Geophysical Journal International

Geophys. J. Int. (2023) **235,** 2640–2670
Advance Access publication 2023 October 04
GJI Seismology

https://doi.org/10.1093/gji/ggad389

Imaging upper mantle anisotropy with traveltime and splitting intensity observations from teleseismic shear waves: insights from tomographic reconstructions of subduction simulations

Brandon P. VanderBeek, Rosalia Lo Bue, Francesco Rappisi and Manuele Faccenda

Dipartimento di Geoscienze, Università degli Studi di Padova, 35131 *Padova PD, Italy. E-mail: brandonpaul.vanderbeek@unipd.it*

Accepted 2023 September 30. Received 2023 September 25; in original form 2022 December 8

SUMMARY

Teleseismic traveltime tomography remains one of the most popular methods for obtaining images of Earth's upper mantle. However, despite extensive evidence for an elastically anisotropic mantle, the isotropic assumption remains commonplace in such imaging studies. This can result in significant model artefacts which in turn may yield misguided inferences regarding mantle dynamics. The nature of anisotropy-induced apparent velocity anomalies has been well-documented in *P*-wave imaging and various strategies have been proposed to constrain both isotropic and anisotropic heterogeneity from these data. In contrast, few studies have explored the consequences for shear wave tomography and no practical framework for the anisotropic inversion of *S*-wave delays exists. Here, we propose a new method for constraining arbitrarily oriented hexagonal anisotropy using both traveltime and splitting intensity observations from direct *S* phases. Our approach accounts for polarization and finite-frequency effects and allows for isotropic starting models. The imaging method is validated through the tomographic analysis of a realistic synthetic dataset produced from waveform simulations through a geodynamic model of subduction. Results illustrate that neglecting anisotropy produces distortions in slab geometry and the appearance of sub- and supraslab low-velocity zones. Anisotropic inversions remove these artefacts while also constraining geodynamically relevant fabric properties including dip.

Key words: Composition and structure of the mantle; Seismic anisotropy; Seismic tomography; Shear body waves.

1 INTRODUCTION

Constraining the anisotropic structure of Earth's interior remains a challenge in seismology and is crucial for understanding mantle dynamics. By coupling micromechanical deformation models with petrologic observations (e.g. Nicolas & Christensen 1987; Kaminski *et al.* 2004; Boneh *et al.* 2015; Wenk 2016), geodynamic simulations can yield detailed predictions for the anisotropic elastic structure of the upper mantle from global (e.g. Becker 2008; Becker *et al.* 2008) to local (e.g. Blackman & Kendall 2002; Faccenda 2014) scales. The anisotropic properties record the convective history of the mantle and the resulting mineral fabrics reflect a number of important properties including temperature, pressure and deviatoric stress conditions, mineral assemblages, presence of partial melt and volatile content (Karato *et al.* 2008; Hansen *et al.* 2021). Moreover, elastic anisotropy strongly effects wave propagation by causing direction-dependent variations in seismic velocity that are comparable to those due to thermal or composition changes (Anderson 2007). Thus, one can study mantle dynamics by exploiting the anisotropic properties of the seismic wavefield (Simmons *et al.* 2006; Long & Becker 2010; Wang & Becker 2019; Becker & Lebedev 2021).

Anisotropy has long been considered in surface wave tomography (Montagner & Nataf 1986; Montagner 2007) revealing large-scale patterns in mantle flow (e.g. Zhou *et al.* 2006; Long & Becker 2010; Becker *et al.* 2014). Shear wave splitting (SWS) is routinely analysed to constrain lateral variations in depth-averaged anisotropic structure at the scale of instrument spacing within an array (Savage 1999; Long & Silver 2009). However, it remains commonplace to assume an isotropic Earth in body wave tomography particularly with teleseismic data. This assumption is problematic for a number of reasons. (i) It is inconsistent with the aforementioned surface wave and SWS observations which are often derived from seismic data collected in the same geographic area, if not from the same array, studied for body wave imaging. This makes meaningful comparisons of models and integration of multiple seismic observables difficult. (ii) Body wave arrivals are strongly affected by anisotropy (Kendall 1994; Levin *et al.* 1996; Wu & Lees 1999; Blackman &

Kendall 2002; Sieminski *et al.* 2007) and neglecting this effect on apparent wave speeds could lead to significant imaging artefacts as is well-documented for *P* waves (Sobolev *et al.* 1999; Menke 2015; Bezada *et al.* 2016; VanderBeek & Faccenda 2021; Lo Bue *et al.* 2022; Rappisi *et al.* 2022). (iii) Consequently, interpretation of isotropic velocity models could lead to misguided inferences on mantle dynamics. (iv) Finally, body waves are particularly useful for constraining elastic anisotropy owing to their high-frequency content and diverse directional sampling of the mantle (Beller & Chevrot 2020). While surface waves are well-suited for constraining depth-dependent structure, the long periods required for mantle imaging result in poor lateral resolution. Shear wave splitting studies primarily focus on core-converted *S* phases (i.e. XK(K)S) because their polarization is known and the contribution of source-side anisotropy is removed after passing though the outer core. However, their near-vertical incidence angles within the upper mantle result in poor depth resolution. The horizontal and vertical propagation paths of surface and XK(K)S waves, respectively, provide poor constraints on the dip of anisotropic fabrics (Chevrot & van der Hilst 2003; Romanowicz & Yuan 2012; Mondal & Long 2019) which are relevant for evaluating geodyanmic predictions. Clearly, extending the study of anisotropy to a broader range of body wave phases would resolve these limitations.

Several researchers have investigated strategies for incorporating hexagonal anisotropy (e.g. Hearn 1996; Eberhart-Phillips & Henderson 2004; Munzarová *et al.* 2018; VanderBeek & Faccenda 2021; Wang & Zhao 2021) and, less commonly, lower-order elastic symmetry (Mochizuki 1995; Grésillaud & Cara 1996; Wu & Lees 1999; Koulakov *et al.* 2009; Liu & Zhao 2016) in the inversion of *P*-wave traveltimes. Recently, (VanderBeek & Faccenda 2021) demonstrated the importance of solving for anisotropy, especially fabric dip, in the construction of teleseismic *P*-wave models in subduction environments. However, the consequences of assuming isotropy in shear wave traveltime tomography is poorly documented and a practical ray-based framework for the anisotropic inversion of such data does not exist for passive-source mantle imaging. Here we address these issues using a realistic synthetic seismic dataset produced by propagating teleseismic wavefields though an elastically anisotropic geodynamic model of subduction. After reviewing prior work on retrieving anisotropy from *S* waves (Section 2), we derive approximate analytic expressions for the propagation velocity and splitting intensity of a shear wave travelling through a hexagonally anisotropic media with arbitrary symmetry axis orientation (Section 3). From these expression, we detail an iterative linearized inversion strategy for imaging the mean shear velocity and the strength and 3-D orientation of anisotropy from these observables (Section 4). Using SPECFEM (Komatitsch & Tromp 1999; Chen & Tromp 2007) generated waveforms, we validate our new tomography method (Sections 5–7). Our results illustrate that traditional isotropic *S*-wave tomography is prone to strong imaging artefacts when anisotropic fabric is present in the imaging target but ignored in the structural inversion. By inverting for anisotropic parameters, our method nearly eliminates these artefacts while also reconstructing geodynamically relevant anisotropic heterogeneity.

2 BACKGROUND: ANISOTROPIC IMAGING WITH SHEAR WAVES

To give context to the present work, we briefly summarize previous efforts at retrieving anisotropic structure from shear waves. The literature on seismic anisotropy is dense and so we focus this review on passive source mantle imaging for which our tomography method was developed. However, we acknowledge the extensive body of research pertaining to crustal seismic anisotropy (e.g. Crampin & Peacock 2008; Crampin & Gao 2009).

Analysis of shear waves for the purpose of constraining mantle anisotropy is almost exclusively restricted to the measurement of shear wave splitting parameters. After passing through an anisotropic medium, a shear wave splits into two orthogonally polarized quasi-shear phases that propagate at different velocities. The polarization direction (usually that of the fast-polarized waveform) and time separation between the two phases define the splitting parameters that provide a measure of the apparent azimuth and strength of the anisotropic medium. Since the early work of Ando *et al.* (1983), Vinnik *et al.* (1989) and Silver & Chan (1988) measurement of SWS parameters has emerged as a popular means of investigating mantle anisotropy, particularly at regional-to-local scales (Savage 1999; Long & Silver 2009). The method is especially robust for constraining anisotropy because the birefringence phenomenon is insensitive to isotropic heterogeneity (Chevrot 2006). However, the observed splitting parameters reflect the accumulated effect of anisotropy as the wave propagates from source to receiver. For this reason, phases recorded at steep incidence angles, typically SK(K)S or intraslab earthquakes beneath nearby seismic arrays, are studied. In this case, splitting measurements represent a depth integrated view and lateral variations in anisotropic structure can be investigated at the scale of the instrument spacing. However, direct interpretation of such integrated measurements can be complicated by the presence of vertical heterogeneity (e.g. Silver & Savage 1994; Silver & Long 2011; Romanowicz & Yuan 2012). In such cases, backazimuthal (or, more precisely, incoming polarization-dependent) variations in splitting parameters can be used to identify layered anisotropy as well as dipping fabrics. Although, such analysis is generally limited to the identification of discrete two-layer models.

As an alternative to the splitting parameters, Chevrot (2000) proposed quantifying anisotropy via the splitting intensity which measures the amplitude of the transverse component waveform as a function of incoming polarization direction. Splitting intensity measurements follow sinusoidal trends with incoming polarization direction that can be used to identify the strength, azimuth and dip of anisotropic fabrics. Unlike the splitting parameters, splitting intensity is commutative (Chevrot 2000; Silver & Long 2011) meaning its value does not depend on the order of anisotropic heterogeneity traversed by the wave. While the commutativity of splitting intensity implies that it cannot be used to diagnose layered anisotropy, this property does make it amenable to linearized inversion methods common in traveltime tomography.

Ultimately, we wish to constrain a 3-D view of anisotropy. To this end, there exists a number of tomographic imaging methods for the inversion of splitting parameters or intensity. Babuška *et al.* (1993) and Šílený & Plomerová (1996) used a grid-search approach to find the orientation of a presumed elastic tensor that best explained *P* wave delays and splitting parameters in selected geographic domains. Later, Ryberg *et al.* (2005) used full-waveform simulations with direct search methods to invert splitting parameters for the size and position of azimuthally anisotropic domains in 2-D. Abt & Fischer (2008) developed an iterative linearized inversion for splitting parameters using local earthquakes in arbitrarily oriented hexagonal and orthorhombic anisotropic media. Shear wave splitting was modelled using a ray based matrix propagation method (Rümpker & Silver 1998; Fischer *et al.* 2000) with a finite difference scheme for

computing parameter sensitivities. In practice, data limitations restricted the method to primarily recovering azimuthally anisotropic parameters in hexagonally symmetric media (Abt *et al.* 2009). To pose a more linear imaging problem, Zhang *et al.* (2007) focused on imaging only the strength of anisotropy while assuming a fixed symmetry axis orientation.

Inversions based on splitting parameters have a number of disadvantages. The non-commutativity of the observations exacerbates the non-linearity of the imaging problem. Consequently, solutions obtained from linearized inversions may strongly depend on the starting model. The splitting parameters themselves cannot be directly calculated but must be measured from synthetic waveforms. While a number of methodologies exist for making these measurements (Vecsey *et al.* 2008; Long & Silver 2009), they can yield different results (e.g. Long & van der Hilst 2005; Wüstefeld & Bokelmann 2007; Vecsey *et al.* 2008) further complicating the tomographic analysis. Finally, frequency-dependent scattering effects cannot easily be approximated and require more expensive waveform simulations.

Anisotropic inversions based on splitting intensities largely avoid these issues. Following the initial conception by Chevrot (2000), a series of subsequent works laid the foundation for performing finite-frequency splitting intensity tomography (Favier & Chevrot 2003; Chevrot & van der Hilst 2003; Chevrot *et al.* 2004; Chevrot 2006; Long & Silver 2008; Chevrot & Monteiller 2009) with focus on SKS phases in hexagonally symmetric elastic media. Sieminski *et al.* (2007) generalized the concept of splitting intensity and traveltime for shear waves propagating in arbitrary anisotropic media. Through computation of adjoint kernels, they demonstrated that these observables are generally sensitive to a large number of the 21 elastic coefficients. Perhaps for this reason, we are not aware of any application of the Sieminski *et al.* (2007) method to shear wave tomography.

The major limitation of splitting intensity for imaging anisotropic structure is the requirement that the delay time between fast- and slow-polarized waves remains small relative to the dominant period of the waveform. This requirement is not prohibitively restrictive for mantle imaging where periods of $\sim$10 s are generally analysed while delays rarely exceed 2 s. Iterative linearized inversion schemes are generally parametrized using an anisotropic strength parameter(s) and the azimuth and, to a lesser extent, the dip of a hexagonal symmetry axis. Although a perturbation to the splitting intensity can be linearly related to the anisotropic perturbations along the ray path or integral over the finite-frequency kernel, the imaging problem is still non-linear in that parameter sensitivities depend on the current model. Recently, Mondal & Long (2019) implemented a model space search approach for estimating 3-D azimuthal hexagonal anisotropy structure. Such methods are better suited for non-linear minimization but become computationally expensive for high-dimensional problems. Application of splitting intensity tomography has largely been restricted to observations from SK(K)S phases (Monteiller & Chevrot 2011; Lin *et al.* 2014; Mondal & Long 2020; Confal *et al.* 2023). While the finite-frequency kernels offer some depth sensitivity, the near-vertical propagation through the imaging target ultimately limits vertical resolution and offers poor constraints on fabric dip (Chevrot & van der Hilst 2003; Romanowicz & Yuan 2012). Additionally, splitting intensity tomography requires good back-azimuthal coverage which can be difficult to achieve when only SK(K)S phases are considered (Chevrot 2000).

Despite the attention given to shear wave splitting analysis and the success of such methods in revealing the widespread anisotropic nature of the upper mantle (e.g. Savage 1999; Long & Silver 2009; Long & Becker 2010; Long 2013), tomographic inversion of shear wave arrivals has almost exclusively been carried out assuming an isotropic earth. This is likely due to a number of reasons. For example, it is unclear how to best measure and predict shear wave traveltimes when two quasi-shear phases are recorded. Anisotropic effects are polarization-dependent and polarizations are not routinely measured in tomographic studies. There is likely some reluctance to include more parameters in an already underdetermined inverse problem. However, this is necessary to understand potential isotropic–anisotropic structure trade-offs that are relevant for model interpretation. Additionally, for many imaging targets the isotropic assumption is difficult to justify as there are often independent observations supporting the presence of anisotropy (e.g. geodynamic predictions, mineral physics arguments, tectonic fabrics, shear wave splitting observations and surface wave anisotropy).

Ray-theoretical expressions for shear velocities in anisotropic media are well-established (e.g. Chapman & Shearer 1989; Chapman & Pratt 1992). However, these assume that both quasi-shear phases can be identified and their arrival times measured. This is not generally true due to polarization and frequency-dependent propagation effects. Consequently, there have been relatively few attempts at incorporating anisotropy in teleseismic *S*-wave tomography. Some examples include Lloyd & van der Lee (2008) who explored the potential bias anisotropy could introduce into isotropic *S*-wave velocity models. While they argue that anisotropy-induced artefacts are likely small, they did not consider depth-dependent or dipping anisotropy and neglected the effect of shear wave polarization on apparent velocities. Wu & Lees (1999) developed a tensoral representation for the velocity at which the first arriving *S*-wave travels. However, this theoretical *S*-arrival may not exist depending on the initial polarization direction and constraining this velocity surface requires solving for 12 independent variables. Hammond & Toomey (2003) jointly inverted teleseismic *P* and *S* delays with SKS splitting observations measured across the East Pacific Rise. By measuring *S* delay times in the slow polarization directions inferred from SKS splitting parameters, they were able to approximate shear velocities using the same functional form as for *P* waves; this implicitly assumes propagation through a single anisotropic layer. They constrained the geometry, strength, and orientation of three 2-D anisotropic domains via grid-search combined with linearized inversions for isotropic heterogeneity carried out for each proposed anisotropic model. Following a similar measurement strategy, Eilon *et al.* (2016) inverted delay times of fast- and slow-polarized teleseismic *S* phases together with splitting times for 3-D variations in fast- and slow-shear wave velocity beneath the Woodlark Rift. However, their method requires one to prescribe the orientation of hexagonal azimuthal anisotropy prior to inversion. Liu & Zhao (2016) proposed a method to invert both *P*- and *S*-wave traveltimes for azimuthal anisotropic parameters beneath Japan. They assumed that the directional dependence of shear velocity follows the same sinusoidal trend as for *P* waves with an added dependence on the shear wave polarization. In practice, the polarization of the shear waves in their dataset were not estimated and instead all were assumed to be radially polarized.

Body wave anisotropy has been considered in adjoint waveform tomography. However, imaging is generally restricted to radial anisotropy parametrizations which explicitly assume a vertical symmetry axis (e.g. Fichtner *et al.* 2010; Zhu *et al.* 2017; Rodgers *et al.* 2022). Recently, Zhu *et al.* (2020) applied adjoint waveform inversion for azimuthally anisotropic parameters beneath North America. Owing to the relatively long periods used ($>$20 s), these models are

largely constrained by surface waves and *P*-wave anisotropy was not considered. Incorporating shorter periods at which body waves are commonly observed remains a challenge due to computational costs. However, such data will be important to fully characterize the anisotropic properties of the Earth. This point was illustrated by Beller & Chevrot (2020) who demonstrated with synthetic data the ability of full-waveform inversions (in which both the phase and amplitude of waveforms are fit) applied to teleseismic *P* and *S* waveforms (5–20 s period) to recover 21-component elastic tensors in a regional-scale model. Of course, with imperfect data coverage there is the potential for large trade-offs between model parameters and evaluating uncertainty for non-linear waveform inversion methodologies remains a challenge due to expensive simulation times. For this reason, simplified anisotropic imaging strategies with reduced parameter spaces are beneficial as they allow for greater exploration of the model space (e.g. Bodin & Sambridge 2009; Mondal & Long 2019).

In summary, anisotropic imaging with shear waves has largely been restricted to the interpretation and/or inversion of splitting observations predominantly from XK(K)S phases. However, traveltimes also carry information on anisotropic heterogeneity that when ignored could manifest as isotropic artefacts in tomographic models. Prior attempts to include anisotropy in shear wave imaging generally restricted fabric orientations to the horizontal or vertical plane. Existing methods based on traveltime observations do not appropriately account for polarization effects and the shear velocity approximations used may not be applicable for propagation through anisotropically heterogeneous media. Although FWI is the most promising tool for high-resolution anisotropic imaging, its widespread application remains limited by the availability of computational resources and sufficiently dense high-quality seismic recordings. To address these issues, we develop a teleseismic shear wave imaging method that leverages both finite-frequency traveltime and splitting intensity observations to constrain arbitrarily oriented hexagonal anisotropy. Both observables are obtained following minor modifications to a typical workflow for measuring cross-correlated delay times and accounts for incoming *S*-wave polarization. To describe hexagonal anisotropy with an arbitrary symmetry axis requires seven parameters (five elastic coefficients and two orientation parameters). However, we can accurately predict *S*-wave observables with just four variables that constrain the mean shear velocity, anisotropic strength, and symmetry axis orientation. Our tomographic method is tested using SPECFEM simulated waveforms propagated through a geodynamic model of subduction. Thus, our synthetic dataset is independent of assumptions in our imaging algorithm.

3 SHEAR WAVE OBSERVABLES IN ANISOTROPIC MEDIA

We derive expressions for two scalar observables which we term the principal delay and principal splitting intensity for a shear wave with arbitrary propagation and polarization directions traversing a heterogeneous anisotropic medium with hexagonal symmetry. We show that these quantities are linearly related to the generalized shear wave delay and splitting intensity defined by Sieminski *et al.* (2007). Both observables can be integrated along the seismic ray path and we provide an approximation to account for finite-frequency effects. We conclude with a discussion on the practical limitations of these observables for anisotropic *S*-wave imaging.

3.1 Derivation of principal traveltime and splitting intensity

We consider a linearly polarized shear wave entering an anisotropic medium. Following Silver & Chan (1991), we model shear wave splitting by projecting a waveform onto the two quasi-shear wave polarization directions, qS' and qS'', and time shifting these components according to their velocities. In a coordinate system defined such that $\hat{\mathbf{x}}_1$ parallels the incoming *S*-wave polarization direction and $\hat{\mathbf{x}}_3$ parallels the propagation direction (Fig. 1a), the split waveform is given by,

$$\mathbf{U}(\omega) = \mathbf{R}^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{R}\mathrm{U}_0(\omega)\hat{\mathbf{x}}_1, \tag{1}$$

where $U_0(\omega)$ is the incoming shear waveform in the frequency domain with angular frequency ω; the superscript 'T' denotes the matrix transpose. The matrix $\mathbf{R}$ rotates $x_{1,2}$-coordinates to the qS'- and qS''-polarization directions and is defined as,

$$\mathbf{R} = \begin{bmatrix} \cos(\psi' - \zeta) & \sin(\psi' - \zeta) \\ -\sin(\psi' - \zeta) & \cos(\psi' - \zeta) \end{bmatrix}, \tag{2}$$

where ψ' and ζ are the qS' and initial polarization azimuths, respectively, measured from the Q-channel in the ray-normal QT-plane (Fig. 1a). The splitting operator, $\boldsymbol{\Gamma}$, applies the time delays $\Delta t'$ and $\Delta t''$ to the qS'- and qS''-aligned waveforms,

$$\boldsymbol{\Gamma} = \begin{bmatrix} \exp(-i\omega\Delta t') & 0 \\ 0 & \exp(-i\omega\Delta t'') \end{bmatrix} \tag{3}$$

Two assumptions are implicit in eq. (1). (i) It neglects phase shifts due to interactions with the free surface which is justified considering the shallow incidence angles of teleseismic ray paths (Evens 1984). (ii) Deviations in polarization from the ray-normal QT-plane are neglected which is justified when considering weakly anisotropic media. No assumptions regarding the anisotropic symmetry system have yet been made.

