{U}^{\mathrm{T}}(\phi). \tag{A6}$$

The propagator for the anisotropic, heterogeneous interval $0 \leq z \leq d$ can be constructed by approximating the heterogeneity as a stack of homogeneous layers and multiplying propagators in the form of (A6). Better analytical insight into the effects of the heterogeneity can be obtained by proceeding in a different fashion, however. We make the ansatz (cf. Kennett 1983, p. 53),

$$\mathscr{P}(z, z') = \mathscr{U}(\phi(z))\mathscr{D}\mathscr{E}(z, z_0)\mathscr{Q}(z, z_0)\mathscr{D}^{-1}\mathscr{U}^{\mathrm{T}}(\phi(z_0)), \tag{A7}$$

and find that (A7) obeys the system equation $\partial_z \mathscr{P} = \mathscr{A}\mathscr{P}$ if and only if $\mathscr{Q}$ satisfies

$$\partial_z \mathscr{Q}(z, z_0) = \mathscr{S}(z)\mathscr{Q}(z, z_0), \tag{A8a}$$

$$\mathscr{Q}(z_0, z_0) = \mathbf{I}, \tag{A8b}$$

where

$$\mathscr{S}(z) = \begin{bmatrix} 0 & a(z) & 0 & b^*(z) \\ -a^*(z) & 0 & b^*(z) & 0 \\ 0 & b(z) & 0 & a^*(z) \\ b(z) & 0 & -a(z) & 0 \end{bmatrix} \frac{d\phi}{dz}. \tag{A9}$$

The depth-dependent coefficients a and b depend exponentially on the wavenumber $\Delta k = k_2 - k_1$ and the wavenumber average $\bar{k} = (k_1 + k_2)/2$, respectively:

$$a(z) = \frac{\bar{v}}{\sqrt{v_1 v_2}} e^{\mathrm{i}\Delta k z}, \tag{A10a}$$

$$b(z) = \frac{\Delta v}{2\sqrt{v_1 v_2}} e^{2\mathrm{i}\bar{k} z}. \tag{A10b}$$

The propagator $\mathscr{Q}(z, z_0)$ is a 4×4 differential scattering matrix for the eigenwaves, which we write in block form as

$$\mathscr{Q} = \begin{bmatrix} \mathbf{Q}_{++} & \mathbf{Q}_{+-} \\ \mathbf{Q}_{-+} & \mathbf{Q}_{--} \end{bmatrix}. \tag{A11}$$

The submatrices $\mathbf{Q}_{++}$ and $\mathbf{Q}_{+-}$ describe, respectively, the forward scattering of downgoing (+) and upgoing (−) eigenwaves by gradients in $\phi(z)$, and $\mathbf{Q}_{-+}$ and $\mathbf{Q}_{+-}$ describe the corresponding backward scattering. These scattering operators

satisfy the reciprocal relations

$$\mathbf{Q}_{--} = \mathbf{Q}_{++}^{*}, \tag{A12a}$$

$$\mathbf{Q}_{+-} = \mathbf{Q}_{-+}^{*}. \tag{A12b}$$

From eqs (A8) to (A10) we can see that the strength of forward scattering in a depth increment dz is proportional to $a_0 \exp(i\Delta k z)d\phi(z)$, while the strength of the back scattering goes like $b_0 \exp(2i\bar{k}z)d\phi(z)$. If the relative difference in the eigenvelocities is small, then forward scattering will tend to dominate because $b_0 \ll a_0 \sim 1$. Moreover, in this situation of small anisotropy, $\Delta k \ll 2\bar{k}$, so that the back-scattering kernel will be more oscillatory and its integral contributions will tend to cancel if $\phi(z)$ is smooth.