Provided that $\Delta t'$ and $\Delta t''$ are small with respect to ω such that the first-order Taylor expansion, $\exp(-i\omega\Delta t') \approx 1 - i\omega\Delta t'$, is valid, then eqs (1)–(3) yield the following expression for the split waveform components,

$$\mathrm{U}_1(\omega) \approx \left(1 - i\omega[\Delta t' \cos^2(\beta) + \Delta t'' \sin^2(\beta)]\right)\mathrm{U}_0(\omega) \tag{4}$$

and

$$\mathrm{U}_2(\omega) \approx \frac{1}{2}\sin(2\beta)[\Delta t'' - \Delta t']i\omega\mathrm{U}_0(\omega), \tag{5}$$

where β has been substituted for $\psi' - \zeta$. It is clear from eq. (4), that U_1 is simply U_0 subject to the delay,

$$\Delta t = \Delta t'' + (\Delta t' - \Delta t'')\cos^2(\beta). \tag{6}$$

Eq. (5) predicts that the energy scattered to the x_2-direction is the scaled time-derivative of U_0. Following Chevrot (2000), the scaling factor is referred to as the splitting intensity,

$$\Delta s = \frac{1}{2}(\Delta t'' - \Delta t')\sin(2\beta). \tag{7}$$

Eqs (6) and 7 define the principal anisotropic delay and principal splitting intensity referred to at the beginning of this section. Here 'principal' is referring to the fact that these observables are defined in a coordinate system aligned with the the principal components of *S*-wave propagation and displacement. These quantities are linearly related to the generalized delay time (o_t) and splitting intensity (o_s) presented by Sieminski *et al.* (2007) who derived these observables from the S_h and S_v components of an isotropic reference waveform

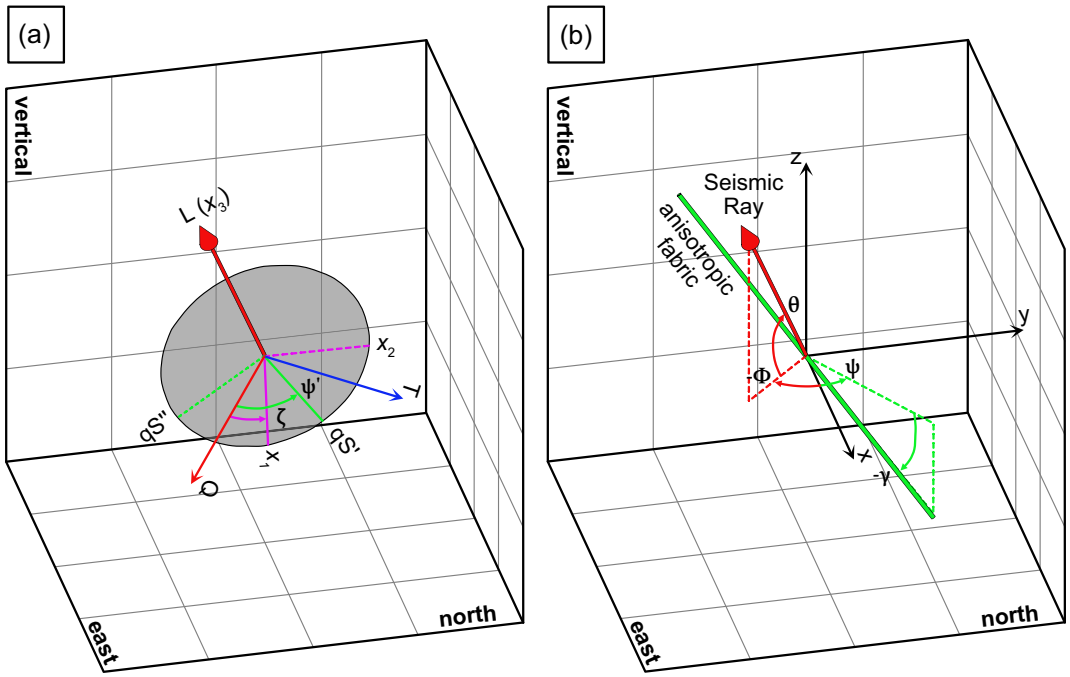

Figure 1. Coordinate systems and parameters used to model shear wave anisotropy. (a) The ray-aligned (QTL) and polarization-aligned ($x_{1,2,3}$) coordinate systems. The red arrow depicts the *S*-wave ray path (*L*-axis). The solid pink bar labelled x_1 shows the incoming shear wave polarization direction, ζ, in the ray-normal QT-plane (grey shaded region); x_2 (dashed pink line) is orthogonal to x_1 and x_3. The solid and dashed green lines show the qS' and qS'' quasi-shear wave polarization directions, respectively. Note that qS' parallels the projection of the symmetry axis (Fig. 1b) into the QT-plane and is defined by the azimuth ψ'. (b) The *S*-wave ray path (red arrow) is shown in relation to the orientation of the anisotropic symmetry axis (green bar). The ray and symmetry axis azimuths (ϕ and ψ) and elevations (θ and γ) are measured in a right-handed coordinate system.

perturbed by an anisotropic scatterer. From Sieminski *et al.* (2007), we may define,

$$o_t = \frac{1}{2}(\Delta t' + \Delta t'')\cos(2\zeta) + (\Delta t' - \Delta t'')\cos(2\psi') \tag{8}$$

and

$$o_s = \frac{1}{2}(\Delta t' + \Delta t'')\sin(2\zeta) + (\Delta t' - \Delta t'')\sin(2\psi'). \tag{9}$$

It is clear that $(\Delta t, \Delta s)$ and (o_t, o_s) are related via the transformation,

$$\begin{bmatrix}\Delta t \\ \Delta s\end{bmatrix} = \begin{bmatrix}\cos(2\zeta) & \sin(2\zeta) \\ \sin(2\zeta) & -\cos(2\zeta)\end{bmatrix}\begin{bmatrix}o_t \\ o_s\end{bmatrix}. \tag{10}$$

While eqs (6) and (7) were derived for a single anisotropic interval, the extension to multiple layers is straight forward. Provided that the splitting intensity remains small such that the *S*-wave remains approximately linearly polarized in the direction ζ, the anisotropic observables may be integrated along the ray path. For a path composed of M-segments these are defined,

$$t = \sum_{j=1}^{M} dL_j \left[u_j'' + \left(u_j' - u_j''\right)\cos^2(\beta_j)\right] \tag{11}$$

and

$$s = \frac{1}{2}\sum_{j=1}^{M} dL_j (u_j'' - u_j')\sin(2\beta_j), \tag{12}$$

where dL_j is the length and u_j' and u_j'' are the slowness values (i.e. inverse of velocity) of the qS'- and qS''-polarized waves through the *j*th ray path segment. Because the splitting intensity depends only on the difference between u'' and u', it's largely insensitive to changes in the isotropic shear wave speed.

We want to relate u' and u'' to the anisotropic properties of the medium. The elastic properties of Earth's mantle are well-approximated by a weakly anisotropic tensor with hexagonal symmetry (e.g. Becker *et al.* 2006). Following Thomsen (1986), the velocity of qS''- and qS'-polarized waves in such a media are well-described by the following sinusoidal functions,

$$u'' = u\left[1 + f''\cos(2\alpha)\right]^{-1} \tag{13}$$

and

$$u' = u\frac{(1+f')}{(1+f'')}\left[1 + f'\cos(4\alpha)\right]^{-1}, \tag{14}$$

where u is the mean slowness; f' and f'' are the anisotropic fractions that define the amplitude of the directional velocity variations of the qS'- and qS''-polarized waves and can be positive or negative depending on the particular cause of anisotropy in the medium; and α is the angle between the seismic ray and the hexagonal symmetry axis (Fig. 1b). Strictly speaking, α should be the phase rather than the ray angle with respect to the symmetry axis. However, these are approximately equivalent in anisotropic media characterized by weak ($\lessapprox$10 per cent) directional velocity variations. The relationships between u, f'' and f' to the original Thomsen parameters are provided in Appendix A.

The angle α is related to the dot-product between the ray and symmetry axis unit vectors,

$$\cos(\alpha) = \left[\cos(\phi - \psi)\cos(\theta)\cos(\gamma) + \sin(\theta)\sin(\gamma)\right]. \tag{15}$$

where ϕ and θ are the azimuth and elevation of the ray path and ψ and γ are the azimuth and elevation of the hexagonal symmetry axis (Fig. 1b). Recognizing that the qS'-polarization, ψ', parallels the projection of the anisotropic symmetry axis into the ray-normal

QT-plane (Fig. 1b), the angle β may be defined as,

$$\beta = \arctan\left[\frac{-\sin(\phi-\psi)\cos(\gamma)}{\cos(\phi-\psi)\sin(\theta)\cos(\gamma)-\cos(\theta)\sin(\gamma)}\right]-\zeta. \quad (16)$$

Eqs (11)–(16) provide a framework for predicting traveltimes and splitting intensities through a heterogeneous and weakly anisotropic media with hexagonal symmetry. Calculation of these observables depends on the spherical parameters describing the ray path (dl, ϕ, θ) and five parameters to describe shear wave hexagonal anisotropy (u, f'', f', ψ, γ).

3.2 Approximating finite-frequency effects

Eqs (11) and (12) define ray-theoretical observables. However, teleseismic imaging studies often rely on traveltimes derived from cross-correlations of band-limited waveforms resulting in frequency-dependent measurements (Marquering *et al.* 1999; Dahlen *et al.* 2000). It is possible to derive explicit finite-frequency sensitivity kernels for t and s following previous studies (e.g. Dahlen *et al.* 2000; Chevrot 2006; Calvet *et al.* 2006; Mondal & Long 2019) or analytically using adjoint methods (e.g. Sieminski *et al.* 2007). However, this is beyond the scope of the present work. Instead, to account for finite-frequency effects, we apply the so-called anisotropic heuristic finite-frequency kernels (HFFKs) discussed in VanderBeek & Faccenda (2021). These are constructed by assuming the Born sensitivity kernels developed by Dahlen *et al.* (2000) for isotropic scatterers can be extended to anisotropic models by simply making the slowness at the scatterer a function of the ray orientation, polarization direction, and the five parameters u, f', f'', ψ and γ. In effect, we are neglecting the influence of mode conversions as well as the mid- and near-field contributions to kernel structure (Favier *et al.* 2004). Under this assumption, the anisotropic finite-frequency traveltime is,

$$t = \tilde{t} + \sum_{j=1}^{M} K_j\left[u''_j + \left(u'_j - u''_j\right)\cos^2(\beta_j) - \tilde{u}_j\right] \quad (17)$$

where $\tilde{t}$ is the traveltime through a reference isotropic model defined by $\tilde{u}_j$; K_j is the sensitivity of the Born slowness kernel at the jth node. Considering that the splitting intensity is the difference between the traveltimes of the qS' and qS'' waveforms scaled by $\sin(2\beta)/2$, our assumption regarding the treatment of anisotropic scatterers leads to the finite-frequency splitting intensity equation,

$$s = \frac{1}{2}\sum_{j=1}^{M} K_j\left(u''_j - u'_j\right)\sin(2\beta_j) \quad (18)$$

To simplify the computation of K_j, we use the following approximation for the kernel's cross-sectional shape within the first Fresnel zone proposed by Schmandt & Humphreys (2010) and neglect the rapidly diminishing sensitivity outside this region,

$$K(x_r, x_n) = \frac{K_0}{\pi R_f^2(x_r)}\sin\left(\pi\frac{x_n^2}{R_f^2(x_r)}\right), \quad (19)$$

where x_r and x_n are coordinates measured in the ray-parallel and ray normal directions, respectively, with $x_r = 0$ located at the source; R_f is the radius of the first Fresnel zone; and K_0 is a dimensionless scalar chosen such that the integral of the slowness kernel is equivalent to the ray-theoretical sensitivity (i.e. the ray length; Hung *et al.* 2000). For a direct teleseismic phase, a reasonable approximation for R_f is

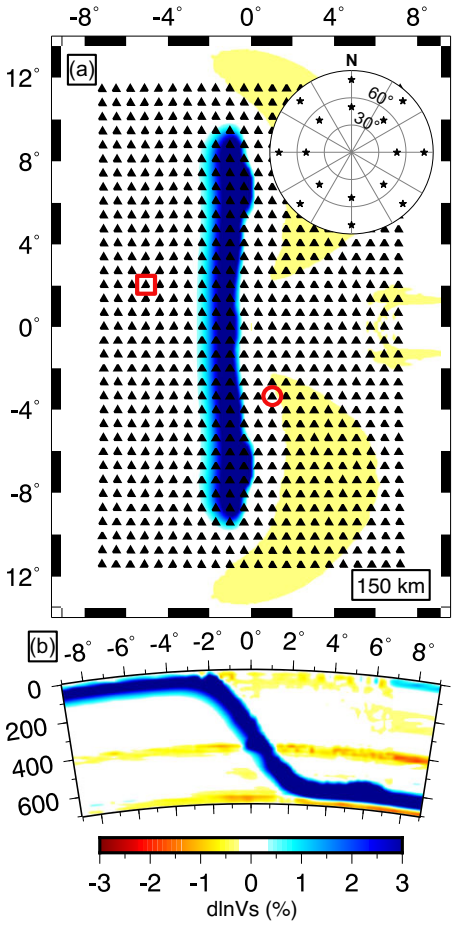

Figure 2. Isotropic structure, array geometry and distribution of teleseismic sources considered in this study. (a) Seismic stations (black triangles) are uniformly spaced 75 km apart and plotted over isotropic velocity heterogeneity in the true model at 150 km depth. Inset shows location of teleseismic sources (stars) relative to the experiment centre. Sources are located at distances of 50° and 80° and evenly distributed in backazimuth. An east–west cross-section through the centre of the true isotropic model at 0°N is shown in (b). Note that the isotropic structure is symmetric about 0°N. Red square and circle identify stations shown in Figs 5(a),(c) and (b),(d), respectively.

given by,

$$R_f(x_r) = \left[\frac{T x_r(L-x_r)}{L\tilde{u}(x_r)}\right]^{1/2}, \quad (20)$$

where T is the dominant period and L is the total length of the ray path. Fig. 3 of VanderBeek & Faccenda (2021) shows a comparison between the Born sensitivity kernel and the HFFK. Given these approximations (eqs 19 and 20), this kernel implementation is effectively a form of Fresnel volume tomography (Vasco *et al.* 1995; Zelt & Chen 2016) intended to impose physically based smoothing criteria on the solution and provide a framework for the inversion of multifrequency observables at reduced computational expense. Although not intended to reproduce the accuracy of the full Born kernels, VanderBeek & Faccenda (2021) demonstrated that the HFFKs can improve traveltime predictions for teleseismic P waves in both isotropic and anisotropic models compared to ray theory. In Section 5.2, we demonstrate the same is true for the principal S-wave observables.

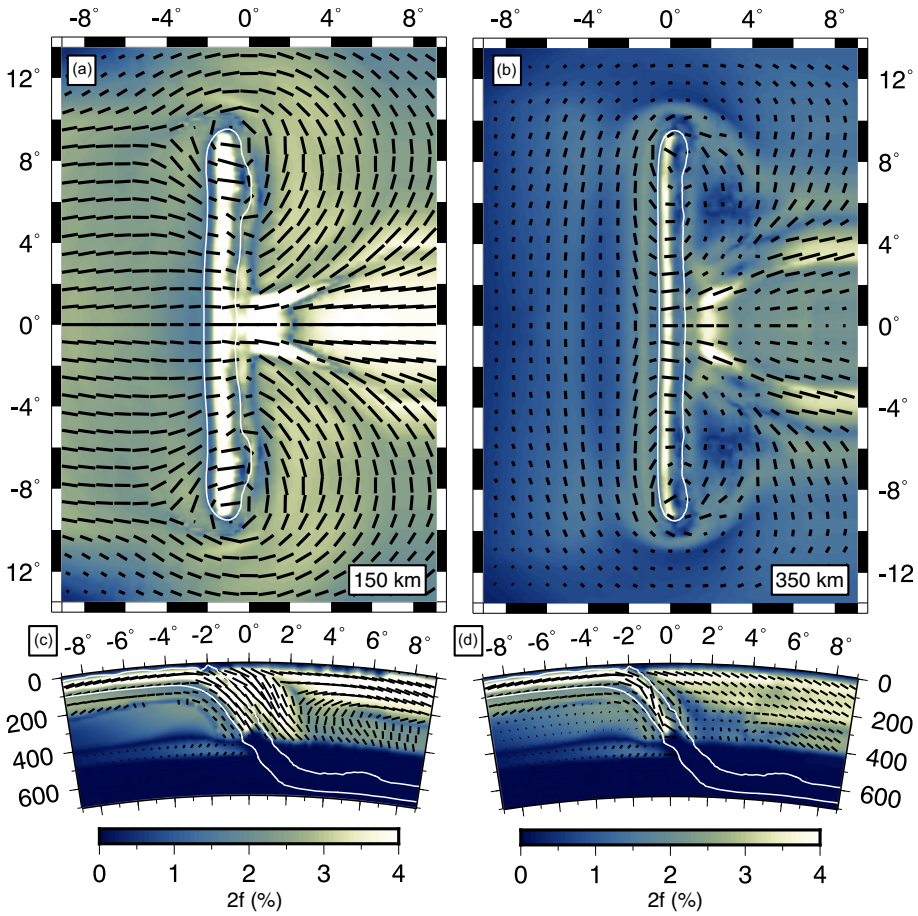

Figure 3. Cross-section through the true anisotropic model. Hexagonal symmetry axis vectors are plotted over the peak-to-peak magnitude of *S*-wave anisotropy (i.e. $2f''$) at (a) 150 km and (b) 350 km depth. East–west cross-section are shown at (c) 0°N and (d) 4°30'S. Symmetry axis vectors are scaled by the anisotropic magnitude and projected onto the cross-section plane. Note that the anisotropic structure is symmetric about 0°N.

3.3 Limitations

An important limitation of the proposed shear wave observables is that they are only applicable when the integrated splitting intensity at any point along the ray path is small relative to the period of the wave. The validity of this approximations was investigated in detail by Chevrot *et al.* (2004) using spectral element modelling. They found that provided $T/|\Delta t'' - \Delta t'| \gtrapprox 2\pi/\sqrt{2}$ eqs (4)–(7) are valid. In Appendix B, we describe a number of simple tests performed to evaluate the accuracy of eqs (6) and (7) as a function of the strength of anisotropy and deviations from linear polarization. For these tests, vertically travelling synthetic waveforms are modelled using a matrix propagation method (eqs 1–3; Rümpker & Silver 1998) applied to layered anisotropic models. Ray-theoretical delay times and splitting intensities were computed via the integration of eqs (6)–(7) through the layered models and compared to those measured on the simulated waveforms. Specifically, we explored how the accuracy of these expressions varies in response to (i) the strength of shear wave splitting (Figs B1 and B2), (ii) layered anisotropic heterogeneity (Figs B3 and B4) and (iii) elliptical shear wave polarizations (Figs B5–B7). Consistent with Chevrot *et al.* (2004), these results demonstrate that the prediction errors remain small (approximately one order of magnitude less than the measured observable) provided that $T/|\Delta t'' - \Delta t'| \gtrapprox 5$. Furthermore, the equations are applicable even when the incoming shear wave exhibits elliptical particle motion provided the principal displacement direction can be estimated; see Appendix B for details. Note that these tests do not consider finite-frequency kernels or wave propagation in 3-D heterogeneous media. The accuracy of our modelling strategy in response to such effects is discussed in Section 5.2 where we compare predicted shear wave observables to those measured on SPECFEM generated waveforms.

We do not expect that requiring $T/|\Delta t'' - \Delta t'| \gtrapprox 5$ will prohibit the general application of shear wave observables to teleseismic imaging. The global analysis of SKS splitting parameters by Walpole *et al.* (2014) found the average delay time between fast- and slow-polarized waveforms is 1.65 ± 0.62 s (± standard deviation). Because SKS waves reflect the integrated anisotropic structure of the upper mantle (Sieminski *et al.* 2007), we may expect similar splitting times for direct *S*-phases. Thus, the principal anisotropic delay and splitting intensity equations are likely applicable to periods $\gtrapprox 8$ s which corresponds to the mid- to lower- end of the period range commonly considered for teleseismic *S*-wave tomography. Furthermore, if the upper mantle is dominated by horizontally aligned symmetry axes, as predicted by plate-driven flow, then shear wave splitting should be weaker for direct *S*-phases because their propagation directions would align more closely with the fabric orientations (i.e. smaller angle α). Consequently, the difference between qS' and qS'' velocities would be reduced. The more laterally propagating *S*-phase is also likely to sample a greater diversity of anisotropic fabrics relative to the SKS phase which will generally result in a reduced splitting effect (Rümpker & Silver 1998).

4 ANISOTROPIC INVERSION METHOD

In Section 3, we related shear wave traveltimes and splitting intensities to anisotropic variables describing a hexagonally symmetric elastic medium. Here we describe an iterative inversion strategy to recover the 3-D distribution of these anisotropic parameters from the shear wave observables. We use an iterative Gauss–Newton approach to solve for the perturbations to a starting model that minimize a least-squares objective function of the form (e.g. Tarantola & Valette 1982; Aster *et al.* 2005),

$$\chi = \mathbf{r}(\mathbf{m}, \Delta\mathbf{m})^T \mathbf{C}_d^{-1} \mathbf{r}(\mathbf{m}, \Delta\mathbf{m}) + \mu^2 \Delta\mathbf{m}^T \mathbf{C}_m^{-1} \Delta\mathbf{m} + \lambda^2 (\mathbf{L}\Delta\mathbf{m})^T (\mathbf{L}\Delta\mathbf{m}) \quad (21)$$

where $\mathbf{r}$ is a (N × 1) vector of data residuals which is a non-linear function of the starting model, $\mathbf{m}$, and the (M × 1) cumulative model perturbation vector, $\Delta\mathbf{m}$; $\mathbf{C}_d$ is the (N × N) data covariance matrix which we assume to be diagonal and composed of the squared data uncertainties; $\mathbf{C}_m$ is the (M × M) model covariance matrix; $\mathbf{L}$ is an (M × M) matrix that defines the finite difference 3-D Laplacian operator that constrains $\Delta\mathbf{m}$ to be spatially smooth; lastly μ and λ are Lagrangian multipliers that limit the size and roughness of the model perturbation vector. The last two terms in eq. (21) are required to regularize the otherwise under-determined and ill-posed inverse problem.

Rather than solve for perturbations to the four spherical anisotropy terms (f', f'', ψ and γ) or the anisotropy vector (e.g. Chevrot 2006; Chevrot & Monteiller 2009; Wang & Zhao 2021), we parametrize the inversion following VanderBeek & Faccenda (2021) using the mean slowness (u) and the three anisotropic variables,

$$A = |f''| \cos^2(\gamma) \cos(2\psi) \quad (22)$$

$$B = |f''| \cos^2(\gamma) \sin(2\psi) \quad (23)$$

$$C = \sqrt{|f''|} \sin(\gamma) \quad (24)$$

The perturbation vector then consists of four sets of parameters, $\Delta\mathbf{m} = [\Delta\mathbf{u}; \Delta\mathbf{A}; \Delta\mathbf{B}; \Delta\mathbf{C}]$, that are discretized in space. To limit the number of free parameters, we take advantage of the strong linear correlation between f' and f'' (e.g. Becker *et al.* 2006). From the results of our micromechanical fabric modelling (Section 5), we use the ratio $f''/f' = -4.75$. Expressions for u'', u' and β as functions of A, B and C are provided in Appendix A. In this parametrization, the sign of f'' is lost and must be defined based on prior knowledge of the imaging target. A positive anisotropic fraction corresponds to materials with a seismically fast symmetry axis (e.g. olivine A-Type fabrics) while a negative sign implies a seismically slow symmetry axis (e.g. aligned fractures).