At zero frequency, $\mathscr{S} = \mathscr{S}_0$ depends on z only through $\dot{\phi} \equiv d\phi/dz$. Therefore, $\mathscr{S}_0$ commutes with its integral, and the solution to (11) is $\mathscr{Q}(z, z_0) = \exp(\mathscr{S}\Delta\phi)$, where $\Delta\phi = \phi(z) - \phi(z_0)$. Using the fact that

$$\mathscr{S}_0^{2n} = (-1)^n \Delta\phi^{2n} \mathbf{I} \quad \text{and} \quad \mathscr{S}_0^{2n+1} = (-1)^n \Delta\phi^{2n} \mathscr{S}_0,$$

we can sum the exponential series. This yields a good approximation to the propagator across a layer that is thin compared with a wavelength; that is, for $\bar{k}\Delta z \ll 1$,

$$\mathbf{Q}_{++}(z + \Delta z, z) \approx \begin{bmatrix} \cos\Delta\phi & a(z)\sin\Delta\phi \\ -a^*(z)\sin\Delta\phi & \cos\Delta\phi \end{bmatrix}, \tag{A13a}$$

$$\mathbf{Q}_{-+}(z + \Delta z, z) \approx \begin{bmatrix} 0 & b(z)\sin\Delta\phi \\ b(z)\sin\Delta\phi & 0 \end{bmatrix}. \tag{A13b}$$

We note that this approximation is independent of the form of $\phi(z)$ and, for example, does not require $\phi(z)$ to be a smooth function of depth. Indeed, it provides the generalization of the propagator to self-affine (fractal) media for which $\dot{\phi}$ may not be well defined. Eqs (A7) and (A13) are the basis of our computational algorithm.

For an upgoing wave $\mathbf{u}_I(z) \sim \exp(-i\bar{k}z)$ incident at the base of the anisotropic layer, the displacement–stress vector in the half-space can be expressed as

$$f'(d) = \begin{bmatrix} (\mathbf{I} + \mathbf{R}) \\ i\omega\bar{v}(\mathbf{I} - \mathbf{R}) \end{bmatrix} u_I(d), \tag{A14}$$

where $\mathbf{R}$ is a 2×2 matrix of reflection coefficients. Satisfying the zero-traction boundary conditions at the surface yields

$$\mathbf{R} = \mathbf{U}_d(i\omega\bar{v}\hat{\mathbf{P}}_{\tau\tau} - \hat{\mathbf{P}}_{\tau u})^{-1}(i\omega\bar{v}\hat{\mathbf{P}}_{\tau\tau} + \hat{\mathbf{P}}_{\tau u})\mathbf{U}_d^T. \tag{A15}$$

From here on, $\mathbf{U}_z = \mathbf{U}(\phi(z))$ and it is understood that, unless otherwise specified, the propagators are taken from the base of the anisotropic layer to the surface; for example, $\hat{\mathbf{P}}_{uu} \equiv \mathbf{U}_0^T \mathbf{P}_{uu}(0, d)\mathbf{U}_d$. In this notation, the free-surface displacement vector is

$$\mathbf{u}(0) = \mathbf{U}_0[\hat{\mathbf{P}}_{uu} + i\omega\bar{v}\hat{\mathbf{P}}_{u\tau} + (\hat{\mathbf{P}}_{uu} - i\omega\bar{v}\hat{\mathbf{P}}_{u\tau})\hat{\mathbf{R}}]\mathbf{U}_d^T \mathbf{u}_I'(d). \tag{A16}$$

The symmetries in (A2), (A4) and (A12) can be used to express the propagator submatrices as the following (real-valued) expressions:

$$\hat{\mathbf{P}}_{uu} = \mathbf{D}_u \mathscr{R}e[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_u^{-1}, \tag{A17a}$$

$$\hat{\mathbf{P}}_{u\tau} = i\mathbf{D}_u \mathscr{I}m[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_\tau^{-1}, \tag{A17b}$$

$$\hat{\mathbf{P}}_{\tau u} = i\mathbf{D}_\tau \mathscr{I}m[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_u^{-1}, \tag{A17c}$$

$$\hat{\mathbf{P}}_{\tau\tau} = \mathbf{D}_\tau \mathscr{R}e[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_\tau^{-1}. \tag{A17d}$$

In the case of a homogeneous layer, these expressions reduce to $\hat{\mathbf{P}}_{uu}^h = \hat{\mathbf{P}}_{\tau\tau}^h = \mathrm{diag}[\cos k_j(z - z_0)]$, $\hat{\mathbf{P}}_{u\tau}^h = \mathrm{diag}[(\omega v_j)^{-1}\sin k_j d]$, and $\hat{\mathbf{P}}_{\tau u}^h = \mathrm{diag}[-(\omega v_j)\sin k_j d]$.