Although solving for these *ABC* parameters still poses a non-linear inverse problem, we have found that this parametrization provides stable solutions for isotropic starting models. This is because the traveltime or splitting intensity partial derivatives that direct the inversion are generally non-zero valued with respect to A and B even where the model is isotropic. This allows for the simultaneous determination of $\Delta\mathbf{u}$ and an azimuthally anisotropic model defined by $\Delta\mathbf{A}$ and $\Delta\mathbf{B}$. This initial solution provides an adequate starting point for subsequent iterations that update the fabric dip via adjustments to $\Delta\mathbf{C}$ (VanderBeek & Faccenda 2021). In comparison, differentiating the shear wave observables with respect to the spherical anisotropy parameters yields null sensitivity to ψ and γ where the model is isotropic. Consequently, the inversion can fail to converge to the correct solution if the symmetry axis orientations in the starting model are not sufficiently close to the true orientations (Munzarová *et al.* 2018). Alternatively, one could parametrize anisotropy using the components of the symmetry axis vector (e.g. Chevrot 2006; Chevrot & Monteiller 2009; Wang & Zhao 2021). However, this too results in null first-order sensitivity to anisotropy for isotropic starting models and multistage inversion schemes, accurate anisotropic starting models, or more computationally expensive solvers that exploit second-order derivatives are required to retrieve robust solutions.

To minimize eq. (21), we use the LSQR algorithm (Paige & Saunders 1982) to iteratively solve the following system of equations linearized about the current model,

$$\begin{bmatrix} \mathbf{C_d}^{-1/2}\mathbf{J} \\ \mu_u \mathbf{C}_u^{-1/2} \\ \mu_a \mathbf{C}_a^{-1/2} \\ \mu_a \mathbf{D}_f \\ \lambda_u \mathbf{L}_u \\ \lambda_a \mathbf{L}_a \end{bmatrix} \delta\mathbf{m}_l = \begin{bmatrix} \mathbf{C_d}^{-1/2}\mathbf{r_l} \\ -\mu_u \mathbf{C}_u^{-1/2} \Delta\mathbf{m}_{l-1} \\ \mathbf{0} \\ -\mu_a \mathbf{D}_f \Delta\mathbf{m}_{l-1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad (25)$$

where $\delta\mathbf{m}_l = [\delta\mathbf{u}_l; \delta\mathbf{A}_l; \delta\mathbf{B}_l; \delta\mathbf{C}_l]$ are the incremental parameter changes to be determined at the lth-iteration and $\Delta\mathbf{m}_{l-1}$ are the cumulative model perturbations from the previous $l-1$ iterations. The (N × M) matrix $\mathbf{J}$ is the data Jacobian. Explicit formula for the elements of $\mathbf{J}$ are provided in Appendix A. The model covariance and Laplacian matrices in eq. (21) have been separated into the rows that constrain the norm and smoothness of the mean slowness ($\mathbf{C}_u$, $\mathbf{L}_u$) and anisotropic parameters ($\mathbf{C}_a$, $\mathbf{L}_a$) and are weighted by the isotropic (μ_u, λ_u) and anisotropic (μ_a, λ_a) Lagrangian multipliers. The right-hand side vector is defined such that the cumulative slowness perturbations are minimized on each iteration while the incremental perturbations are minimized for the anisotropic parameters (i.e. zero-valued right-hand side) at each iteration. This is by design. Because the data are non-linear functions of A, B and C, the anisotropic perturbations may not converge when the solution is restricted to making small linearized steps from the starting values. This is not an issue for the mean slowness as the data are linear in u for a fixed ray geometry. To prevent the solution from favouring anisotropic over isotropic perturbations as the iterations progress, we minimize the cumulative perturbations to $\mathbf{f}''$ through the introduction of the sparse (M × M) matrix $\mathbf{D}_f$ obtained by differentiating $\Delta\mathbf{f}''_{l-1}$ with respect to $\Delta\mathbf{A}_{l-1}$, $\Delta\mathbf{B}_{l-1}$ and $\Delta\mathbf{C}_{l-1}$; see Appendix C for details. Because the anisotropic imaging problem is non-linear, the partial derivatives and regularization coefficients in eq. (25) are re-evaluated at each iteration based on the updated model and a new incremental perturbation vector is determined. Iterations proceed until the reduction in residual variance with respect to the precedent solution is no longer significant at the 95 per cent confidence level as determined by an F-test. For the anisotropic inversions performed in this study, convergence is reached within four iterations.

In addition to solving for the four anisotropic parameters described above, $\mathbf{J}$ is further augmented to include event static terms for traveltimes and splitting intensity. These effectively demean the observation residuals associated with a given event and are intended to remove the influence of structure outside the imaging volume where ray paths converge resulting in poor data coverage (Aki *et al.* 1977; Lévêque & Masson 1999; Masson & Romanowicz 2017). Note that such event demeaning applied to teleseismic traveltimes is equally applicable to splitting intensity. Because event statics are overdetermined, no regularization is applied. The inversion of demeaned data provides information only on relative changes in seismic velocity (Aki *et al.* 1977; Lévêque & Masson 1999; Masson & Romanowicz 2017). The consequences of this for resolving anisotropic structure is detailed by VanderBeek & Faccenda (2021). In general, provided the model is sufficiently well-sampled and is

anisotropically heterogeneous, then anisotropic fabrics are accurately recovered. Potential biases resulting from demeaning can be investigated through specific resolution tests.

5 SYNTHETIC TOMOGRAPHY EXPERIMENT: TARGET MODEL AND SEISMIC DATA

To test our anisotropic shear wave tomography method, we use the synthetic waveform data sets created by VanderBeek & Faccenda (2021) and briefly summarize their key features below. The teleseismic wavefield is modelled using the spectral element code SPECFEM 3D (Komatitsch & Tromp 1999; Chen & Tromp 2007) with the AxiSEM grid-injection technique (Nissen-Meyer *et al.* 2014; Monteiller *et al.* 2013, 2015, 2021). Outside the subduction zone model, 3-D wave propagation through a spherically symmetric Earth is solved using the AxiSEM software (Nissen-Meyer *et al.* 2014) with isotropic elastic properties defined by the IASP91 reference model (Kennett & Engdahl 1991). The displacements and tractions of the incoming teleseismic wavefield are stored on the boundaries of the 3-D subduction zone mesh and wave propagation is continued via the SPECFEM 3D solver; see Monteiller *et al.* (2013, 2015, 2021) for further details. Three-component seismograms are recorded by an array of 770 receivers equally spaced 75 km apart (Fig. 2a) yielding a station density comparable to the USArray. In total 16 double-couple sources are modelled; 8 at a range of 50° and another 8 at 80° equally distributed in backazimuth (Fig. 2a). The dominant period of the waveforms is 15 s and all S waves have an initial polarization angle of $\zeta = 60^\circ$. Importantly, the shear wave observables derived from this data set are completely independent of our tomography method.

The anisotropic elastic model used in computing the waveforms is the result of a geodynamic simulation of a slab (1000-km-long half-width) subducting freely in response to its negative buoyancy. Fully anisotropic elastic tensors are predicted via micromechanical modelling of polymineralic aggregates advected through the simulated mantle flow field (see Kaminski *et al.* 2004; Faccenda & Capitanio 2013; Faccenda 2014); resolution of anisotropic elastic model is 10 km. Parameters controlling the fabric development were calibrated based on high-strain laboratory experiments (Boneh *et al.* 2015). The tensors are subsequently simplified by extracting the dominant hexagonally symmetric component following the method of Browaeys & Chevrot (2004). In summary, each elastic tensor is rotated into the appropriate symmetry Cartesian coordinate system (SCCS) identified from the eigenvectors of the 3×3 Voigt stiffness tensor. Next, the hexagonal approximation of the tensor is obtained from a series of orthogonal projections. Finally, the resulting hexagonal tensor is rotated back to the geodynamic model coordinates by reversing the rotation into the SCCS. This simplification was made to maintain focus on the accuracy of the imaging methodology specifically designed to approximate hexagonal anisotropy. A second waveform data set was created using only the isotropic component of the elastic tensors.

The isotropic component of the synthetic subduction zone is shown in Fig. 2 and contains only one significant anomaly–the seismically fast slab. The anisotropic component of the synthetic model is shown in Fig. 3 and contains five significant imaging targets that provide insight into the structure and dynamics of the subduction zone. (i) Throughout the upper 300 km, toroidal mantle flow generates a circular pattern in the symmetry axis orientations around the slab edges (Fig. 3a). (ii) At greater depths beneath the incoming plate there is a region of trench-parallel anisotropy reflecting pressure-driven flow from slab roll-back (Fig. 3b). (iii) Surrounding the subducting lithosphere, flow entrainment produces anisotropic symmetry axes that follow the trajectory of the descending plate (Figs 3c and d). (iv) The subducting lithosphere also contains frozen-in anisotropic fabrics characterized by ∼4 per cent S-wave speed variations oriented east–west. (v) Mantle circulation within the wedge generates a corner-flow type pattern in anisotropy near the mid-plate that evolves towards the edges (Figs 3c and d).

5.1 Measurement of principal shear wave observables

Our tomographic method requires that S-wave traveltimes are measured in the direction of polarization which is not a coordinate system routinely considered in the analysis of body wave delays. If an appropriate reference waveform is known *a priori*, then the generalized shear wave observables may be measured using the cross-correlations defined by Sieminski *et al.* (2007) and transformed into the principal observables following eq. (10). However, such a reference is generally not known and we outline a measurement strategy that relies only on the observed waveforms to generate data suitable for our imaging method. We assume seismograms are initially processed to remove instrument responses, filtered (here we use a second-order band-pass filter with corners at 15 and 40 s), and rotated into ray-aligned QTL-coordinates (Fig. 1a). For a given event, we then measure the principal traveltimes in four steps:

(i) Relative traveltime delays on the transverse channel are measured with respect to a reference radial Earth velocity model using the multichannel cross-correlation method of VanDecar & Crosson (1990).

(ii) The aligned waveforms from Step 1 are stacked to provide an estimate of the incoming shear wave. The arrival time of the stacked waveform is picked and combined with the relative delays to yield traveltime measurements (e.g. Lou *et al.* 2013).

(iii) The principal polarization direction is measured via eigendecomposition of the trace covariance matrix (e.g. Flinn 1965) computed using the stacked three-component waveform from Step 2. In computing the trace covariance, a full period of the waveform is used (15 s in this study). The polarization angle, ζ, is estimated as the azimuth of the eigenvector with the largest magnitude eigenvalue. This angle is stored for each event and used in the forward modelling of principal traveltimes and splitting intensities.

(iv) Seismograms are rotated about the L-axis into the principal polarization direction and the principal delay times are measured via cross-correlation of the waveforms on the x_1-channel (Fig. 1a). According to eq. (4), these waveforms should be well-correlated as they reflect time shifted versions of the initial shear wave. A new pick on the stacked x_1-component waveform is made to yield the traveltimes.

These steps are illustrated in Fig. 4.

Once the waveforms are aligned on the x_1-channel, the principal splitting intensity is readily computed following Chevrot (2000),

$$s_{ij} = \frac{\dot{\mathbf{d}}_{ij} \cdot \mathbf{w}_{ij}}{\|\dot{\mathbf{d}}_{ij}\|^2}, \tag{26}$$

where $\dot{\mathbf{d}}_{ij}$ is the time-derivative of the x_1-component and $\mathbf{w}_{ij}$ is the x_2-component of the waveform observed at the ith-station for the

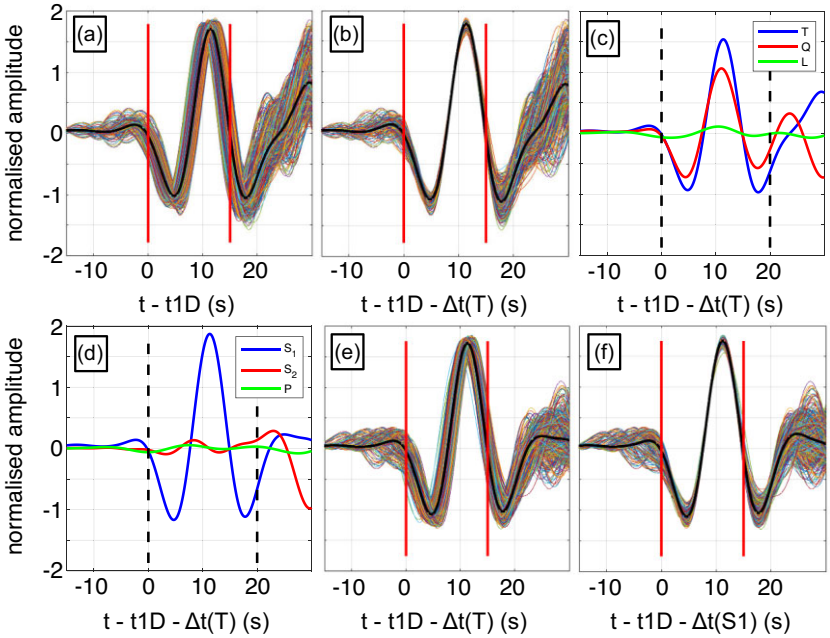

Figure 4. Illustration of multichannel cross-correlation (MCC) analysis. (a) Alignment of transverse component waveforms based on IASP91 traveltime predictions (t1D). The source is located 80° from the centre of the array at a backazimuth of 90°; thin coloured lines are the seismograms recorded at each receiver position in Fig. 2. (b) Waveform alignment after several iterations of MCC; $\Delta t(T)$ in the x-axis title refers to the delay times measured on the transverse channel. (c) Stacked three-component seismogram in TQL coordinates after transverse channel alignment. (d) Stacked three-component seismogram after rotation into the principal coordinate system. (e) Alignment of S waves observed in principal direction (S_1-component; Fig. 1) using the transverse component delays; waveforms observed in this orientation are no longer aligned due to seismic anisotropy. (f) Re-alignment of waveforms following MCC measurement of S_1-component delays, $\Delta t(S1)$. Vertical lines in all panels define the window used in the cross-correlation and polarization analyses. Trace amplitudes are normalized by the RMS amplitude in the window.

jth event. We compute s_{ij} in a 15 s (one period) window beginning at the S-wave onset. Note that the sign change in our eq. (26) with respect to eq. (7) of Chevrot (2000) is due to our definition of a positive delay as corresponding to a later arrival; the scaling factor of 2 has also been removed such that the principal splitting intensity varies between $\pm 0.5(\Delta t'' - \Delta t')$ in accordance with our eq. (7).

Examples of split shear waves produced from our SPECFEM waveform simulations are shown in Figs 5(a) and (b) for stations on the incoming plate and above the mantle wedge. As predicted by eqs (4) and (5), the x_1-component waveform is a time-shifted version of the equivalent isotropic trace and the x_2-component is consistent with the time-derivative of the x_1-signal. Examples of single-station splitting intensity measurements (computed via eq. 29) as a function of polarization-azimuth are plotted for these stations in Figs 5(c) and (d). On the incoming plate where the anisotropic signal is dominated by a shallow layer of horizontal east–west oriented fabrics (Fig. 5c), the splitting intensity follows a clear $\sin(2\zeta)$ trend. Beneath the wedge where 3-D anisotropic heterogeneity is stronger both laterally and vertically, the splitting intensity signal is more complex and generally weaker (Fig. 5d) as the diversity of fabrics tends to destructively interfere.

5.2 Error estimates in measured and predicted shear wave observables

To appropriately assess the fit of the tomographic models recovered from the synthetic data sets, we must know the error in our measured shear observables and the accuracy to which we can predict them using the methods in Section 3. We first evaluate the error inherent to our synthetic data set introduced by our measurement procedure. To estimate these errors, we measure delay times and splitting intensities as outlined in Section 5.1 of shear waves propagated through the 1-D velocity model IASP91 (Kennett & Engdahl 1991) using the same source–receiver geometry as for the subduction zone model. In this case, all measured delay times and splitting intensities should be zero. However, we measure a RMS delay of 100 ms for shear waves observed on the transverse channel and 188 ms for those measured in the estimated principal polarization direction. The RMS splitting intensity measured is 157 ms suggesting that only a minor amount of apparent splitting may be generated due to interactions with velocity interfaces or our measurement procedure.

We attribute the increased RMS delay of the polarization-aligned measurements to waveform distortions caused by mode conversions. These errors are comparable to the time step used in the SPECFEM simulations (100 ms) but larger than the accuracy of TauP calculations for direct body wave traveltimes (∼50 ms; Crotwell *et al.* 1999). Relative phase arrival times measured from SPECFEM seismograms may differ slightly arising from differences in the discretization of the IASP91 model. Additionally, delays are measured using the array-based multichannel cross-correlation method of VanDecar & Crosson (1990) rather than cross-correlating with a reference waveform computed for every station. While minor, changes in the waveform across the array can yield an inconsistent set of delays for station pairs from which the best-fitting relative delay times are determined.

Next we assess the accuracy of our forward modelling procedure and the HFFK approximation by comparing data predicted through

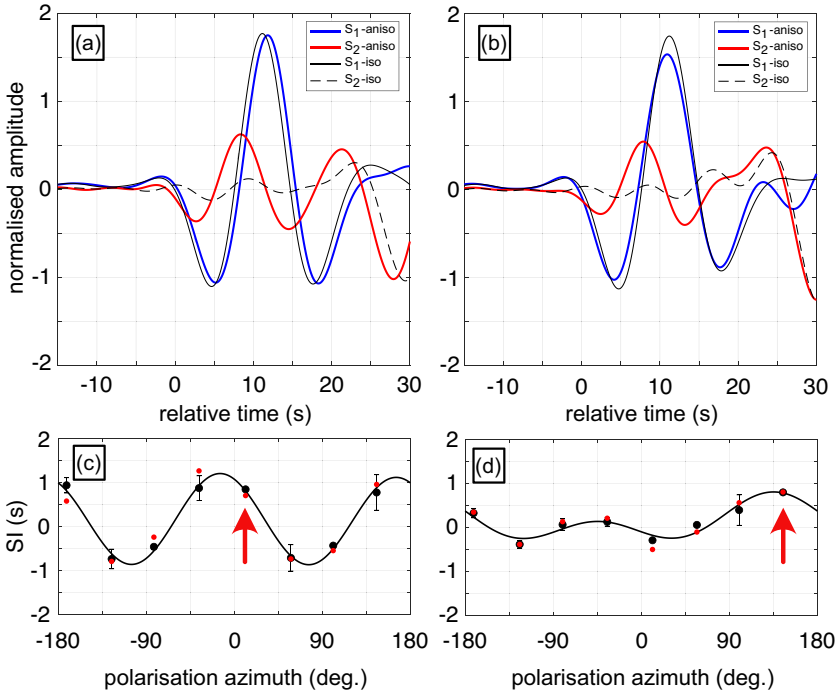

Figure 5. Effect of anisotropy on shear waveforms with examples of single-station splitting intensity measurements and predictions. Shear waveforms propagated through the anisotropic (blue = x_1-component; red = x_2-component) and purely isotropic (solid black = x_1-component; dashed black = x_2-component) subduction zone models are shown for stations located (a) on the incoming plate and (b) above the mantle wedge; see Fig. 2 for locations. Time of 0 s corresponds to the arrival of the isotropic *S* wave. Splitting intensity measurements (black points) and predictions (red points) for the station located (c) on the incoming plate and (d) above the mantle wedge are plotted as a function of the *S*-wave polarization direction for events located 80° from the centre of the array. For reference, the best-fitting sinusoidal curve composed of the first two azimuthal harmonics is plotted. Red arrows in (c) and (d) indicate the splitting intensity measurements obtained from the waveforms in (a) and (b), respectively. Vertical bars represent the measurement standard error from eq. (B1) of Chevrot (2000).

the true anisotropic tomography model to the shear wave observables measured from SPECFEM generated waveforms. The RMS error between the measured and predicted ray-theoretical traveltimes is 527 ms. This error is reduced to 330 ms when traveltimes are computed using our HFFK approximation. The HFFKs offer a less drastic improvement in the prediction of splitting intensities yielding an RMS error of 278 ms compared to 291 ms for ray-theory. While appreciable, these errors are far below the signal in the shear-wave observables considering the RMS delay time with respect to IASP91 (Kennett & Engdahl 1991) is 1,211 ms and the RMS splitting intensity is 600 ms. Additionally, these errors are small with respect to the accuracy of modern teleseismic shear wave velocity models which generally yield S delay time residuals with an RMS >400 ms. Estimated standard errors for splitting intensity measurements from high-quality SKS recordings are ∼250 ms (Monteiller & Chevrot 2010).

In regards to the origin of the errors in our forward modelling approach, they are not the result of anomalously strong anisotropy as all measured splitting intensities are <2.5 s and thus satisfy the criteria $T/\|\Delta t'' - \Delta t'|\gtrapprox 5$. Rather, they arise from inaccuracies in our simplified approximation to the far-field Born kernels (Section 3.2). Chevrot *et al.* (2004) demonstrated that the true far-field kernels predict splitting intensities to within 5 per cent of values measured from full-waveform synthetics. They also found that geometrical ray-theory increasingly overpredicts the anisotropic effect of anomalies with spatial scales <0.75× the Fresnel width (∼200 km for thistudy). Furthermore, when significant anisotropy exists within 2-wavelengths of the station, neglecting mid- and near-field contributions of the Green's tensor can introduce errors up to 20 per cent (Favier *et al.* 2004). This is indeed the case for the synthetic subduction zone model in which 5 per cent shear wave anisotropy extends to 40 km depth and the seismic wavelength is ∼60 km. Considering that the accuracy of our kernel implementation falls somewhere between geometrical ray theory and the far-field Born kernels, our error estimates of ∼300 ms are not surprising. The general consequence of such inaccuracies is poorer spatial resolution and an underestimation of anisotropic magnitudes. To verify that such errors do not introduce significant artefacts, we performed an anisotropic inversion of self-consistent data—synthetic delay times and splitting intensities predicted using our forward modelling approximation rather than from full-waveform simulations. In this way, the predicted observables are exact. The results (Figs S1–S2) are nearly identical to those obtained from SPECFEM generated observables (presented in Section 7) highlighting that data coverage is the primary factor limiting model recovery.

We continue with our kernel approximation but do not attempt to fit the observations beyond the their accuracy. Despite our simplifications, the results in Section 7 demonstrate that geodynamically significant anisotropic features of the subduction zone model are constrained. A more accurate treatment of observable sensitivity will only serve to improve our imaging strategy.

6 SELECTION OF INVERSION PARAMETERS

Solving eq. (25) requires a number of user-defined inputs. Specifically, the starting model, spatial discretization of inversion parameters, and uncertainties in the data and model parameters must be defined. Optimal Lagrangian multipliers that balance data fit against the model norm and roughness must also be identified. Here we detail the selection of these parameters for our synthetic inversions.