A useful approximation, almost always employed in vertical shear wave splitting analysis, is to ignore back-scattering and reverberations within the anisotropic layer. This amounts to ignoring terms of order $\Delta v/\bar{v}$ in eq. (A17). Under this approximation, $a_0 = 1$ and $b_0 = 0$, so that $\mathbf{Q}_{+-} = \mathbf{Q}_{-+} = 0$. A little algebra obtains

$$\left.\begin{array}{l} \mathbf{R} = \mathbf{U}_d \mathbf{Q}_{++}^T \mathbf{E}^2 \mathbf{Q}_{++} \mathbf{U}_d^T, \\ \mathbf{u}(0) = 2\mathbf{U}_0 \mathbf{E} \mathbf{Q}_{++} \mathbf{U}_d^T \mathbf{u}_I'(d). \end{array}\right\} \text{(no back-scattering)}. \begin{array}{l} \text{(A18a)} \\ \text{(A18b)} \end{array}$$

Eq. (A18b) shows that, when back-scattering can be ignored, the eigenwave propagator is just $\mathbf{EQ}_{++}$ It will be convenient to pull out the phase factor corresponding to the mean traveltime through the layer, $\bar{t} = \bar{k}d/\omega$, and rewrite these expressions in terms of the eigenwave splitting matrix,

$$\mathbf{H} = \exp(-i\omega\bar{t})\mathbf{E} = \begin{bmatrix} \exp(-i\omega\Delta t/2) & 0 \\ 0 & \exp(i\omega\Delta t/2) \end{bmatrix}, \tag{A19}$$

which is unimodular; that is, $\det[\mathbf{H}] = 1$. We define the splitting operator,

$$\mathbf{\Gamma} = \mathbf{U}_0 \mathbf{H} \mathbf{Q}_{++} \mathbf{U}_d^T. \tag{A20}$$

The surface displacement is thus $\mathbf{u}(0) = 2\exp(i\omega\bar{t})\mathbf{\Gamma}\mathbf{u}_I'(d)$, and the reflection matrix is $\mathbf{R} = \exp(2i\omega\bar{t})\mathbf{\Gamma}^T\mathbf{\Gamma}$. The factor of two in the former comes from the constructive interference of the upward-going wave and its surface reflection. In the case of a homogeneous layer, $\mathbf{Q}_{++} = \mathbf{I}$, and (A20) becomes

$$\mathbf{\Gamma}_h(\phi, \Delta t) = \mathbf{U}(\phi)\mathbf{H}(\Delta t)\mathbf{U}^T(\phi) \text{(homogeneous layer)}. \tag{A21}$$

All of the matrices in (A20) are both unitary and unimodular; for example, $\mathbf{\Gamma}^{-1} = \mathbf{\Gamma}^\dagger \equiv (\mathbf{\Gamma}^*)^T$, $\det[\mathbf{\Gamma}] = 1$. (The unimodularity of $\mathbf{Q}_{++}$ follows from $\mathrm{tr}[\mathscr{S}_{++}(z)] = 0$; see Kennett 1983, p. 42.) Therefore, all of the matrix operations associated with forward-scattering belong to the group SU(2). This symmetry can be used to simplify the analysis. Any member of this group can be written in it terms of two complex numbers,

$$\begin{bmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{bmatrix}, \quad \text{where} \quad |\alpha|^2 + |\beta|^2 = 1. \tag{A22}$$