6.1 Starting model and parameter discretization

The calculation of shear wave observables is performed using a regular grid with uniform 10 km spacing between nodes. Each node is defined by the five anisotropic variables u, f, f', ψ and γ. This vector-valued field defines the forward model $\mathbf{m}$. The model is centred at 0°N, 0°E and spans 2000 km in the x-direction, 3000 km in the y-direction and 700 km in depth. To account for Earth's curvature in our Cartesian model geometry, an Earth-flattening transform is applied to the seismic velocities and model depths (Müller 1971). For all inversions, the starting model is isotropic and defined by the 1-D far-field velocity profile taken from the geodynamic simulation which closely resembles IASP91 (Kennett & Engdahl 1991) but without a crustal layer. We use the TauP toolkit (Crotwell *et al.* 1999) to predict ray paths using the initial 1-D velocity profile. In this study, the use of 1-D rays is justified given the relatively long period of the waveforms and the magnitude and dimensions of both isotropic and anisotropic heterogeneity. Finite-frequency traveltimes through 3-D velocity models are computed following eq. (17) where $\tilde{t}$ is the TauP predicted traveltime through the starting model slowness profile $\tilde{\mathbf{u}}$. Frequency-dependent splitting intensities are computed from eq. (18).

We solve for perturbations to the mean slowness ($\Delta\mathbf{u}$) and the three anisotropic parameters ($\Delta\mathbf{A}$, $\Delta\mathbf{B}$ and $\Delta\mathbf{C}$) on a regular grid with uniform 50 km spacing. Together, these parameters make up the perturbational model $\Delta\mathbf{m}$. To restrict the number of free parameters, we limit perturbations to A, B and C to the upper 500 km where anisotropy is strongest in our geodynamic model. This choice has a minor effect on the recovered model and acts primarily to limit vertical smearing of anisotropic structure (e.g. VanderBeek & Faccenda 2021). Event static terms for the traveltime and splitting intensity observations are included in the inversion. These act to demean the observations and, in real-world application, effectively remove any signal coming from structure outside the imaging volume (e.g. Aki *et al.* 1977; Lévêque & Masson 1999; Masson & Romanowicz 2017).

6.2 Regularization parameters

We assume a diagonal data covariance matrix, $\mathbf{C}_d$, where all elements are 0.3^2 s^2 based on our traveltime and splitting intensity error estimates discussed in Section 5.2. Defining the matrix $\mathbf{C}_m$ that controls the solution damping assumes one has knowledge of the parameter covariances. This is rarely the case and these matrices are generally assumed to be diagonal with coefficients corresponding to an estimate of individual squared parameter uncertainties. However, these too are poorly defined and trade-off directly with the Lagrangian multipliers ($\lambda_{u,a}$) and parameter discretization (finer discretizations will generally require weaker damping to accommodate a proportionally larger perturbational vector norm). Rather than *a priori* define $\mathbf{C}_m$, we use a Levenberg–Marquardt style damping such that the diagonal of $\mathbf{C}_m^{-1}$ is given by the mean of the squared parameter sensitivities (i.e. squared row norm of $\mathbf{J}$; Moré 1978). The mean-squared sensitivity is evaluated separately over each parameter set ($\Delta\mathbf{u}$, $\Delta\mathbf{A}$, $\Delta\mathbf{B}$ and $\Delta\mathbf{C}$) which renders the inverse problem scale-invariant. This is beneficial when, as in our case, the parameters have different expected magnitudes. The diagonal elements of $\mathbf{C}_u^{-1/2}$ are also weighted by $\overline{u}/u_k$ where $\overline{u}$ is the mean slowness in the model. This prevents biasing the solution towards models with stronger fractional velocity perturbations at greater depths (i.e. higher-velocity/lower-slowness regions; Toomey *et al.* 1994). The diagonal elements $\mathbf{C}_m^{-1/2}$ are also used to weight the rows of $\mathbf{L}_m$ and $\mathbf{D}_f$. In this way, all regularization equations have comparable influence on the objective function prior to the selection of the user-defined $\lambda_{u,a}$ and $\mu_{u,a}$ weights.

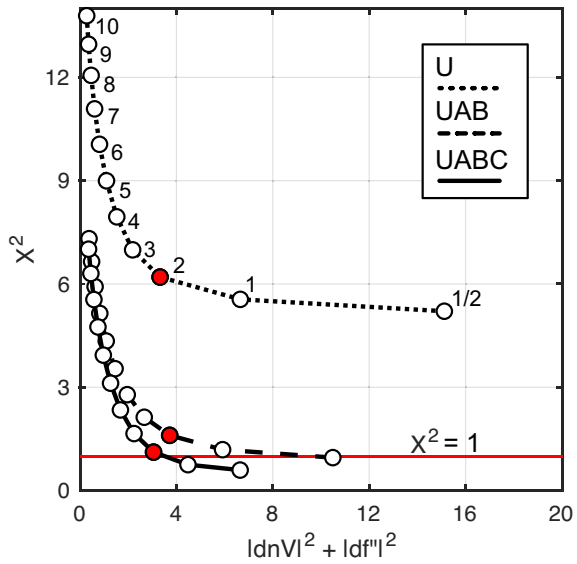

Figure 6. Selected trade-off curves for determining damping weights ($\mu_{u,a}$). The normalized data variance ($\chi^2 = \mathbf{r}^T\mathbf{C}_d^{-1}\mathbf{r}/N$) is plotted against the squared norm of the relative isotropic (d$\ln V$) and anisotropic (df'') perturbations. Trade-off curves are plotted for inversions in which we solved for purely isotropic (U; dotted line), azimuthally anisotropic (UAB; dashed line), and all four anisotropic (UABC; solid line) parameters. The damping weights ($\mu_{u,a}$) corresponding to each solution are indicated on the isotropic L-curve. The smoothing weights for all inversions are defined such that $\lambda_{u,a}/\mu_{u,a} = 10$. Isotropic and anisotropic perturbations were equally damped ($\mu_u = \mu_a$). The preferred solutions are indicated in red and all correspond to $\mu_{u,a} = 2$. Red line at $\chi^2 = 1$ is the ideal data variance given the estimated data uncertainty.

Appropriate Lagrangian multipliers were selected via the construction of L-curves (e.g. Aster *et al.* 2005) in which the data variance ($\mathbf{r}^T\mathbf{C}_d^{-1}\mathbf{r}/N$) is plotted against the model variance ($\Delta\mathbf{m}^T\mathbf{C}_m^{-1}\Delta\mathbf{m}/M$). Solutions near the corner of the L-curve are considered ideal as a further increase in model complexity (i.e. increased model variance) does not substantially improve the data prediction while simpler models rapidly increase the residual norm. To identify an appropriate ratio of λ_u-to-μ_u, we follow the same strategy outlined by VanderBeek & Faccenda (2021). First, we constructed several L-curves from isotropic inversions using fixed ratios of λ_u/μ_u and systematically varied μ_u. From this analysis, we found $\lambda_u/\mu_u = 10$ provides reasonably smooth solutions without compromising data fit. Within the set of isotropic inversions corresponding to $\lambda_u/\mu_u = 10$, we found $\mu_u = 2$ corresponds to the corner of the L-curve (Fig. 6). Not wanting to favour isotropic over anisotropic perturbations, we equally damp both parameter sets by setting $\mu_a =$

μ_u (the magnitudes of these weights are directly comparable due to the way in which $\mathbf{C}_m^{-1/2}$ was defined) and adopt the same smoothing-to-damping ratios of $\lambda_a/\mu_a = \lambda_u/\mu_u = 10$. After performing a series of anisotropic inversions across different μ_u values, we found that $\mu_u = \mu_a = 2$ provides optimal solutions (Fig. 6).

7 TOMOGRAPHIC RESULTS

In the following sections, we present a series of tomographic models constructed under different imaging assumptions. First, we consider purely isotropic inversions to evaluate the nature of anisotropy-induced artefacts. Next, we present azimuthally anisotropic solutions followed by inversions that include dipping fabrics.

7.1 Isotropic solutions

We begin by evaluating the ideal recovery of isotropic anomalies by inverting traveltimes derived from waveforms propagated through the isotropic subduction zone model. We refer to this model as the Iso-U solution (Figs 7a, b and 8a–c) it illustrates the resolution of the slab anomaly subject to our forward-modelling approximations, imperfect data coverage, and inversion regularization but without complications from anisotropy. The geometry of the descending plate is accurately recovered albeit with a reduced magnitude and spatial smoothing expected given the imaging constraints. Due to the relative nature of the traveltime residuals, some low-velocity artefacts appear around the fast-slab anomaly. However, the magnitude of these artefacts is generally weak (<1 per cent) and spatially distributed throughout the mantle surrounding the descending plate.

A number of isotropic imaging artefacts emerge when anisotropic heterogeneity is present in the target model but neglected in the imaging. The tomographic model recovered from inverting *S*-wave traveltimes measured on the transverse component of waveforms propagated through the anisotropic subduction zone (Aniso-U solution) is shown in Figs 7(c), (d) and 8(d)–(f). There is a localization and increase in magnitude of low-velocity artefacts, particularly around the slab edges (Figs 7c and d). In cross-section (Figs 8d–f), sub- and supraslab low-velocity anomalies are present and exhibit along strike changes in geometry and intensity. We also find a general increase in the magnitude of the slab high-velocity anomaly and the geometry of the descending plate is poorly recovered along the central portion of the subduction zone (Fig. 8e).

The increase in anomaly magnitudes in the Aniso-U relative to the Iso-U solution reflects the mapping of anisotropic heterogeneity into isotropic perturbations. In particular, the steeply dipping fabrics within and around the descending plate act to increase the apparent velocity of the mantle in this region. This relative increase in velocity is further exaggerated by the presence of more horizontal symmetry axes in the subslab region which reduce the propagation velocity of shear waves travelling at high incidence angles. The dependence of the *S*-wave velocity on both the ray propagation direction and the incoming polarization direction results in an asymmetric distribution of artefacts about 0°N despite the symmetry present in the geodynamic model and uniform back-azimuthal distribution of events.

7.2 Anisotropic solutions

It is clear from Section 7.1 that neglecting anisotropy in teleseismic shear wave imaging can lead to significant artefacts in the recovered tomographic model. Now we evaluate the ability of our anisotropic imaging method to discern isotropic from anisotropic heterogeneity and yield more accurate images of the upper mantle. First, we consider an azimuthally anisotropic solution (Aniso-UAB) in which the symmetry axes are confined to the horizontal plane by only solving for perturbations to u, A and B. We find that the magnitude of low-velocity artefacts present in the Aniso-U solution are reduced in the Aniso-UAB model (Figs 7e, f and 8g–i). However, significant velocity reductions are still imaged in the subslab region in the southern half of the synthetic study area. We also find that the geometry of the slab is better reconstructed along the central portion of the subduction zone when azimuthal anisotropy is included in the inversion.

Recovery of the true isotropic structure continues to improve when symmetry axis dip is included in the inversions via the C parameter. This Aniso-UABC solution is shown in Figs 7(g), (h) and 8(j)–(l) and is very similar to the Iso-U solution which represents the best-case recovery of isotropic heterogeneity. The magnitude of spurious low-velocity zones are further decreased and the magnitude of the slab anomaly is more uniform in the Aniso-UABC model. These improvements in isotropic recovery are clearly illustrated in Fig. S3 in which differences in velocity anomalies between the Iso-U and the Aniso-U, Aniso-UAB, and Aniso-UABC solutions are plotted.

In addition to improving the recovery of isotropic heterogeneity, the Aniso-UAB and Aniso-UABC solutions accurately image many of the anisotropic features present in the target model. Both the azimuthally anisotropic model (Figs 9a and b) and the inversion including dip (Figs 9c and d) image well the toroidal flow pattern present throughout the mantle wedge but the Aniso-UABC solution better recovers this structure at depths below ∼300 km. In the Aniso-UAB model, anisotropic fabrics are poorly recovered where the true symmetry axes are steeply dipping (Figs 10a–c). In contrast, the Aniso-UABC model accurately images such regions. In particular, the increased anisotropic strength and steeply dipping fabrics associated with the descending plate are clear in the Aniso-UABC solution (Figs 10d–f). Along strike changes in anisotropic heterogeneity within the mantle wedge (Figs 3c and d) are also evident in the Aniso-UABC solution (Figs 10d–f). Through the centre of the subduction zone, the transition from a thick region of entrained mantle flow to more horizontal fabrics in the shallow mantle wedge is recovered. The thick block of steeply dipping fabrics near 0°N thins towards the ends of the subducting plate becoming more localized to the slab. Both solutions poorly image anisotropic structure within and beneath the incoming plate. While the shallow trench-normal fabrics are visible in the tomographic models (Fig. 9), the magnitude is underestimated and no deeper trench-parallel fabrics are present. This is not surprising considering that the anisotropic traveltime and splitting intensity signals will tend towards zero for teleseismic paths traversing two overlying anisotropic regions with orthogonal fabrics. Finally, we note that the Aniso-UAB and Aniso-UABC solutions were constrained by both relative traveltimes and splitting intensities. We also performed inversions using only traveltimes and found that these data were sufficient to constrain the anisotropic parameters. In practice, splitting intensity data may be important for distinguishing between isotropic and anisotropic heterogeneity when sampling of the medium is more biased since these data are generally insensitive to changes in isotropic velocity (Chevrot 2006).

To illustrate the influence of our finite-frequency kernel approximation, we present a ray-theoretical solution of the Aniso-UABC inversion in Figs S4 and S5. As expected, ray theory further underestimates the magnitude of anomalies. This is particularly evident in

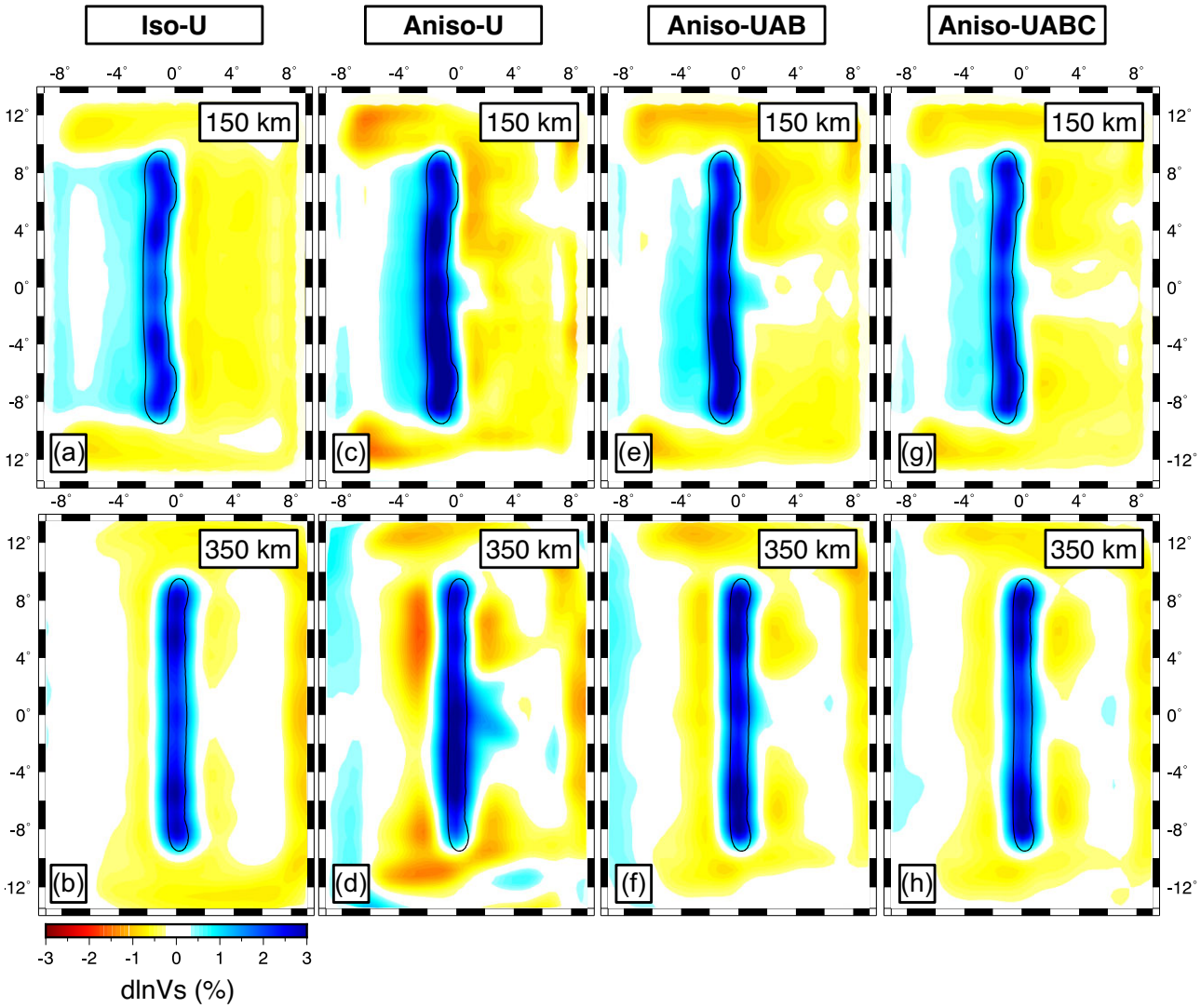

Figure 7. Comparison of recovered isotropic heterogeneity. The isotropic inversion of relative traveltimes generated by the isotropic subduction zone model (Iso-U) is shown in (a, b). The remaining panels (c–h) show the inversion of relative traveltimes generated by the anisotropic subduction zone model. (c, d) The Aniso-U solution solved only for isotropic (u) parameters. (e, f) The Aniso-UAB solution solved for u and the azimuthally anisotropic A, B parameters. (g, h) The Aniso-UABC solution included fabric dip (i.e. simultaneous inversion for u, A, B and C parameters). Solutions are plotted at 150 km (top row) and 350 km (bottom row) depth. The target fast slab anomaly is the same for all inversions and is outlined in black.

the anisotropy field (Fig. S5). However, the geometry of the recovered anomalies is similar between the models. Improving resolution requires leveraging multifrequency observables.

8 DISCUSSION

8.1 Comparison of tomographic models

Our synthetic tomographic inversions demonstrate that accounting for anisotropy in teleseismic *S*-wave imaging is crucial for accurately recovering isotropic heterogeneity and adequately fitting delay times and splitting intensity observations. This is clearly illustrated in Fig. 11 where the normalized data residual is plotted against the normalized difference between the isotropic perturbations recovered in the Iso-U solution with those recovered in the Aniso-U, Aniso-UAB and Aniso-UABC solutions. As a more complete description of anisotropy is included in the tomography, the data fit and isotropic perturbations converge towards the Iso-U model. While the Aniso-UAB and Aniso-UABC solutions similarly fit the data and yield similar isotropic perturbations, including fabric dip further reduces the magnitude of subslab low-velocity artefacts (Figs 8 and S3). Better fitting solutions are expected as the number of free parameters in the inversion increases. However, improved data variance reduction in the Aniso-UAB and Aniso-UABC solutions cannot only be attributed to increased model complexity. As shown in Fig. 6, the norm of fractional perturbations to seismic velocity, due to both changes in mean slowness and anisotropy, are comparable across all solutions. Thus, the better fitting Aniso-UAB and Aniso-UABC models do not simply reflect a larger model perturbation vector norm.

As is common in damped tomographic inversions, the magnitude of both isotropic and anisotropic heterogeneity is generally underestimated. All models reconstruct the slab anomaly to within 50–75 per cent of its full amplitude. Both the Aniso-UAB and Aniso-UABC solutions underestimate f' on average by 0.5 per cent (Figs 12a and b). This under-recovery may be influenced by a num-

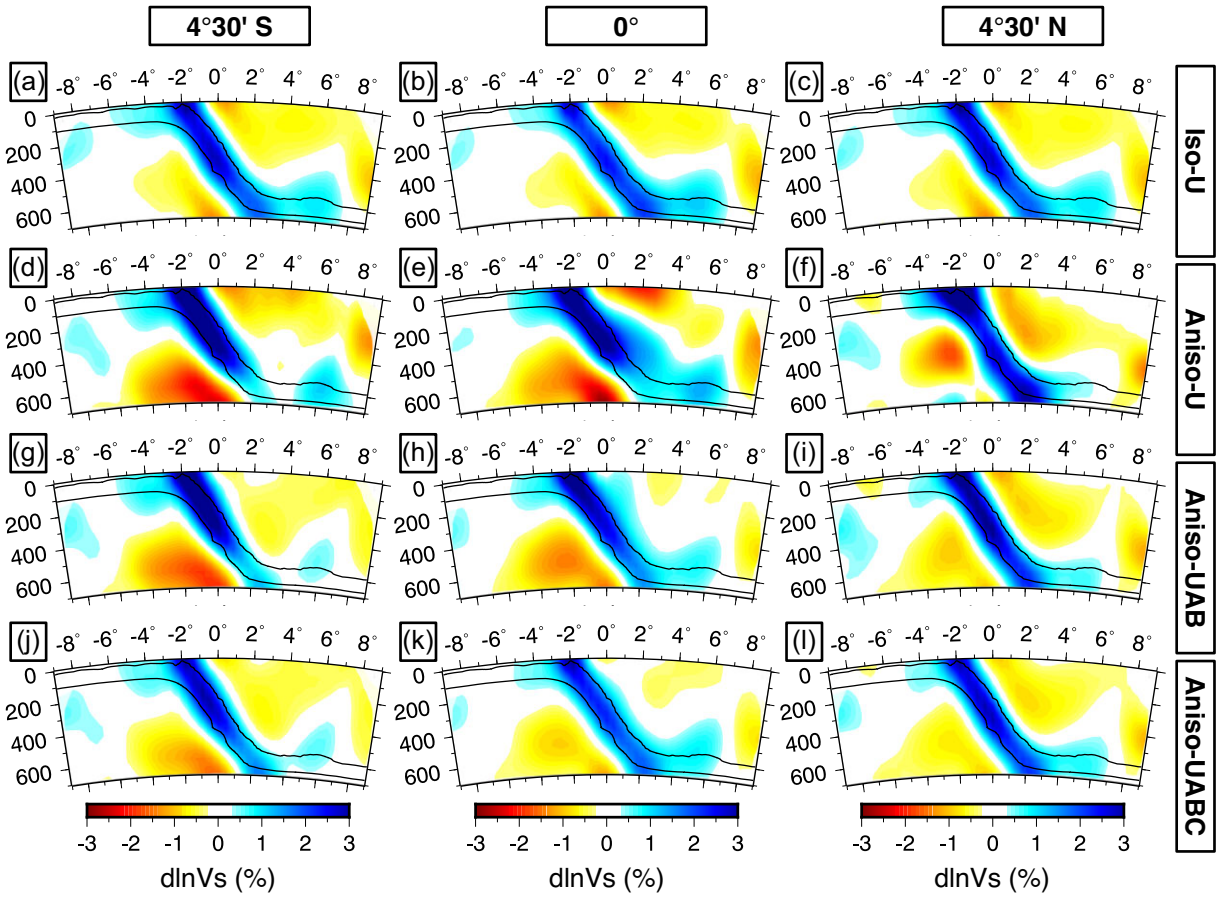

Figure 8. Comparison of recovered isotropic heterogeneity. West-to-East cross-sections are show at −4.5° (left-hand column), 0°N (centre column), and 4.5° (right-hand column). The (a–c) Iso-U, (d–f) Aniso-U, (g–i) Aniso-UAB and (j–l) Aniso-UABC solutions are plotted in each row; see Fig. 7 caption and text for description of differences between models. The target fast slab anomaly is the same for all inversions and is outlined in black.

ber of factors in addition to regularization effects. (i) The limited sampling of incidence angles by the teleseismic waves does not fully capture the extent of directional velocity variations leading to weaker anisotropic magnitude estimates that can trade-off with mean velocity. (ii) The vertically integrated effect of depth-dependent anisotropic fabrics generally produces a weaker apparent anisotropic signal (e.g. Appendix B; Rümpker & Silver 1998) that when combined with the poor depth resolution of teleseismic data results in a weaker estimate of the anisotropic magnitude. (iii) Inaccuracies in our kernel approximation may limit anisotropic magnitude recovery (Chevrot *et al.* 2004; Favier *et al.* 2004).