An SU(2) matrix thus depends on three real parameters and can be written in the following general forms (Varshalovich & Moskalev 1988):

$$\mathbf{\Gamma} = \begin{bmatrix} \exp[-i(\alpha+\gamma)/2]\cos\beta/2 & \exp[i(\alpha+\gamma)/2]\sin\beta/2 \\ \exp[-i(\alpha+\gamma)/2]\sin\beta/2 & \exp[i(\alpha+\gamma)/2]\cos\beta/2 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\Omega/2 - i\cos\Theta\sin\Omega/2 & -i\exp(i\Phi)\sin\Theta\sin\Omega/2 \\ -i\exp(-i\Phi)\sin\Theta\sin\Omega/2 & \cos\Omega/1 + i\cos\Theta\sin\Omega/2 \end{bmatrix}. \tag{A23}$$

SU(2) is homomorphic (with a two-fold ambiguity) to $\mathrm{O}^+(3)$, the group of proper orthogonal transformations in 3-space,

which allows the splitting operations to be visualized as 3-D rotations. In the first form in (A23), the parameters (α, β, γ) correspond to the Euler angles of the 3-D rotation; in the second, the 3-D rotation is through an angle Ω about an axis with polar coordinates (Θ, Φ). Eq. (A21) can be recast as

$$\mathbf{\Gamma}_{\mathrm{h}} = \begin{bmatrix} \cos\omega\Delta t/2 - \mathrm{i}\cos 2\phi \sin\omega\Delta t/2 & -\mathrm{i}\sin 2\phi \sin\omega\Delta t/2 \\ -\mathrm{i}\sin 2\phi \sin\omega\Delta t/2 & \cos\omega\Delta t/2 + \mathrm{i}\cos 2\phi \sin\omega\Delta t/2 \end{bmatrix}. \tag{A24}$$

The splitting matrix for a homogeneous layer thus corresponds to a 3-D rotation through an angle $\omega\Delta t$ about an axis located at colatitude 2ϕ and zero longitude

APPENDIX B: FRÉCHET KERNELS FOR A HOMOGENEOUS STARTING MODEL

To calculate the Fréchet kernels defined by (8) and (9), we perturb the splitting orientation function $\phi(z)$ by a small constant amount $\delta\phi$ in a thin layer of thickness δz at a depth $0 < z < d$ and compute the corresponding perturbations $\delta\tilde{\phi}(z)$ and $\delta\tilde{t}(z)$. The kernels $G_\phi(z)$ and $G_t(z)$ are then given as the limiting values of the ratios $\delta\tilde{\phi}(z)/\delta\phi\delta z$ and $\delta\tilde{t}(z)/\delta\phi\delta z$, respectively. This calculation can be done numerically for arbitrary starting models and pulse shapes (e.g. Fig. 7). Here we present an analytical derivation for narrow-band pulses in the special case of a homogeneous starting model.

Because the perturbations are small, the forward-scattering approximation applies. Vertical propagation through a homogeneous anisotropic layer with a fast-axis orientation ϕ_0 and splitting time Δt yields the Fourier-transformed vertical displacement $\mathbf{u}(0, \omega) = 2\exp(\mathrm{i}\omega\bar{t})\mathbf{\Gamma}_{\mathrm{h}}(\phi_0, \Delta t)\mathbf{u}_I(\omega)$, where the homogeneous-layer splitting operator $\mathbf{\Gamma}_{\mathrm{h}}$ is given by eq. (A24). We assume that the incident pulse is radially polarized, $\mathbf{u}_I(\omega) = u_I(\omega)\hat{\mathbf{x}} = [u_I(\omega)\, 0]^{\mathrm{T}}$, and we approximate its energy spectrum by a Gaussian:

$$|u_I(\omega)|^2 = \frac{1}{\sigma\sqrt{8\pi}} \exp[-(\omega - \omega_0)^2/2\sigma^2]$$
$$+ \frac{1}{\sigma\sqrt{8\pi}} \exp[-(\omega + \omega_0)^2/2\sigma^2]. \tag{B1}$$

This spectrum has peaks of half-bandwidth σ centred at frequencies of $\pm\omega_0$, and it is normalized to have unit total energy: $\int_{-\infty}^{\infty} |u_I(\omega)|^2 d\omega = 1$. If the pulse is narrow band in the sense that $\sigma \ll \omega_0$, then the integral of its energy spectrum against any reasonably smooth function can be approximated by integrating a truncated Taylor expansion of the function about the centre frequency ω_0:

$$2\int_0^{\infty} f(\omega)|u_I(\omega)|^2 d\omega$$
$$\approx 2\int_0^{\infty} [f(\omega_0) + \omega\dot{f}(\omega_0) + \omega^2\ddot{f}(\omega_0)]|u_I(\omega)|^2 d\omega$$
$$\approx f(\omega_0) + \sigma^2\ddot{f}(\omega_0). \tag{B2}$$

The terms dropped in this approximation are of order $(\sigma/\omega_0)^4$.

Since the perturbed model is homogeneous above z and below $z + \delta z$, its splitting operator can be written in the form

$$\mathbf{\Gamma} = \mathbf{U}(\phi_0)\mathbf{H}(\Delta t)\mathbf{Q}_{++}\mathbf{U}^{\mathrm{T}}(\phi_0), \tag{B3}$$

where $\mathbf{H}$ is given by (A19). For a constant perturbation $\delta\phi$ in a layer $(z, z + \delta z)$, $\mathbf{Q}$ satisfies (A8) with $a_0 = 1$, $b_0 = 0$, and

$$\dot{\phi}(\zeta) = \delta\phi[\delta(\zeta - z) - \delta(\zeta - z - \delta z)]. \tag{B4}$$

Integrating up from the base of the layer across the discontinuities yields an expression for $\mathbf{Q}_{++}$ that is the product of two matrices in the form of (A13a), one for an azimuthal change of $\delta\phi$ at $z + \delta z$, and one for a change of $-\delta\phi$ at z. Multiplying these out and using the small-angle approximations, we can express the forward-scattering matrix in terms of a perturbation parameter $\delta X \equiv \Delta k\, \delta z\, \delta\phi \exp[\mathrm{i}\Delta k(d - z)]$:

$$\mathbf{Q}_{++} = \begin{bmatrix} 1 & -\mathrm{i}\delta X \\ -\mathrm{i}\delta X^* & 1 \end{bmatrix}. \tag{B5}$$

The apparent splitting parameters minimize the energy on the transverse component of the back-projected displacement field, given by the quadratic form (7). In the present notation, this integral becomes

$$\varepsilon^2(\phi', \Delta t') = 2\int_0^{\infty} |\hat{\mathbf{y}}^{\mathrm{T}}\mathbf{\Gamma}_{\mathrm{h}}^{-1}(\phi', \Delta t')\mathbf{\Gamma}\hat{\mathbf{x}}|^2 |u_I(\omega)|^2 d\omega. \tag{B6}$$

The homogeneous-layer splitting operator in this expression corresponds to the perturbed apparent splitting parameters $\phi' = \phi_0 + \delta\phi'$ and $\Delta t' = \Delta t + \delta t'$, which can be expressed in a form similar to (B3):

$$\mathbf{\Gamma}_{\mathrm{h}}(\phi', \Delta t') = \mathbf{U}(\phi_0)\mathbf{H}(\Delta t)\mathbf{Q}'_{++}\mathbf{U}^{\mathrm{T}}(\phi_0). \tag{B7}$$

Equating (B6) with $\mathbf{\Gamma}_{\mathrm{h}}(\phi', \Delta t) = \mathbf{U}(\phi')\mathbf{H}(\Delta t')\mathbf{U}^{\mathrm{T}}(\phi')$ yields the scattering matrix

$$\mathbf{Q}'_{++} = \begin{bmatrix} c - \mathrm{i}s\cos 2\delta\phi' & -\mathrm{i}s\sin 2\delta\phi' \\ -\mathrm{i}s^*\sin 2\delta\phi' & c^* + \mathrm{i}s^*\cos 2\delta\phi' \end{bmatrix}, \tag{B8}$$

with complex coefficients,

$$c = \exp(\mathrm{i}\omega\Delta t/2)\cos[\omega(\Delta t + \delta t')/2], \tag{B9a}$$

$$s = \exp(\mathrm{i}\omega\Delta t/2)\sin[\omega(\Delta t + \delta t')/2]. \tag{B9b}$$

Eq. (B8) is exact and does not require the perturbations $\delta\phi'$ and $\delta t'$ to be small. When, $\phi_0 = \Delta t = 0$, for example, $c = \cos\omega\delta t'/2$, $s = \sin\omega\delta t'/2$, and (B7) reduces to $\mathbf{\Gamma}_{\mathrm{h}}(\delta\phi', \delta t')$ in the form given by (A24).