We find that upper mantle anisotropy fabric orientations are more accurately imaged when symmetry axis dip is included in the inversion (Figs 12c–f). The Aniso-UABC model constrains symmetry axis orientations with a weighted mean error in ψ and γ of 15° and 12°, respectively. In comparison, the weighted mean error is 17° for ψ and 20° for γ in the Aniso-UAB model. The means are weighted by $\sqrt{f_k'' g_k''}$ where f_k'' and g_k'' are the recovered and true anisotropic magnitudes at the kth parameter. This weighting is used because angular differences become irrelevant as the anisotropic magnitude approaches zero. These errors are generally less than the differences between mantle azimuthal anisotropy constrained by SWS and surface waves and that predicted from mantle convection models (e.g. Conrad *et al.* 2007; Becker *et al.* 2014; Zhou *et al.* 2018). Therefore, body wave-constrained anisotropic fabrics may provide useful constraints for geodynamic models of mantle convection.

It is instructive to compare our anisotropic *S*-wave tomography to the anisotropic *P*-wave tomography of (VanderBeek & Faccenda 2021) who used the same synthetic data set and inversion methodology. Two major differences are apparent. (i) Shear waves appear less sensitive to fabric dip compared to *P* waves. This was also noted by Beller & Chevrot (2020) and is evident from the modest improvement in data fit moving from our Aniso-UAB to Aniso-UABC solutions (Fig. 11). Slab recovery and reduction in low-velocity artefacts are also reduced in azimuthally anisotropic inversions. This is in contrast to the *P*-wave tomography in which solving for fabric dip was key to adequately fitting the delay times and removing erroneous low-velocity zones. (ii) Shear waves better constrain shallow anisotropic heterogeneity. At steep incidence angles, the anisotropic signal in *P* wave delays is reduced. The opposite is generally true for *S* waves provided that the symmetry axes are not at large angles to the plane of *S*-wave polarization. Both this study and that of VanderBeek & Faccenda (2021) consider a relatively narrow range of teleseismic source distances (50° and 80°) which limits the sampling of incidence angles. This adversely affects both the recovery of anisotropic fabrics—which requires good directional coverage—and the spatial resolution of the tomographic model by allowing for greater trade-off among parameters. Directional sampling can be improved simply through a joint analysis of teleseismic *P* and *S* wave observables as they sample orthogonal planes of the same anisotropic fabric. Further improvements in anisotropic model construction can come from the inclusion of steeply propagating XKS

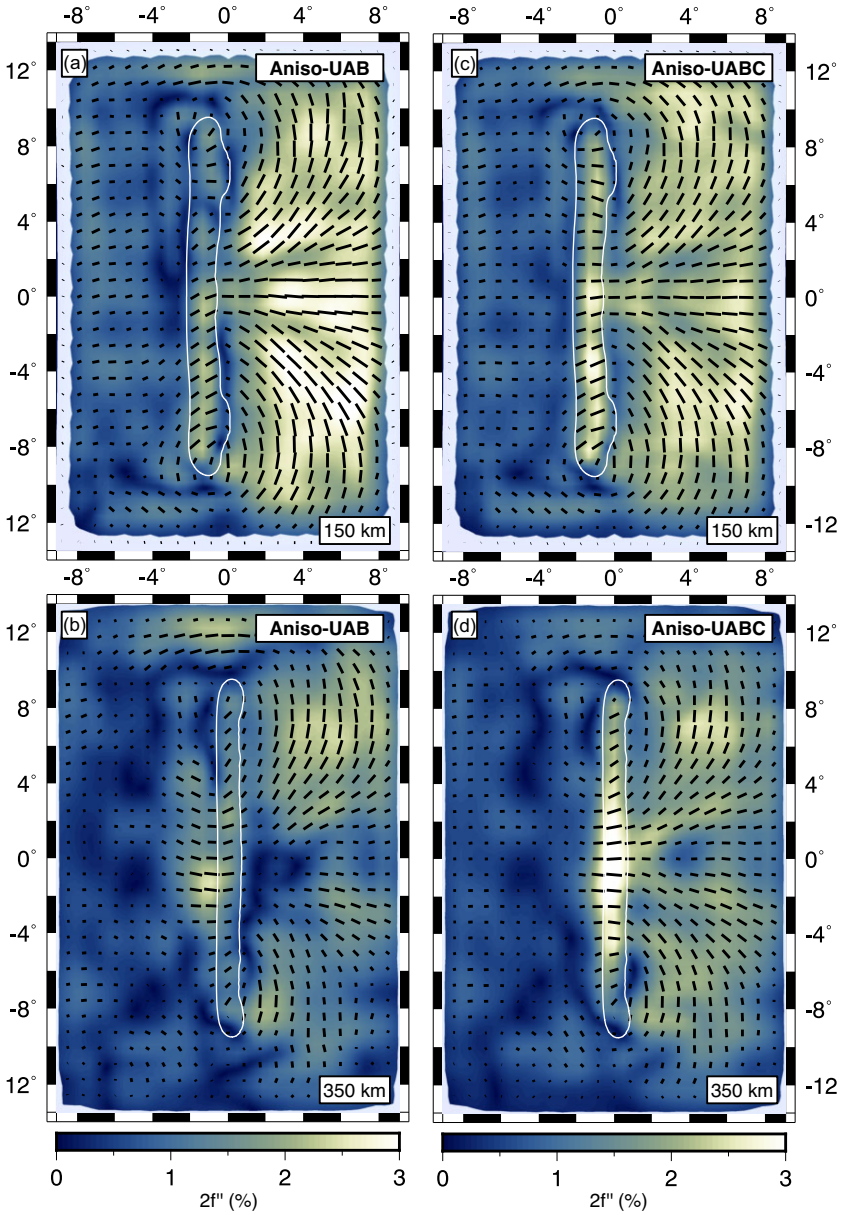

Figure 9. Recovered anisotropic heterogeneity from (a, b) the Aniso-UAB solution and (c, d) the Aniso-UABC solution are shown at 150 km (top row) and 350 km (bottom row) depth. Quivers depict the anisotropy symmetry axes scaled by f'' and projected into the cross-section plane. Colour illustrates the peak-to-peak anisotropic magnitude ($2f''$). White contour outlines the position of the subducting plate.

phases as well as more horizontal travelling local and regional seismic phases.

8.2 Implications for the interpretation of seismic images

Our results have significant implications for the interpretation of tomographic models. Although upper mantle low-velocity anomalies are ubiquitous in *P* and *S* tomographic images, anisotropy is rarely considered in velocity model construction. When anisotropy is considered in imaging, it is often subject to simplifying assumptions regarding the orientation of the symmetry axes. Bezada *et al.* (2016) and VanderBeek & Faccenda (2021) clearly demonstrated that low-velocity artefacts emerge in isotropic *P*-wave tomography due to realistic anisotropic fabrics produced by subduction zone flow fields. Importantly, these studies show that anisotropy-induced artefacts persist even when azimuthal anisotropic inversions are performed and cannot easily be corrected for using SKS splitting observations (O'Driscoll *et al.* 2011; Bezada *et al.* 2016; Mohanty *et al.* 2016; Confal *et al.* 2020). Our results confirm that such artefacts can also be expected in isotropic shear wave velocity models. For the particular imaging geometry considered here, dipping fabrics associated with the descending plate act to increase the propagation speeds of both teleseismic *P* and *S* waves while subhorizontal fabrics beneath the incoming plate and within the mantle wedge have the opposite effect. This results in a similar distribution of artefacts in isotropic tomography of these different seismic phases. The subslab low-velocity zone in our Aniso-U solution tends to be stronger at greater depths where the model is less well-constrained due to limited ray crossing. The appearance of consistent anomalies between independent data sets could mislead researchers as to

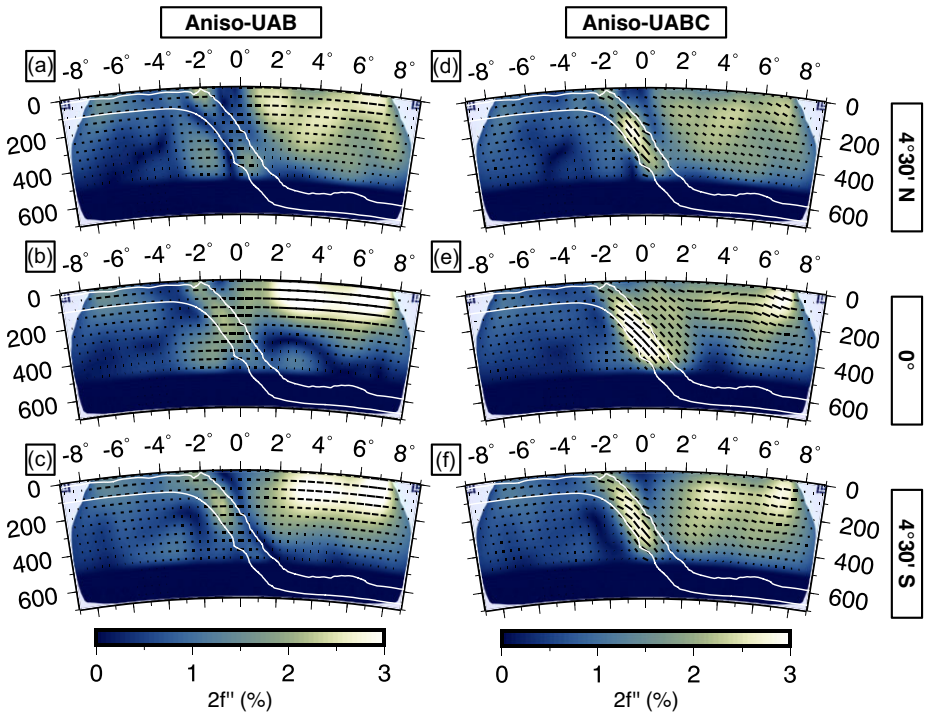

Figure 10. East-west cross-section showing the anisotropic structure recovered in the (a–c) Aniso-UAB and (d–f) Aniso-UABC solution. Cross-section are located at 4.5° (top row), 0°N (middle row) and −4.5° (bottom row). Quivers depict the anisotropy symmetry axes scaled by f'' and projected into the cross-section plane. Colour illustrates the peak-to-peak anisotropic magnitude ($2f''$). White contour outlines the position of the subducting plate.

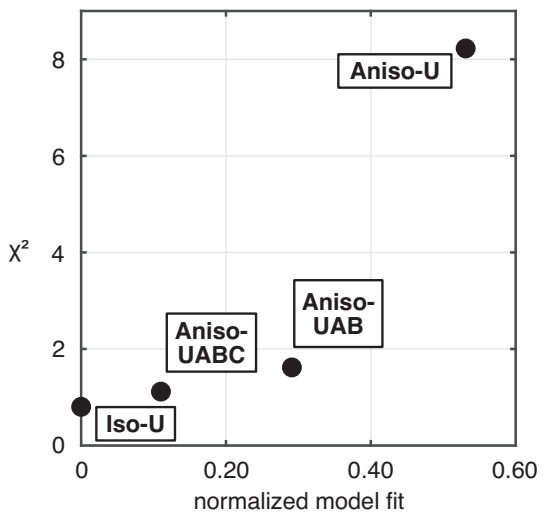

Figure 11. Comparison of isotropic model recovery. The normalized data variance ($\chi^2 = \mathbf{r}^T \mathbf{C}_d^{-1} \mathbf{r}/N$) is plotted against the normalized model fit defined by $||\mathrm{dln}V - \mathrm{dln}V_r||^2/||\mathrm{dln}V_r||^2$, where $\mathrm{dln}V$ is the recovered fractional velocity perturbations and $\mathrm{dln}V_r$ is a reference fractional velocity perturbation model chosen here to be the Iso-U solution. As a more complete description of anisotropy is included in the inversion, the data variance and isotropic heterogeneity approach the Iso-U solution.

the robustness and physical origin of such features. It is important to note that the exact nature of anisotropy-induced artefacts will depend on the array geometry and the distribution of sources and their polarization. The uniform station spacing and backazimuthal coverage considered in the present study likely acts to minimize artefacts relative to a more biased data distribution. As quantitative comparisons and integration of tomographic images becomes more routine (e.g. Shephard *et al.* 2017; Hosseini *et al.* 2018; Golos *et al.* 2020; Marignier *et al.* 2020), it will be important to understand how anisotropy could generate discrepancies among models.

Recent studies on the geodynamic significance of subslab low-velocity zones (SSLVZs) in subduction environments (e.g. Fan & Zhao 2021) is one example where neglecting anisotropy may have significant implications for isotropic model interpretation. In the Andean (Portner *et al.* 2017; Rodríguez *et al.* 2021) and Cascadia (Hawley *et al.* 2016; Bodmer *et al.* 2018) subduction zones, such low-velocity features have been variously attributed to entrainment of hotspot material, decompression melting, and return flow along the base of the subducting plate. Buoyancy forces inferred from these local velocity reductions are hypothesized to influence megathrust behaviour (Bodmer *et al.* 2018; Fan & Zhao 2021) and the topographic evolution of the forearc (Bodmer *et al.* 2020). An alternative explanation may be that such features are simply a consequence of assuming an isotropic earth in a truly anisotropic setting. This explanation is supported by the prevalence of anisotropy-induced low-velocity artefacts present in our synthetic tests as well as those of Bezada *et al.* (2016) and VanderBeek & Faccenda (2021). However, we note that the SSLVZs in Figs 8(d)–(f) and in the *P*-wave study of VanderBeek & Faccenda (2021) occur $\gtrapprox$300 km depth while those found in real data sets generally occur $\lessapprox$300 km. This may reflect differences in array geometries, source distributions, the frequency content of waveforms analysed, subjective choices in the regularization of the

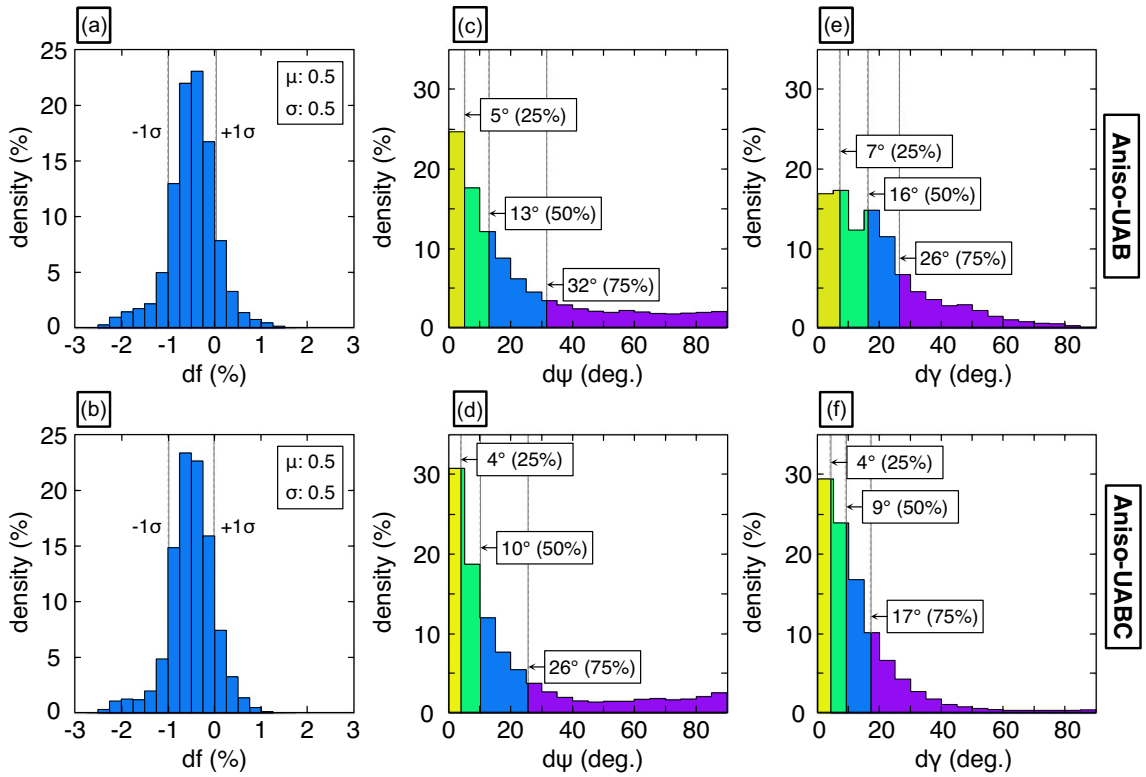

Figure 12. Anisotropy errors in the (top row) Aniso-UAB and (bottom row) Aniso-UABC solutions. Error distributions are shown for the (a, b) anisotropic fraction (f''), (c, d) symmetry axis azimuth (ψ) and (e, f) symmetry axis elevation (γ). The location of the 25-, 50- and 75-percentiles for the errors in ψ and γ are annotated and shaded.

tomographic inversions and, of course, the true nature of heterogeneity sampled. Ultimately, improving our understanding of subduction zone structure, and the nature of SSLVZs in particular, will require the application of both isotropic and anisotropic imaging techniques.

9 CONCLUSIONS

Using synthetic seismic data generated from a geodynamic model of subduction, we have demonstrated that neglecting elastic anisotropy in teleseismic shear wave imaging can introduce significant imaging artefacts. Specifically, anisotropic fabrics produced by the subduction zone flow field can yield substantial sub- and supraslab low velocity anomalies when an isotropic Earth is assumed. Such features could result in misguided interpretations regarding the physical properties and dynamics of the mantle. To reduce such artefacts, we have described and applied an anisotropic imaging method to resolve arbitrarily oriented hexagonal anisotropy from observations of relative shear wave traveltimes and splitting intensities. Not only does the new imaging methodology improve the recovery of true isotropic heterogeneity, it is able to accurately capture 3-D variations in the strength and orientation of anisotropic fabrics. Such constraints provide insights into mantle deformation processes and are valuable for constraining geodynamic models.

ACKNOWLEDGMENTS

We are grateful to Sébastien Chevrot and an anonymous reviewer for their constructive comments that improved the manuscript. This research was performed as part of the NEWTON project and supported by the European Research Council (grant 758199). Waveform modelling was performed using computational resources on the GALILEO supercomputer maintained by the CINECA HPC facility. Figures were created using Generic Mapping Tools (Wessel *et al.* 2019) with colourmaps from Crameri (2018) used for Figs 3, 9 and 10.

DATA AVAILABILITY

No new data were generated or analysed in support of this research. The tomography software developed using the methods detailed in this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abt, D.L. & Fischer, K.M., 2008. Resolving three-dimensional anisotropic structure with shear wave splitting tomography, *Geophys. J. Int.,* **173**(3), 859–886.

Abt, D.L., Fischer, K.M., Abers, G.A., Strauch, W., Protti, J.M. & Gonzalez, V., 2009. Shear wave anisotropy beneath Nicaragua and Costa Rica: implications for flow in the mantle wedge, *Geochem. Geophys. Geosyst.,* **10**(5), doi:10.1029/2009GC002375.

Aki, K., Christoffersson, A. & Husebye, E.S., 1977. Determination of the three-dimensional seismic structure of the lithosphere, *J. geophys. Res.,* **82**(2), 277–296.

Anderson, D.L., 2007. *New Theory of the Earth,* Cambridge Univ. Press.

Ando, M., Ishikawa, Y. & Yamazaki, F., 1983. Shear wave polarization anisotropy in the upper mantle beneath Honshu, Japan, *J. geophys. Res.,* **88**(B7), 5850–5864.

Aster, R.C., Borchers, B. & Thurber, C.H., 2018. *Parameter Estimation and Inverse Problems,* Elsevier.

Babuška, V., Plomerová, J. & Šílený, J., 1993. Models of seismic anisotropy in the deep continental lithosphere, *Phys. Earth planet. Inter.,* **78**(3–4), 167–191.

Becker, T.W., Chevrot, S., Schulte–Pelkum, V. & Blackman, D.K., 2006. Statistical properties of seismic anisotropy predicted by upper mantle geodynamic models, *J. geophys. Res.,* **111**(B8), doi:10.1029/2005JB004095.

Becker, T.W., 2008. Azimuthal seismic anisotropy constrains net rotation of the lithosphere, *Geophys. Res. Lett.,* **35**(5), doi:10.1029/2007GL032928.

Becker, T.W., Kustowski, B. & Ekström, G., 2008. Radial seismic anisotropy as a constraint for upper mantle rheology, *Earth planet. Sci. Lett.,* **267**(1–2), 213–227.

Becker, T.W., Conrad, C.P., Schaeffer, A.J. & Lebedev, S., 2014. Origin of azimuthal seismic anisotropy in oceanic plates and mantle, *Earth planet. Sci. Lett.,* **401,** 236–250.

Becker, T.W. & Lebedev, S., 2021. Dynamics of the upper mantle in light of seismic anisotropy, in *Mantle Convection and Surface Expressions,* pp. 257–282, eds Marquardt, H., Ballmer, M., Cottaar, S. & Konter, J., Geophysical Monograph Series, AGU.

Beller, S. & Chevrot, S., 2020. Probing depth and lateral variations of upper-mantle seismic anisotropy from full-waveform inversion of teleseismic body-waves, *Geophys. J. Int.,* **222**(1), 352–387.

Bezada, M.J., Faccenda, M. & Toomey, D.R., 2016. Representing anisotropic subduction zones with isotropic velocity models: a characterization of the problem and some steps on a possible path forward, *Geochem. Geophys. Geosyst.,* **17**(8), 3164–3189.

Blackman, D.K. & Kendall, J.M., 2002. Seismic anisotropy in the upper mantle 2. Predictions for current plate boundary flow models, *Geochem. Geophys. Geosyst.,* **3**(9), doi:10.1029/2001GC000247.

Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm, *Geophys. J. Int.,* **178**(3), 1411–1436.

Bodmer, M., Toomey, D.R., Hooft, E.E. & Schmandt, B., 2018. Buoyant asthenosphere beneath Cascadia influences megathrust segmentation, *Geophys. Res. Lett.,* **45**(14), 6954–6962.

Bodmer, M., Toomey, D.R., Roering, J.J. & Karlstrom, L., 2020. Asthenospheric buoyancy and the origin of high-relief topography along the Cascadia forearc, *Earth planet. Sci. Lett.,* **531,** doi:10.1016/j.epsl.2019.115965.

Boneh, Y., Morales, L.F., Kaminski, E. & Skemer, P., 2015. Modeling olivine CPO evolution with complex deformation histories: implications for the interpretation of seismic anisotropy in the mantle, *Geochem. Geophys. Geosyst.,* **16**(10), 3436–3455.

Browaeys, J.T. & Chevrot, S., 2004. Decomposition of the elastic tensor and geophysical applications, *Geophys. J. Int.,* **159**(2), 667–678.

Calvet, M., Chevrot, S. & Souriau, A., 2006. P-wave propagation in transversely isotropic media: I. Finite-frequency theory, *Phys. Earth planet. Inter.,* **156**(1–2), 12–20.

Chapman, C.H. & Shearer, P.M., 1989. Ray tracing in azimuthally anisotropic media—II. Quasi-shear wave coupling, *Geophys. J. Int.,* **96**(1), 65–83.

Chapman, C.T. & Pratt, R.G., 1992. Traveltime tomography in anisotropic media—I. Theory, *Geophys. J. Int.,* **109**(1), 1–19.

Chen, M. & Tromp, J., 2007. Theoretical and numerical investigations of global and regional seismic wave propagation in weakly anisotropic earth models, *Geophys. J. Int.,* **168**(3), 1130–1152.

Chevrot, S., 2000. Multichannel analysis of shear wave splitting, *J. geophys. Res.,* **105**(B9), 21 579–21 590.

Chevrot, S. & van der Hilst, R.D., 2003. On the effects of a dipping axis of symmetry on shear wave splitting measurements in a transversely isotropic medium, *Geophys. J. Int.,* **152**(2), 497–505.

Chevrot, S., Favier, N. & Komatitsch, D., 2004. Shear wave splitting in three-dimensional anisotropic media, *Geophys. J. Int.,* **159**(2), 711–720.

Chevrot, S., 2006. Finite-frequency vectorial tomography: a new method for high-resolution imaging of upper mantle anisotropy, *Geophys. J. Int.,* **165**(2), 641–657.

Chevrot, S. & Monteiller, V., 2009. principals of vectorial tomography–the effects of model parametrization and regularization in tomographic imaging of seismic anisotropy, *Geophys. J. Int.,* **179**(3), 1726–1736.

Confal, J.M., Bezada, M.J., Eken, T., Faccenda, M., Saygin, E. & Taymaz, T., 2020. Influence of upper mantle anisotropy on isotropic P-wave tomography images obtained in the Eastern Mediterranean region, *J. geophys. Res.,* **125**(8), e2019JB018559, doi:10.1029/2019JB018559.

Confal, J.M., Baccheschi, P., Pondrelli, S., Karakostas, F., VanderBeek, B.P., Huang, Z. & Faccenda, M., 2023. Reproducing complex anisotropy patterns at subduction zones from splitting intensity analysis and anisotropy tomography, *Geophys. J. Int.,* **235**(2), 1725–1735.

Conrad, C.P., Behn, M.D. & Silver, P.G., 2007. Global mantle flow and the development of seismic anisotropy: differences between the oceanic and continental upper mantle, *J. geophys. Res.,* **112**(B7), doi:10.1029/2006JB004608.

Crameri, F., 2018. Geodynamic diagnostics, scientific visualisation and StagLab 3.0, *Geosci. Model Dev.,* **11**(6), 2541–2562.

Crampin, S. & Peacock, S., 2008. A review of the current understanding of seismic shear-wave splitting in the Earth's crust and common fallacies in interpretation, *Wave Motion,* **45**(6), 675–722.

Crampin, S. & Gao, Y., 2009. A review of a quarter century of International Workshops on Seismic Anisotropy in the crust (0IWSA–12IWSA), *J. Seismol.,* **13**(2), 181–208.

Crotwell, H.P., Owens, T.J. & Ritsema, J., 1999. The TauP Toolkit: flexible seismic travel-time and ray-path utilities, *Seismol. Res. Lett.,* **70**(2), 154–160.

Dahlen, F.A., Hung, S.H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.,* **141**(1), 157–174.

Eberhart-Phillips, D. & Mark Henderson, C., 2004. Including anisotropy in 3-D velocity inversion and application to Marlborough, New Zealand, *Geophys. J. Int.,* **156**(2), 237–254.

Eilon, Z., Abers, G.A. & Gaherty, J.B., 2016. A joint inversion for shear velocity and anisotropy: the Woodlark Rift, Papua New Guinea, *Geophys. J. Int.,* **206**(2), 807–824.

Evans, R., 1984. Effects of the free surface on shear wave trains, *Geophys. J. Int.,* **76**(1), 165–172.

Faccenda, M. & Capitanio, F.A., 2013. Seismic anisotropy around subduction zones: insights from three-dimensional modeling of upper mantle deformation and SKS splitting calculations, *Geochem. Geophys. Geosyst.,* **14**(1), 243–262.

Faccenda, M., 2014. Mid mantle seismic anisotropy around subduction zones, *Phys. Earth planet. Inter.,* **227,** 1–19.

Fan, J. & Zhao, D., 2021. Subslab heterogeneity and giant megathrust earthquakes, *Nat. Geosci.,* **14**(5), 349–353.

Favier, N. & Chevrot, S., 2003. Sensitivity kernels for shear wave splitting in transverse isotropic media, *Geophys. J. Int.,* **153**(1), 213–228.

Favier, N., Chevrot, S. & Komatitsch, D., 2004. Near-field influence on shear wave splitting and traveltime sensitivity kernels, *Geophys. J. Int.,* **156**(3), 467–482.

Fichtner, A., Kennett, B.L., Igel, H. & Bunge, H.P., 2010. Full waveform tomography for radially anisotropic structure: new insights into present and past states of the Australasian upper mantle, *Earth planet. Sci. Lett.,* **290**(3–4), 270–280.

Fischer, K.M., Parmentier, E.M., Stine, A.R. & Wolf, E.R., 2000. Modeling anisotropy and plate-driven flow in the Tonga subduction zone back arc, *J. geophys. Res.,* **105**(B7), 16 181–16 191.

Flinn, E.A., 1965. Signal analysis using rectilinearity and direction of particle motion, *Proc. IEEE,* **53**(12), 1874–1876.

Golos, E.M., Fang, H. & van der Hilst, R.D., 2020. Variations in seismic wave speed and VP/VS ratio in the North American lithosphere, *J. geophys. Res.,* **125**(12), e2020JB020574, doi:10.1029/2020JB020574.

Grésillaud, A. & Cara, M., 1996. Anisotropy and P-wave tomography: a new approach for inverting teleseismic data from a dense array of stations, *Geophys. J. Int.,* **126**(1), 77–91.

Hammond, W.C. & Toomey, D.R., 2003. Seismic velocity anisotropy and heterogeneity beneath the Mantle Electromagnetic and Tomography Experiment (MELT) region of the East Pacific Rise from analysis of P and S body waves, *J. geophys. Res.,* **108**(B4), doi:10.1029/2002JB001789.

Hansen, L.N., Faccenda, M. & Warren, J.M., 2021. A review of mechanisms generating seismic anisotropy in the upper mantle, *Phys. Earth planet. Inter.,* **313,** doi:10.1016/j.pepi.2021.106662.

Hawley, W.B., Allen, R.M. & Richards, M.A., 2016. Tomography reveals buoyant asthenosphere accumulating beneath the Juan de Fuca plate, *Science,* **353**(6306), 1406–1408.

Hearn, T.M., 1996. Anisotropic Pn tomography in the western United States, *J. geophys. Res.,* **101**(B4), 8403–8414.

Hosseini, K., Matthews, K.J., Sigloch, K., Shephard, G.E., Domeier, M. & Tsekhmistrenko, M., 2018. SubMachine: web-based tools for exploring seismic tomography and other models of Earth's deep interior, *Geochem. Geophys. Geosyst.,* **19**(5), 1464–1483.

Hung, S.H., Dahlen, F.A. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes – II. Examples, *Geophys. J. Int.,* **141**(1), 175–203.

Kaminski, E., Ribe, N.M. & Browaeys, J.T., 2004. D-Rex, a program for calculation of seismic anisotropy due to crystal lattice preferred orientation in the convective upper mantle, *Geophys. J. Int.,* **158**(2), 744–752.

Karato, S.I., Jung, H., Katayama, I. & Skemer, P., 2008. Geodynamic significance of seismic anisotropy of the upper mantle: new insights from laboratory studies, *Ann. Rev. Earth planet. Sci.,* **36**(1), 59–95.

Kendall, J.M., 1994. Teleseismic arrivals at a mid-ocean ridge: effects of mantle melt and anisotropy, *Geophys. Res. Lett.,* **21**(4), 301–304.

Kennett, B.L. N. & Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.,* **105**(2), 429–465.

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

Koulakov, I., Jakovlev, A. & Luehr, B.G., 2009. Anisotropic structure beneath central Java from local earthquake tomography, *Geochem. Geophys. Geosyst.,* **10**(2), doi:10.1029/2008GC002109.

Lévêque, J.J. & Masson, F., 1999. From ACH tomographic models to absolute velocity models, *Geophys. J. Int.,* **137**(3), 621–629.

Levin, V., Menke, W. & Lerner-Lam, A., 1996. Seismic anisotropy in the north-eastern US as a source of significant teleseismic P traveltime anomalies, *Geophys. J. Int.,* **126**(2), 593–603.

Lin, Y.-P., Zhao, L. & Hung, S.-H., 2014. Full-wave multiscale anisotropy tomography in Southern California, *Geophys. Res. Lett.,* **41,** 8809–8817.

Liu, X. & Zhao, D., 2016. Seismic velocity azimuthal anisotropy of the Japan subduction zone: constraints from P and S wave traveltimes, *J. geophys. Res.,* **121**(7), 5086–5115.

Lloyd, S.M. & van der Lee, S., 2008. Influence of observed mantle anisotropy on isotropic tomographic models, *Geochem. Geophys. Geosyst.,* **9**(7), doi:10.1029/2008GC001997.

Lo Bue, R., Rappisi, F., Vanderbeek, B.P. & Faccenda, M., 2022. Tomographic image interpretation and central-western Mediterranean-like upper mantle dynamics from coupled seismological and geodynamic modeling approach, *Front. Earth Sci.,* **10,** doi:10.3389/feart.2022.884100.

Long, M.D. & van der Hilst, R.D., 2005. Upper mantle anisotropy beneath Japan from shear wave splitting, *Phys. Earth planet. Inter.,* **151**(3–4), 206–222.

Long, M.D., De Hoop, M.V. & Van Der Hilst, R.D., 2008. Wave-equation shear wave splitting tomography, *Geophys. J. Int.,* **172**(1), 311–330.

Long, M.D. & Silver, P.G., 2009. Shear wave splitting and mantle anisotropy: measurements, interpretations, and new directions, *Surv. Geophys.,* **30**(4), 407–461.

Long, M.D. & Becker, T.W., 2010. Mantle dynamics and seismic anisotropy, *Earth planet. Sci. Lett.,* **297**(3–4), 341–354.

Long, M.D., 2013. Constraints on subduction geodynamics from seismic anisotropy, *Rev. Geophys.,* **51**(1), 76–112.

Lou, X., Van Der Lee, S. & Lloyd, S., 2013. AIMBAT: a Python/Matplotlib tool for measuring teleseismic arrival times, *Seismol. Res. Lett.,* **84**(1), 85–93.

Marignier, A., Ferreira, A.M. & Kitching, T., 2020. The probability of mantle plumes in global tomographic models, *Geochem. Geophys. Geosyst.,* **21**(9), e2020GC009276, doi:10.1029/2020GC009276.

Marquering, H., Dahlen, F.A. & Nolet, G., 1999. Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox, *Geophys. J. Int.,* **137**(3), 805–815.

Masson, Y. & Romanowicz, B., 2017. Box tomography: localized imaging of remote targets buried in an unknown medium, a step forward for understanding key structures in the deep Earth, *Geophys. J. Int.,* **211**(1), 141–163.

Menke, W., 2015. Equivalent heterogeneity analysis as a tool for understanding the resolving power of anisotropic travel–time tomography, *Bull. seism. Soc. Am.,* **105**(2A), 719–733.

Mochizuki, E., 1995. Anisotropic tomography of P-wave traveltimes, *Geophys. J. Int.,* **123**(1), 297–300.

Mohanty, D.D., Singh, A., O'Driscoll, L.J., Ravi Kumar, M., Srinagesh, D. & Humphreys, E.D., 2016. P wave velocity structure below India and Tibet incorporating anisotropic delay time effects, *Geochem. Geophys. Geosyst.,* **17**(3), 725–738.

Mondal, P. & Long, M.D., 2019. A model space search approach to finite-frequency SKS splitting intensity tomography in a reduced parameter space, *Geophys. J. Int.,* **217**(1), 238–256.

Mondal, P. & Long, M.D., 2020. Strong seismic anisotropy in the deep upper mantle beneath the Cascadia backarc: constraints from probabilistic finite-frequency SKS splitting intensity tomography, *Earth planet. Sci. Lett.,* **539,** doi:10.1016/j.epsl.2020.116172.

Montagner, J.P. & Nataf, H.C., 1986. A simple method for inverting the azimuthal anisotropy of surface waves, *J. geophys. Res.,* **91**(B1), 511–520.

Montagner, J.P., 2007. 1.19 Deep earth structure-upper mantle structure: global isotropic and anisotropic elastic tomography, in *Treatise on Geophysics,* 2nd edn, pp. 559–589, ed. Schubert, G., Elsevier.

Monteiller, V. & Chevrot, S., 2010. How to make robust splitting measurements for single-station analysis and three-dimensional imaging of seismic anisotropy, *Geophys. J. Int.,* **182**(1), 311–328.

Monteiller, V. & Chevrot, S., 2011. High-resolution imaging of the deep anisotropic structure of the San Andreas Fault system beneath southern California, *Geophys. J. Int.,* **186**(2), 418–446.

Monteiller, V., Chevrot, S., Komatitsch, D. & Fuji, N., 2013. A hybrid method to compute short-period synthetic seismograms of teleseismic body waves in a 3-D regional model, *Geophys. J. Int.,* **192**(1), 230–247.

Monteiller, V., Chevrot, S., Komatitsch, D. & Wang, Y., 2015. Three-dimensional full waveform inversion of short-period teleseismic wavefields based upon the SEM–DSM hybrid method, *Geophys. J. Int.,* **202**(2), 811–827.

Monteiller, V., Beller, S., Plazolles, B. & Chevrot, S., 2021. On the validity of the planar wave approximation to compute synthetic seismograms of teleseismic body waves in a 3-D regional model, *Geophys. J. Int.,* **224**(3), 2060–2076.

Moré, J.J., 1978. The Levenberg-Marquardt algorithm: implementation and theory, in *Numerical Analysis,* pp. 105–116, Springer.

Müller, G., 1971. Approximate treatment of elastic body waves in media with spherical symmetry, *Geophys. J. Int.,* **23**(4), 435–449.

Munzarová, H., Plomerová, J. & Kissling, E., 2018. Novel anisotropic teleseismic body-wave tomography code AniTomo to illuminate heterogeneous anisotropic upper mantle: Part I—Theory and inversion tuning with realistic synthetic data, *Geophys. J. Int.,* **215**(1), 524–545.

Nicolas, A. & Christensen, N.I., 1987. Formation of anisotropy in upper mantle peridotites—a review, in *Composition, Structure and Dynamics of the Lithosphere-Asthenosphere System,* Vol. **16,** pp. 111–123, eds Fuchs, K. & Froidevaux, C., Geodynamics Series, AGU.

Nissen-Meyer, T., Driel, M.V., Stähler, S., Hosseini, K., Hempel, S., Auer, L., Colombi, A. & Fournier, A., 2014. AxiSEM: broadband 3-D seismic wavefields in axisymmetric media, *Solid Earth,* **5**(1), 425–445.

O'Driscoll, L.J., Humphreys, E.D. & Schmandt, B., 2011. Time corrections to teleseismic P delays derived from SKS splitting parameters and implications for western US P-wave tomography, *Geophys. Res. Lett.,* **38**(19), doi:10.1029/2011GL049031.

Paige, C.C. & Saunders, M.A., 1982. LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software (TOMS),* **8**(1), 43–71.

Portner, D.E., Beck, S., Zandt, G. & Scire, A., 2017. The nature of subslab slow velocity anomalies beneath South America, *Geophys. Res. Lett.,* **44**(10), 4747–4755.

Rappisi, F., VanderBeek, B.P., Faccenda, M., Morelli, A. & Molinari, I., 2022. Slab geometry and upper mantle flow patterns in the central Mediterranean from 3D anisotropic P-wave tomography, *J. geophys. Res.,* **127**(5), e2021JB023488, doi:10.1029/2021JB023488.

Rodgers, A., Krischer, L., Afanasiev, M., Boehm, C., Doody, C., Chiang, A. & Simmons, N., 2022. WUS256: an adjoint waveform tomography model of the crust and upper mantle of the western United States for improved waveform simulations, *J. geophys. Res.,* **127**(7), e2022JB024549, doi:10.1029/2022JB024549.

Rodríguez, E.E. *et al.*, 2021. Mantle dynamics of the Andean Subduction Zone from continent-scale teleseismic S-wave tomography, *Geophys. J. Int.,* **224**(3), 1553–1571.

Romanowicz, B. & Yuan, H., 2012. On the interpretation of SKS splitting measurements in the presence of several layers of anisotropy, *Geophys. J. Int.,* **188**(3), 1129–1140.

Rümpker, G. & Silver, P.G., 1998. Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy, *Geophys. J. Int.,* **135**(3), 790–800.

Ryberg, T., Rümpker, G., Haberland, C., Stromeyer, D. & Weber, M., 2005. Simultaneous inversion of shear wave splitting observations from seismic arrays, *J. geophys. Res.,* **110**(B3), doi:10.1029/2004JB003303.

Savage, M.K., 1999. Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting?, *Rev. Geophys.,* **37**(1), 65–106.

Schmandt, B. & Humphreys, E., 2010. Seismic heterogeneity and small-scale convection in the southern California upper mantle, *Geochem. Geophys. Geosyst.,* **11**(5), doi:10.1029/2010GC003042.

Shephard, G.E., Matthews, K.J., Hosseini, K. & Domeier, M., 2017. On the consistency of seismically imaged lower mantle slabs, *Sci. Rep.,* **7**(1), 1–17.

Sieminski, A., Liu, Q., Trampert, J. & Tromp, J., 2007. Finite-frequency sensitivity of body waves to anisotropy based upon adjoint methods, *Geophys. J. Int.,* **171**(1), 368–389.

Šílený, J. & Plomerová, J., 1996. Inversion of shear-wave splitting parameters to retrieve three-dimensional orientation of anisotropy in continental lithosphere, *Phys. Earth planet. Inter.,* **95**(3–4), 277–292.

Silver, P.G. & Chan, W.W., 1988. Implications for continental structure and evolution from seismic anisotropy, *Nature,* **335**(6185), 34–39.

Silver, P.G. & Chan, W.W., 1991. Shear wave splitting and subcontinental mantle deformation, *J. geophys. Res.,* **96**(B10), 16429–16454.

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**(3), 949–963.

Silver, P.G. & Long, M.D., 2011. The non-commutivity of shear wave splitting operators at low frequencies and implications for anisotropy tomography, *Geophys. J. Int.,* **184**(3), 1415–1427.

Simmons, N.A., Forte, A.M. & Grand, S.P., 2006. Constraining mantle flow with seismic and geodynamic data: a joint approach, *Earth planet. Sci. Lett.,* **246**(1–2), 109–124.

Sobolev, S.V., Grésillaud, A. & Cara, M., 1999. How robust is isotropic delay time tomography for anisotropic mantle?, *Geophys. Res. Lett.,* **26**(4), 509–512.

Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.,* **20**(2), 219–232.

Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics,* **51**(10), 1954–1966.

Toomey, D.R., Solomon, S.C. & Purdy, G.M., 1994. Tomographic imaging of the shallow crustal structure of the East Pacific Rise at 9° 30' N, *J. geophys. Res.,* **99**(B12), 24135–24157.

VanDecar, J.C. & Crosson, R.S., 1990. Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. seism. Soc. Am.,* **80**(1), 150–169.

VanderBeek, B.P. & Faccenda, M., 2021. Imaging upper mantle anisotropy with teleseismic P-wave delays: insights from tomographic reconstructions of subduction simulations, *Geophys. J. Int.,* **225**(3), 2097–2119.

Vasco, D.W., Peterson, J.E. Jr & Majer, E.L., 1995. Beyond ray tomography: Wavepaths and Fresnel volumes, *Geophysics,* **60**(6), 1790–1804.

Vecsey, L., Plomerová, J. & Babuška, V., 2008. Shear-wave splitting measurements—Problems and solutions, *Tectonophysics,* **462**(1–4), 178–196.

Vinnik, L.P., Kind, R., Kosarev, G.L. & Makeyeva, L.I., 1989. Azimuthal anisotropy in the lithosphere from observations of long-period S-waves, *Geophys. J. Int.,* **99**(3), 549–559.

Walpole, J., Wookey, J., Masters, G. & Kendall, J.M., 2014. A uniformly processed data set of SKS shear wave splitting measurements: a global investigation of upper mantle anisotropy beneath seismic stations, *Geochem. Geophys. Geosyst.,* **15**(5), 1991–2010.

Wang, W. & Becker, T.W., 2019. Upper mantle seismic anisotropy as a constraint for mantle flow and continental dynamics of the North American plate, *Earth planet. Sci. Lett.,* **514,** 143–155.

Wang, Z. & Zhao, D., 2021. 3D anisotropic structure of the Japan subduction zone, *Sci. Adv.,* **7**(4), eabc9620, doi:10.1126/sciadv.abc9620.

Wenk, H.R. & Ed., 2016. *Preferred Orientation in Deformed Metal and Rocks: An Introduction to Modern Texture Analysis,* Elsevier.