From (B3) and (B6) and the fact that $\mathbf{Q}'_{++}$ is unitary, we obtain $\mathbf{\Gamma}_{\mathrm{h}}^{-1}(\delta\phi', \delta t')\mathbf{\Gamma} = \mathbf{U}(\phi_0)\mathbf{Q}'^{\dagger}_{++}\mathbf{Q}_{++}\mathbf{U}^{\mathrm{T}}(\phi_0)$. The energy (B6) involves an integration over the (2,1) component of this matrix. The product $\mathbf{Q}'^{\dagger}_{++}\mathbf{Q}_{++}$ can be expressed in the general SU(2) form (A22), where

$$\alpha = c^* + s^*(\delta X^* \sin 2\delta\phi' + \mathrm{i}\cos 2\delta\phi'),$$
$$\beta = -\mathrm{i}\delta Xc^* + s^*(\delta X \cos 2\delta\phi' + \mathrm{i}\sin 2\delta\phi'). \tag{B10}$$

Working out the appropriate matrix element in terms of the real and imaginary parts of these coefficients, we find

$$\varepsilon^2(\delta\phi', \delta t') = 2 \int_0^\infty \{ \mathscr{Re}(\beta)^2 + [\mathscr{Im}(\beta) \cos 2\phi_0 + \mathscr{Re}(\alpha) \sin 2\phi_0]^2 \}$$

$$\times |u_I(\omega)|^2 d\omega. \tag{B11}$$

To find the energy minimum, we differentiate (B11) with respect to the perturbations $\delta\phi'$ and $\delta t'$ and set the results equal to zero, which gives two equations for the apparent splitting parameters. Linearizing these equations in $\delta\tilde{\phi}$ and $\delta\tilde{t}$, we obtain a 2×2 system for the Fréchet kernels:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} G_\phi(z) \\ G_t(z) \end{bmatrix} = \Delta k \begin{bmatrix} D_1(z) \\ D_2(z) \end{bmatrix}, \tag{B12}$$

$$A_{11} = \sin^2 2\phi_0 \int_0^\infty (\omega/\omega_0) |u_I(\omega)|^2 d\omega, \tag{B13a}$$

$$A_{12} = \sin 2\phi_0 \cos 2\phi_0 \int_0^\infty \sin \omega\Delta t |u_I(\omega)|^2 d\omega, \tag{B13b}$$

$$A_{21} = \sin 2\phi_0 \cos 2\phi_0 \int_0^\infty (\omega/\omega_0) \sin \omega\Delta t |u_I(\omega)|^2 d\omega, \tag{B13c}$$

$$A_{22} = \int_0^\infty (1 - \cos \omega\Delta t)^2 |u_I(\omega)|^2 d\omega$$

$$+ \cos^2 2\phi_0 \int_0^\infty \sin^2 \omega\Delta t |u_I(\omega)|^2 d\omega, \tag{B13d}$$

$$D_1(z) = \sin 2\phi_0 \cos 2\phi_0 \int_0^\infty (\omega/\omega_0) \cos \Delta k(d-z) |u_I(\omega)|^2 d\omega, \tag{B14a}$$

$$D_2(z) = \int_0^\infty (1 - \cos \omega\Delta t) \sin \Delta k(d-z) |u_I(\omega)|^2 d\omega$$

$$+ \cos^2 2\phi_0 \int_0^\infty (\omega/\omega_0) \sin \Delta k(d-z) \cos \omega\Delta t' |u_I(\omega)|^2 d\omega. \tag{B14b}$$

Approximating these integrals using (B2) and solving for the kernels leads to the expressions (10) to (14) given in the text.