Wessel, P., Luis, J.F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W.H. & Tian, D., 2019. The generic mapping tools version 6, *Geochem. Geophys. Geosyst.,* **20**(11), 5556–5564.

Wu, H. & Lees, J.M., 1999. Cartesian parametrization of anisotropic traveltime tomography, *Geophys. J. Int.,* **137**(1), 64–80.

Wüstefeld, A. & Bokelmann, G., 2007. Null detection in shear-wave splitting measurements, *Bull. seism. Soc. Am.,* **97**(4), 1204–1211.

Zelt, C.A. & Chen, J., 2016. Frequency-dependent traveltime tomography for near-surface seismic refraction data, *Geophys. J. Int.,* **207**(1), 72–88.

Zhang, H., Liu, Y., Thurber, C. & Roecker, S., 2007. Three-dimensional shear-wave splitting tomography in the Parkfield, California, region, *Geophys. Res. Lett.,* **34**(24), doi:10.1029/2007GL031951.

Zhu, H., Komatitsch, D. & Tromp, J., 2017. Radial anisotropy of the North American upper mantle based on adjoint tomography with USArray, *Geophys. J. Int.,* **211**(1), 349–377.

Zhu, H., Yang, J. & Li, X., 2020. Azimuthal anisotropy of the North American upper mantle based on full waveform inversion, *J. geophys. Res.,* **125**(2), e2019JB018432, doi:10.1029/2019JB018432.

Zhou, Y., Nolet, G., Dahlen, F.A. & Laske, G., 2006. Global upper-mantle structure from finite-frequency surface-wave tomography, *J. geophys. Res.,* **111**(B4), doi:10.1029/2005JB003677.

Zhou, Q., Hu, J., Liu, L., Chaparro, T., Stegman, D.R. & Faccenda, M., 2018. Western US seismic anisotropy revealing complex mantle dynamics, *Earth planet. Sci. Lett.,* **500,** 156–167.

APPENDIX A: ANISOTROPY PARAMETRIZATIONS

A1 Relationship to Thomsen parameters

Thomsen (1986) gives approximate phase velocities for Sv and Sh waves in a vertical transversely isotropic (VTI) media. The expressions are valid for hexagonal symmetry with arbitrary symmetry axis orientations and we may write eq. (16) of Thomsen (1986) using the more general notation,

$$v' = \beta_T \left[1 + \frac{\alpha_T^2}{\beta_T^2} (\epsilon_T - \delta_T) \sin^2(\theta_{ph}) \cos^2(\theta_{ph}) \right], \tag{A1}$$

and

$$v'' = \beta_T \left[1 + \gamma_T \sin^2(\theta_{ph}) \right], \tag{A2}$$

where v' corresponds to the phase velocity of the shear wave polarized in the plane containing the symmetry axis and v'' is the phase velocity of the orthogonally polarized shear wave; θ_{ph} is the angle between the wavefront normal and the hexagonal symmetry axis (i.e. the phase angle). The Thomsen parameters (α_T, β_T, ϵ_T, γ_T and

δ_T) are related to the coefficients of a hexagonally symmetric elastic tensor,

$$\alpha_T = \sqrt{\frac{C_{33}}{\rho}}, \tag{A3}$$

$$\beta_T = \sqrt{\frac{C_{44}}{\rho}}, \tag{A4}$$

$$\epsilon_T = \frac{C_{11} - C_{33}}{2C_{33}}, \tag{A5}$$

$$\gamma_T = \frac{C_{66} - C_{44}}{2C_{44}}, \tag{A6}$$

and

$$\delta = \frac{(C_{13} + C_{44})^2 - (C_{33} - C_{44})^2}{2C_{33}(C_{33} - C_{44})}. \tag{A7}$$

Expanding the $\cos^2(\theta_{ph})$ and $\sin^2(\theta_{ph})$ terms in eqs (A1) and (A2) as functions of $\cos(2\theta_{ph})$ and $\cos(4\theta_{ph})$ yields,

$$v' = \overline{v}' \left[1 + f' \cos(4\theta_{ph})\right], \tag{A8}$$

and

$$v'' = v \left[1 + f'' \cos(2\theta_{ph})\right], \tag{A9}$$

where the mean velocities ($\overline{v}'$, v) and anisotropic fractions (f', f'') are the following combinations of the Thomsen parameters,

$$\overline{v}' = \beta_T \left[1 + \frac{\alpha_T^2}{\beta_T^2} \frac{(\epsilon_T - \delta_T)}{8}\right], \tag{A10}$$

$$v = \beta_T \left[1 + \frac{\gamma_T}{2}\right], \tag{A11}$$

$$f' = -\frac{\alpha_T^2(\epsilon_T - \delta_T)}{8\beta_T^2 + \alpha_T^2(\epsilon_T - \delta_T)}, \tag{A12}$$

and

$$f'' = -\frac{\gamma_T}{(2 + \gamma_T)}. \tag{A13}$$

Given that $v(\theta_{ph} = 0) = v'(\theta_{ph} = 0)$, we can substitute $\overline{v}' = v(1 + f'')/(1 + f')$ into eq. (A9) removing the dependence on $\overline{v}'$,

$$v' = v \frac{(1 + f'')}{(1 + f')} \left[1 + f' \cos(4\theta_{ph})\right]. \tag{A14}$$

In this way, only three parameters (v, f' and f'') are required to characterize shear wave anisotropy instead of the five Thomsen parameters.

A2 Parametrization using A, B and C terms

Before presenting the qS-wave velocity equations and their derivatives using the *ABC* parametrization, it's useful to define the relationships that convert between *ABC*, vectoral (n_1, n_2, n_3) and spherical (f'', ψ, γ) representations of the anisotropic strength and symmetry axis; see Table A1. These will be used in the following derivations when convenient to simplify the equations.

To derive explicit expressions for the slowness partial derivatives with respect to the u, A, B and C parameters (eqs A23 and A24), we first define the quasi-shear wave velocities as a function of these variables. Expanding $\cos(2\alpha)$ and $\cos(4\alpha)$ terms in eqs (13) and (14) into functions of $\cos(\alpha)$ and noting that $\cos(\alpha)$ is equivalent to the dot product between the ray ($\hat{\mathbf{r}}$) and symmetry axis ($\hat{\mathbf{n}} = \mathbf{n}/\sqrt{|f''|}$) unit vectors yields

$$v'' = v \left[1 + 2f''(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})^2 - f''\right] \tag{A15}$$

Table A1. Relationships between the *ABC*, vectoral and spherical parametrizations of anisotropic strength and orientation. In equations involving A, B, and C terms we use G for $\sqrt{A^2 + B^2}$. Under the *ABC* parametrization, the symmetry axes $\mathbf{n}$ and $-\mathbf{n}$ are equivalent. For this reason, the expressions for $n_1(A, B, C)$ and $n_2(A, B, C)$ contain the '$\pm_{1,2}$' symbol which take the sign of $n_{1,2}$. The appropriate signs can always be recovered from the *ABC* parametrization by first converting to spherical parameter and then vectoral parameter. The sign of f'' is lost in the *ABC* and vectoral parametrization hence the '$\pm$' symbol in the expressions for f''.

ABC	f'', ψ, γ	n_1, n_2, n_3
A	$\lvert f''\rvert \cos^2(\gamma)\cos(2\psi)$	$n_1^2 - n_2^2$
B	$\lvert f''\rvert \cos^2(\gamma)\sin(2\psi)$	$2n_1n_2$
C	$\sqrt{\lvert f''\rvert}\sin(\gamma)$	n_3
Vectoral	f'', ψ, γ	A, B, C
n_1	$\sqrt{\lvert f''\rvert}\cos(\gamma)\cos(\psi)$	$\pm_1\sqrt{(G + A)/2}$
n_2	$\sqrt{\lvert f''\rvert}\cos(\gamma)\sin(\psi)$	$\pm_2\sqrt{(G - A)/2}$
n_3	$\sqrt{\lvert f''\rvert}\sin(\gamma)$	C
Spherical	n_1, n_2, n_3	A, B, C
f''	$\pm\left(n_1^2 + n_2^2 + n_3^2\right)$	$\pm(G + C^2)$
ψ	$\arctan_2[n_2/n_1]$	$0.5\arctan_2[B/A]$
γ	$\arctan_2\left[n_3/\sqrt{n_1^2 + n_2^2}\right]$	$\arctan_2\left[C/\sqrt{G}\right]$

and

$$v' = v \frac{(1 + f'')}{(1 + f')} \left[1 + 8f'(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})^4 - 8f'(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})^2 + f'\right], \tag{A16}$$

To keep the velocity equations compact and make the derivation of the partial derivatives more tractable, we will introduce several intermediate terms. Given the relationships in Table A1 and substituting $X = 2|f''|(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})^2$, eqs (A15) and (A16) may be written as,

$$v'' = v \left[1 \pm \left(X - G - C^2\right)\right] \tag{A17}$$

and

$$v' = v \frac{1 \pm \left(G + C^2\right)}{1 \pm h\left(G + C^2\right)} \left[1 \pm h\left(2\frac{X^2}{G + C^2} - 4X + G + C^2\right)\right], \tag{A18}$$

where $h = f'/f''$ is selected based on prior knowledge of the imaging target and assumed constant. The sign of the anisotropic fraction, f'', has been factored out resulting in the $\pm$ sign which must be chosen *a priori* given the expected origin of anisotropy (e.g. olivine or fracture alignment). Finally, we have introduced the G term defined as,

$$G = \left(A^2 + B^2\right)^{1/2}. \tag{A19}$$

Using the definitions in Table A1, the dot product in $X = 2|f''|(\hat{\mathbf{r}} \cdot \hat{\mathbf{n}})^2$ can be expanded and written as a function of A, B and C,

$$X = A\left(r_1^2 - r_2^2\right) + 2Br_1r_2 + G\left(r_1^2 + r_2^2\right) + 2\sqrt{2}CYr_3 + 2C^2r_3^2, \tag{A20}$$

where

$$Y = \pm_1(G + A)^{1/2}r_1 \pm_2 (G - A)^{1/2}r_2. \tag{A21}$$

By design, the *ABC*-parametrization cannot distinguish between $\mathbf{n}$ and $-\mathbf{n}$ resulting in the sign ambiguity represented by $\pm_{1,2}$ which takes the sign of the $n_{1,2}$ components of the symmetry axis vector. However, these sign ambiguities will be resolved when deriving the partial derivatives—the only calculations for which eqs (A17)–(A21) are used.

A3 Anisotropic slowness partial derivatives

The elements of the Jacobian, **J**, in eq. (25) are computed by differentiating the residual vector **r** with respect to the current model. The model **m** used to predict the observations and the perturbation model $\Delta\mathbf{m}$ defining the resolution of the solution need not be defined on the same grid in which case the Jacobian elements are given by the linear mapping,

$$J_{ik} = \sum_{j=1}^{M} w_{jk} \frac{\partial r_i}{\partial m_j} \tag{A22}$$

where w_{jk} is the linear interpolation weight between the jth-model and kth-perturbational node (e.g. Toomey *et al.* 1994). From eqs (17) and (18), the partial derivatives, $\partial r_i/\partial m_j$, for the traveltime and splitting intensity are,

$$\frac{\partial t_i}{\partial m_j} = K_j \left[\frac{\partial u''_j}{\partial m_j} + \left(\frac{\partial u'_j}{\partial m_j} - \frac{\partial u''_j}{\partial m_j} \right) \cos^2(\beta_j) - \frac{\partial \beta_j}{\partial m_j} \left(u'_j - u''_j \right) \sin(2\beta_j) \right] \tag{A23}$$

and

$$\frac{\partial s_i}{\partial m_j} = \frac{K_j}{2} \left[\left(\frac{\partial u''_j}{\partial m_j} - \frac{\partial u'_j}{\partial m_j} \right) \sin(2\beta_j) + 2 \frac{\partial \beta_j}{\partial m_j} (u''_j - u'_j) \cos(2\beta_j) \right]. \tag{A24}$$

Note that replacing K_j with the ray segment length ΔL_j yields the ray-theoretical partials.

The slowness partial derivatives in eqs (A23) and (A24) are constructed via the chain rule,

$$\frac{\partial u''}{\partial m} = \frac{\partial u''}{\partial v''} \frac{\partial v''}{\partial m} = -\frac{\partial v''}{\partial m} (v'')^{-2} \tag{A25}$$

and

$$\frac{\partial u'}{\partial m} = \frac{\partial u'}{\partial v'} \frac{\partial v'}{\partial m} = -\frac{\partial v'}{\partial m} (v')^{-2}. \tag{A26}$$

From eqs (A26)–(A25) and (A17)–(A18), the u'' and u' partial derivatives are,

$$\frac{\partial u''}{\partial u} = \frac{u''}{u}, \tag{A27}$$

$$\frac{\partial u''}{\partial A,\, B} = \mp \frac{(u'')^2}{u} \left(\frac{\partial X}{\partial A,\, B} - \frac{\partial G}{\partial A,\, B} \right), \tag{A28}$$

$$\frac{\partial u''}{\partial C} = \mp \frac{(u'')^2}{u} \left(\frac{\partial X}{\partial C} - 2C \right), \tag{A29}$$

$$\frac{\partial u'}{\partial u} = \frac{u'}{u}, \tag{A30}$$

$$\frac{\partial u'}{\partial A,\, B} = \mp u' \left\{ h \frac{u'}{u} \frac{(1+f'')}{(1+f')} \left([8\cos^2(\alpha) - 4] \frac{\partial X}{\partial A,\, B} + [1 - 8\cos^4(\alpha)] \frac{\partial G}{\partial A,\, B} \right) + \left(\frac{1}{(1+f'')} - \frac{h}{(1+f')} \right) \frac{\partial G}{\partial A,\, B} \right\} \tag{A31}$$

and

$$\frac{\partial u'}{\partial C} = \mp 2u' \left\{ h \frac{u'}{u} \frac{(1+f'')}{(1+f')} \left([4\cos^2(\alpha) - 2] \frac{\partial X}{\partial C} + [1 - 8\cos^4(\alpha)] C \right) + \left(\frac{1}{(1+f'')} - \frac{h}{(1+f')} \right) C \right\} \tag{A32}$$

In the above expressions, substitutions from Table A1 for terms involving X, G, A, B and C were made where convenient. The partial derivatives for the intermediate variable G (eq. A19) are,

$$\frac{\partial G}{\partial A} = \frac{A}{\sqrt{A^2 + B^2}} = \cos(2\psi), \tag{A33}$$

and

$$\frac{\partial G}{\partial B} = \frac{B}{\sqrt{A^2 + B^2}} = \sin(2\psi). \tag{A34}$$

For the Y variable (eq. A21) the partial derivatives are,

$$\frac{\partial Y}{\partial A} = \pm_1 \frac{(\partial G/\partial A + 1)}{2\sqrt{G + A}} r_1 \pm_2 \frac{(\partial G/\partial A - 1)}{2\sqrt{G + A}} r_2 = \frac{[r_1 \cos(\psi) - r_2 \sin(\psi)]}{\sqrt{2|f''|}\cos(\gamma)} \tag{A35}$$

and

$$\frac{\partial Y}{\partial B} = \pm_1 \frac{(\partial G/\partial B)}{2\sqrt{G + A}} r_1 \pm_2 \frac{(\partial G/\partial B)}{2\sqrt{G + A}} r_2 = \frac{[r_1 \sin(\psi) + r_2 \cos(\psi)]}{\sqrt{2|f''|}\cos(\gamma)} \tag{A36}$$

Differentiating the X term (eq. A20) is trivial given the expressions in eqs (A33)–(A36) and after replacing the ray unit vector ($\hat{\mathbf{r}}$) with its spherical representation (ϕ, θ) and some algebraic manipulation yields,

$$\frac{\partial X}{\partial A} = [\cos(2\psi) + \cos(2\phi)] \cos^2(\theta) + \cos(\psi + \phi) \tan(\gamma) \sin(2\theta), \tag{A37}$$

$$\frac{\partial X}{\partial B} = [\sin(2\psi) + \sin(2\phi)] \cos^2(\theta) + \sin(\psi + \phi) \tan(\gamma) \sin(2\theta), \tag{A38}$$

and

$$\frac{\partial X}{\partial C} = 4C \sin^2(\theta) + 2\sqrt{|f''|} \cos(\psi - \phi) \cos(\gamma) \sin(2\theta). \tag{A39}$$

The appearance of the $\tan(\gamma)$ term in eqs (A37) and (A38) causes the partials to become undefined when the symmetry axis is vertical ($\gamma = \pm 90^\circ$). We avoid this issue simply by forcing $|\gamma| \leq 85^\circ$ when evaluating these expressions.

The last ingredient required to compute the traveltime and splitting intensity sensitivities (eqs A23–A24) are the partial derivatives of the angle β. A rotation from the reference coordinate system to the ray-aligned coordinate system (Fig. 1) gives the QTL-components of the symmetry axis vector as,

$$\mathbf{g} = \begin{bmatrix} n_1 \sin(\theta)\cos(\phi) + n_2 \sin(\theta)\sin(\phi) - n_3 \cos(\theta) \\ -n_1 \sin(\phi) + n_2 \cos(\phi) \\ n_1 \cos(\theta)\cos(\phi) + n_2 \cos(\theta)\sin(\phi) - n_3 \sin(\theta). \end{bmatrix} \tag{A40}$$

The angle β is then,

$$\beta = \arctan\left[\frac{g_2}{g_1} \right] - \zeta, \tag{A41}$$

and from the chain rule the partial derivatives can be written as,

$$\frac{\partial \beta}{\partial m} = \frac{1}{(g_1^2 + g_2^2)} \left[g_1 \frac{\partial g_2}{\partial n_2} \frac{\partial n_2}{\partial m} - g_2 \frac{\partial g_1}{\partial n_1} \frac{\partial n_1}{\partial m} \right], \tag{A42}$$

where the $\partial n_i/\partial m$ terms can be determined from the relationships in Table A1. Evaluating eq. (A42) for $m = A$, B and C gives,

$$\frac{\partial \beta}{\partial A} = \frac{-[\sin(\theta)\sin(2\psi) - \cos(\theta)\tan(\gamma)\sin(\psi + \phi)]}{2|f''|\sin^2(\alpha)}, \tag{A43}$$

$$\frac{\partial \beta}{\partial B} = \frac{[\sin(\theta)\cos(2\psi) - \cos(\theta)\tan(\gamma)\cos(\psi + \phi)]}{2|f''|\sin^2(\alpha)}, \tag{A44}$$

and

$$\frac{\partial \beta}{\partial C} = \frac{\cos(\theta)\cos(\gamma)\sin(\psi - \phi)}{\sqrt{|f''|}\sin^2(\alpha)}. \tag{A45}$$

Upon inspection of eqs (A43)–(A45) it's clear that these expression become undefined as (1) the angle γ approaches 90°, (2) the angle α approaches 0°, and/or (3) f'' approaches zero. As previously mentioned, the first issue is avoided simply by restricting γ to the interval $\pm 85^\circ$ when evaluating the partial derivatives. To resolve the last two issues we note that the partial derivatives of β are always multiplied by $\pm(u'' - u')$ in eqs (A23) and (A24). Furthermore, we

can assume that $hf'' << 1$ allowing us to approximate the difference between u' and u'' as,

$$(u'' - u') \approx \frac{2f''u_i \sin^2(\alpha)[1 - 4h\cos^2(\alpha)]}{[1 + f''\cos(2\alpha)][1 + f'']}. \quad \text{(A46)}$$

Upon multiplication of eqs (A43)–(A45) with this approximation for $(u' - u'')$, the problematic division by terms involving f'' and $\sin(\alpha)$ is resolved.

Eqs (A27)–(A32), (A37)–(A39) and (A43)–(A46) provide the necessary ingredients to evaluate the partials derivatives in eqs (A23)–(A24) required for the construction of **J** (eq. A22).

APPENDIX B: EVALUATION OF ASSUMPTIONS IN DERIVATION OF SHEAR WAVE OBSERVABLES

To assess when the assumptions of weak splitting and linear polarization may fail, we compare anisotropic shear wave observables measured from waveforms modelled using eq. (1)–(3) to predictions from eqs (6)–(7). The matrix propagation method (i.e. eqs 1–3) of Rümpker and Silver (1998) is used to rotate and delay a Ricker wavelet as it propagates vertically through a horizontally layered anisotropic medium. The effects of 3-D heterogeneity and frequency-dependent sensitivity are neglected in this geometric treatment of splitting. Each anisotropic layer is defined by a splitting amplitude, $|\Delta t'' - \Delta t'|$, and a fast-polarization azimuth, ψ'. The anisotropic delay time is measured via cross-correlation with the initial Ricker wavelet on the x_1-component and the splitting intensity is measured by projecting the derivative of the x_1-component onto the x_2-component of the split waveform (Fig. 1; eq. 29). In this simple modelling framework and a given vertical distribution of ψ', the results depend only on the ratio between the dominant period, T, and the vertically integrated splitting amplitude. Thus, the results presented are normalized by T.

We first consider a homogeneous medium characterized by weak ($|\Delta t'' - \Delta t'| = T/10$), intermediate ($|\Delta t'' - \Delta t'| = T/5$), and strong splitting amplitude ($|\Delta t'' - \Delta t'| = T/2.5$) with $\psi' = 0^\circ$. The measured and predicted delay times and splitting intensities are shown in Fig. B1. As expected, the error in the observables grows with increasing magnitude of splitting intensity. However, provided that $|T/(\Delta t'' - \Delta t')| \gtrapprox 5$, the prediction errors are an order of magnitude smaller than the measured delay time or splitting intensity (Fig. B2). We also observe at high splitting amplitudes, the sinusoidal trend in the delay times becomes more complex though its phase and amplitude remain consistent with the prediction of eq. (6). In comparison, the sinusoidal trend in the splitting intensity agrees well with eq. (7) but the amplitude is overpredicted. This is due to the decreasing overlap between x_1- and x_2-component waveforms with increasing splitting intensity.

In a heterogeneous medium, the particle motion and principal polarization direction of the shear wave will be continually modified as it interacts with different anisotropic layers. To explore this effect we considered two layered anisotropic models each containing 10 intervals. In the first model (Figs B3a and b), the splitting amplitude of each layer is the same and scaled such that the integrated $|\Delta t'' - \Delta t'|$ is $T/5$. The fast polarization azimuth varies linearly from 0° at the base to 60° at the surface. In comparison to the homogeneous case with comparable splitting amplitude (Figs B1c and d), the amplitude of the sinusoidal trends in delay time and splitting intensity are reduced as each layer tends to destructively interfere and the phase of the trends corresponds to the average ψ' within the model (30°). The weaker anisotropic signal results in comparable but slightly reduced prediction errors for the stratified model (Figs B4a,b) compared to the homogeneous case (Figs B2c and d).

In the second model (Figs B3c and d), random splitting amplitudes and polarization azimuths are assigned to the layers. The less coherent anisotropic structure will generally result in more destructive interference and a weaker integrated splitting effect. Therefore, we scaled the splitting amplitude of all layers such that the maximum integrated splitting intensity is $5T$. In comparison to the homogeneous case (Figs B1c and d), we see a moderate increase in the prediction errors but the absolute error remains generally one order of magnitude less than the measured values (Figs B4c and d). The increased errors reflect the fact that for an N-layer model the splitting amplitude of a given interval is generally greater than $5T/N$ to counteract the destructive interference and result in the desired integrated splitting intensity. Consequently, the polarization of the S-waveform is more strongly perturbed and deviates from our assumption of linearity and constant ζ. Nonetheless, eqs (6)–(7) appear to be applicable for strongly varying anisotropic structure along a ray path.

The two previous examples considered a linearly polarized wave entering an anisotropic model. A particularly relevant case for regional teleseismic tomography is when the incoming wave exhibits significant deviations from linear polarization due to splitting by anisotropic structure outside the local imaging volume. To simulate this situation, we split the initial Ricker wavelet following eq. (1) using $|\Delta t'' - \Delta t'| = T/5$ and a fast-polarization azimuth rotated 22.5° and 45° from the initial linear polarization direction (henceforth referred to as waveform-22 and waveform-45, respectively). As a result of the initial splitting, waveform-22 (Figs B5a and b) exhibits clear elliptical particle motion with a principal polarization direction rotated $\sim 16^\circ$ from the initial linear polarization direction. The waveform-45 (Figs B5c and d) has nearly circular particle motion with the principal polarization direction remaining at 0°.

Despite both waveforms not conforming to the assumption of linear polarization, our approximations still provide an accurate description of the principal delay and splitting intensity (Figs B6 and B7). However, there is a mean offset between predicted and measured splitting intensity. In the case of regional teleseismic imaging, this offset can be accounted for by demeaning the splitting intensity measurements for each event as is common practice for delay time measurements. There is one other important caveat; to correctly predict the phase of the sinusoidal trends shown in the measurements, the correct incoming principal polarization direction must be known (Figs B5b and d). For a regional teleseismic study, this can be obtained via a principal component analysis of the waveform generated by stacking the S-phases observed across an array for a given event—the approach used in our synthetic tomography (Section 5.1).

In summary, we have investigated the accuracy of eqs (6)–(7) in response to increasing splitting strength of homogenous and layered medium and to elliptical particle motion. These tests confirm that provided the integrated splitting intensity remains less than $\sim 5T$, the error in our expressions is generally an order of magnitude smaller than the measured anisotropic observable. Furthermore, provided the principal polarization direction of the incoming wave is known, the splitting intensity and delay time of waveforms exhibiting strongly elliptical particle motions can be reliably estimated. These tests did not consider the effects of more complicated wave phenomenon such as scattering and mode conversions from 3-D heterogeneity. The potential error from such processes is discussed in Section 5.2 where we compare measured principal shear

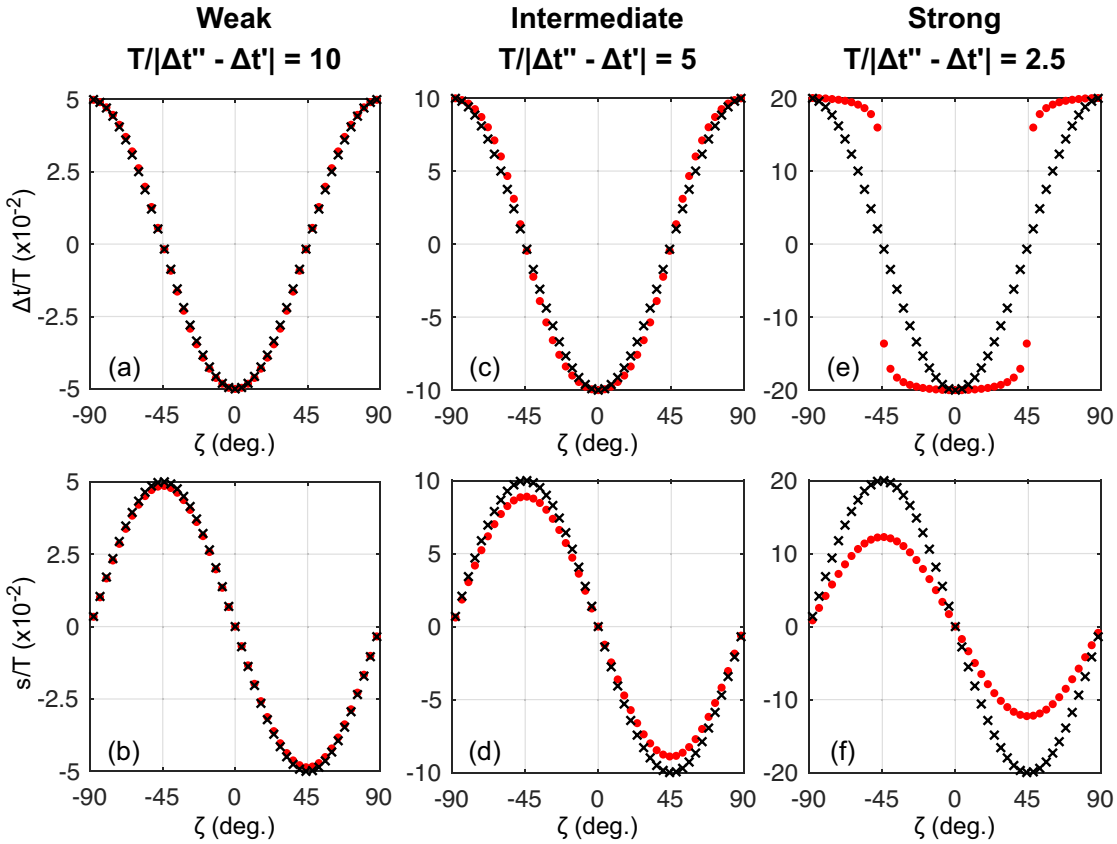

Figure B1. Predicted (black crosses) and measured (red points) principal anisotropic delay (top row) and splitting intensity (bottom row) as a function of the incoming *S*-wave polarization azimuth, ζ, for a homogeneous medium. Results are shown for three models characterized by (a, b) weak, (c, d) intermediate and (e, f) strong anisotropy with a fast-polarization direction corresponding to $\zeta = 0$. The strength of anisotropy is defined by the delay time between the fast- and slow-polarized quasi-shear phases ($|\Delta t'' - \Delta t'|$) relative to the dominant waveform period (T). Note that plotted delay times and and splitting intensities are normalized by T.

wave observables from SPECFEM simulated waveforms to those predicted by our approximations.

APPENDIX C: MINIMIZATION OF ANISOTROPIC MAGNITUDE

Here we describe the construction the (P x M) matrix $\mathbf{D}_f$ in eq. (25) that minimizes the perturbation to the anisotropic magnitude. For the perturbational model vector $\Delta\mathbf{m} = [\Delta\mathbf{u}; \Delta\mathbf{A}; \Delta\mathbf{B}; \Delta\mathbf{C}]$ where each parameter set is defined on a grid of P-nodes discretized in space for a total of M = 4P unknowns, the anisotropic magnitude perturbation at the pth node location is,

$$\Delta f''_p = \sqrt{\Delta m^2_{p+\mathrm{P}} + \Delta m^2_{p+2\mathrm{P}}} + \Delta m^2_{p+3\mathrm{P}}. \tag{C1}$$

Taking the total derivative of $\Delta f''_p$ gives the coefficients of the pth-row of the matrix $\mathbf{D}_f$,

$$d_{p,p+\mathrm{P}} = \frac{\Delta m_{p+\mathrm{P}}}{\sqrt{\Delta m^2_{p+\mathrm{P}} + \Delta m^2_{p+2\mathrm{P}}}} = \cos(2\Delta\psi_p), \tag{C2}$$

$$d_{p,p+2\mathrm{P}} = \frac{\Delta m_{p+2\mathrm{P}}}{\sqrt{\Delta m^2_{p+\mathrm{P}} + \Delta m^2_{p+2\mathrm{P}}}} = \sin(2\Delta\psi_p), \tag{C3}$$

and

$$d_{p,p+3\mathrm{P}} = 2\Delta m_{p+3\mathrm{P}}; \tag{C4}$$

all other elements are zero-valued. At each iteration $\mathbf{D}_f$ is computed for the current value of $\Delta\mathbf{m}$.

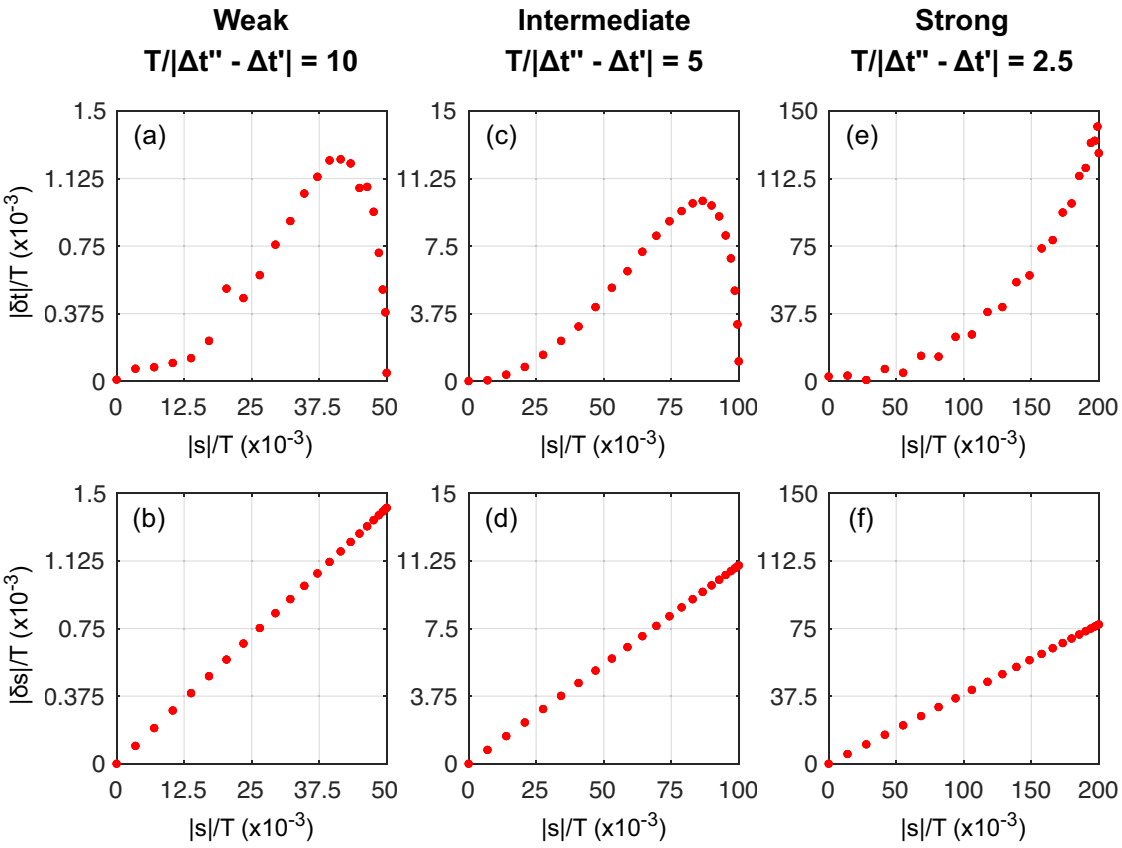

Figure B2. Absolute errors between the predicted and measured (top row) delay times and (bottom row) splitting intensities shown in Fig. B1. Errors are plotted as a function of the predicted splitting intensity magnitude (eq. 7).

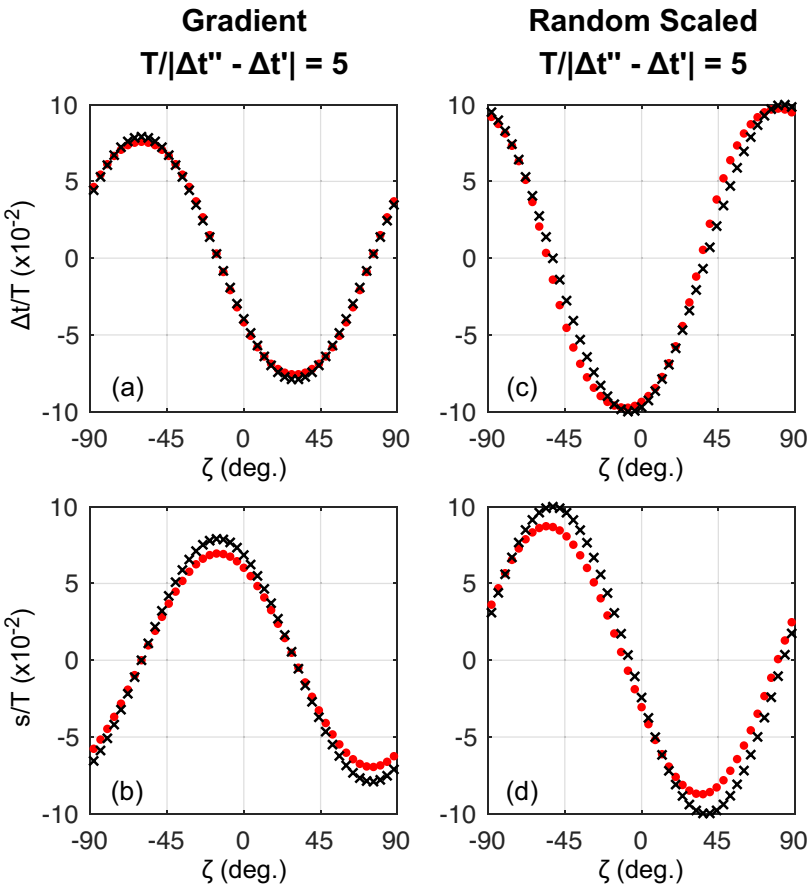

Figure B3. Predicted (black crosses) and measured (red points) principal anisotropic delay (top row) and splitting intensity (bottom row) as a function of the incoming *S*-wave polarization azimuth, ζ, for a stratified medium composed of 10 layers. In (a, b) results are shown for a model with a linear gradient in the fast-polarization azimuth varying from $\zeta = 0^\circ$ at the base to $\zeta = 60^\circ$ at the surface. The anisotropic strength is the same for each layer and defined such that the integrated delay between the fast- and slow-polarized quasi-shear phases, $|\Delta t'' - \Delta t'|$, is $T/5$, where T is the waveform period. The model corresponding to panels (c, d) contains layers with a randomly assigned fast-polarization direction and anisotropic strength. In contrast to the linear gradient model, the anisotropic strength of all layers is scaled such that the maximum integrated splitting intensity is $T/10$. Note that delay times and and splitting intensities are normalized by T.

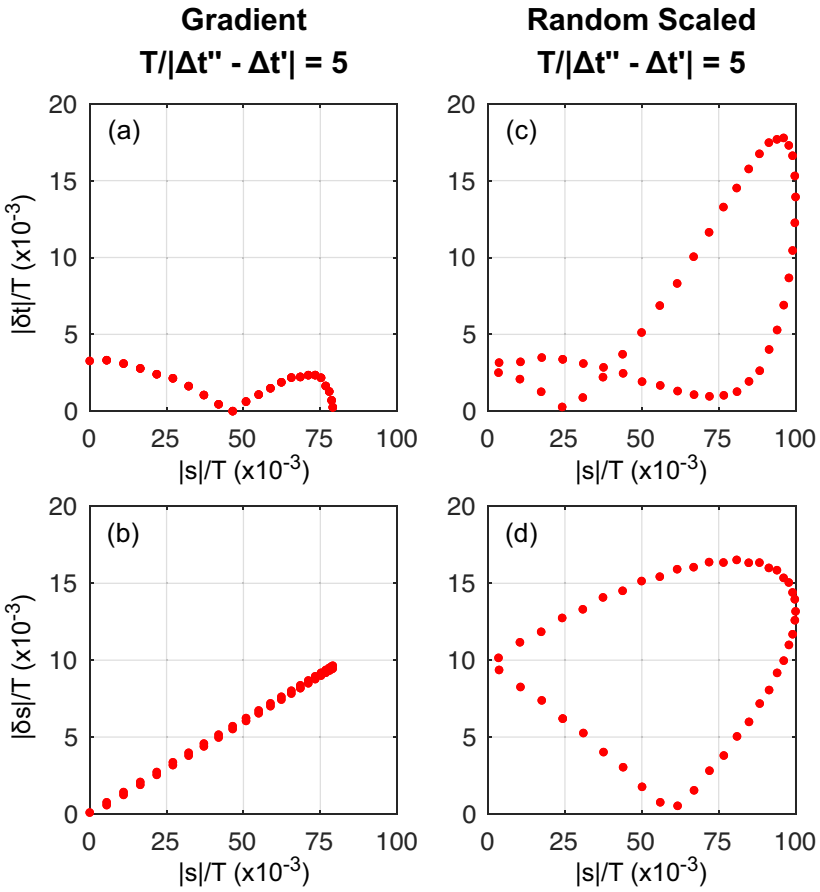

Figure B4. Absolute errors between the predicted and measured (top row) delay times and (bottom row) splitting intensities shown in Fig. A3. Errors are plotted as a function of the predicted splitting intensity magnitude (eq. 7).

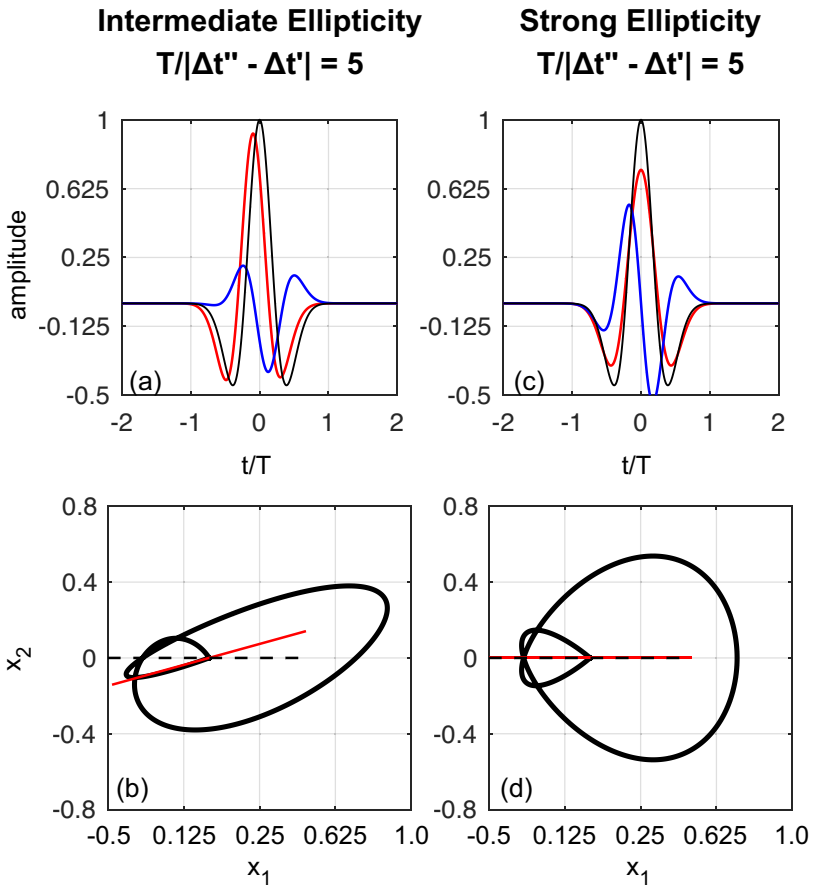

Figure B5. Waveforms exhibiting an (a) intermediate and (b) high degree of ellipticity. Black curve corresponds to the initial linearly polarized Ricker wavelet; red curve shows the split waveform observed in the principal polarization direction and the blue waveform is the displacement in the orthogonal direction. Note the time axis is normalized by the waveform period, *T*. Particle motion diagrams for the waveforms in (a) and (b) are shown in (c) and (d), respectively. In (c,d), the dashed black line depicts the linear polarization direction of the initial Ricker wavelet while the red line depicts the principal polarization direction of the split waveform determined via eigendecomposition of the trace covariance matrix.

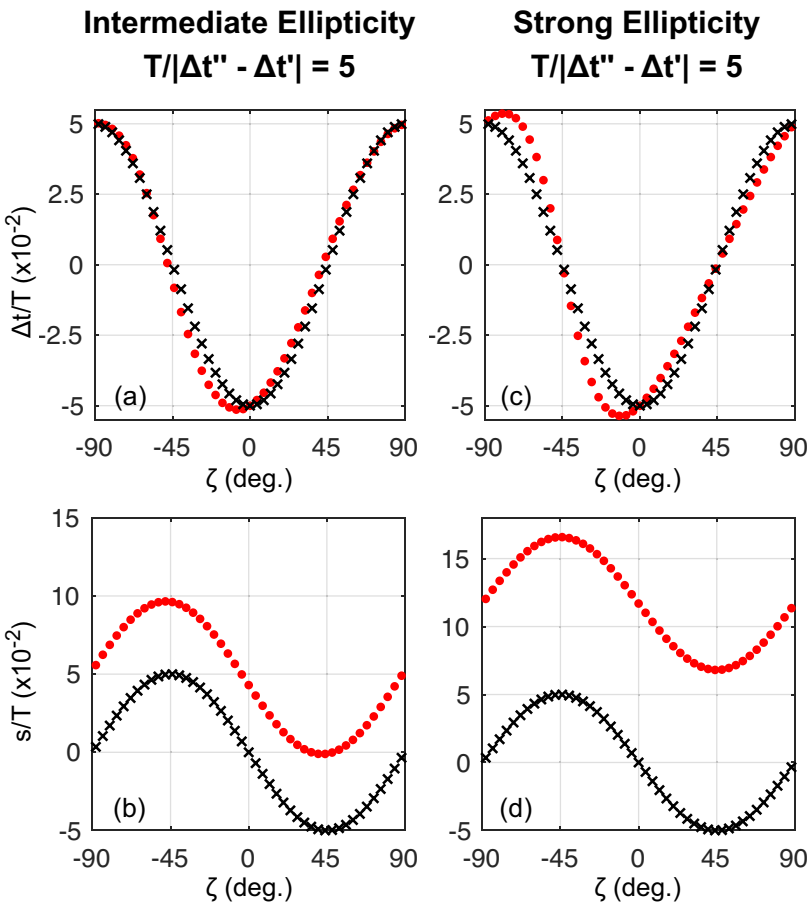

Figure B6. Predicted (black crosses) and measured (red points) principal anisotropic delay (top row) and splitting intensity (bottom row) as a function of the incoming *S*-wave principal polarization azimuth, ζ, for the elliptically polarized waveforms shown in Fig. A5. The anisotropic model is the same as that described in Figs B1(c) and (d). Note that delay times and and splitting intensities are normalized by the waveform period, *T*.

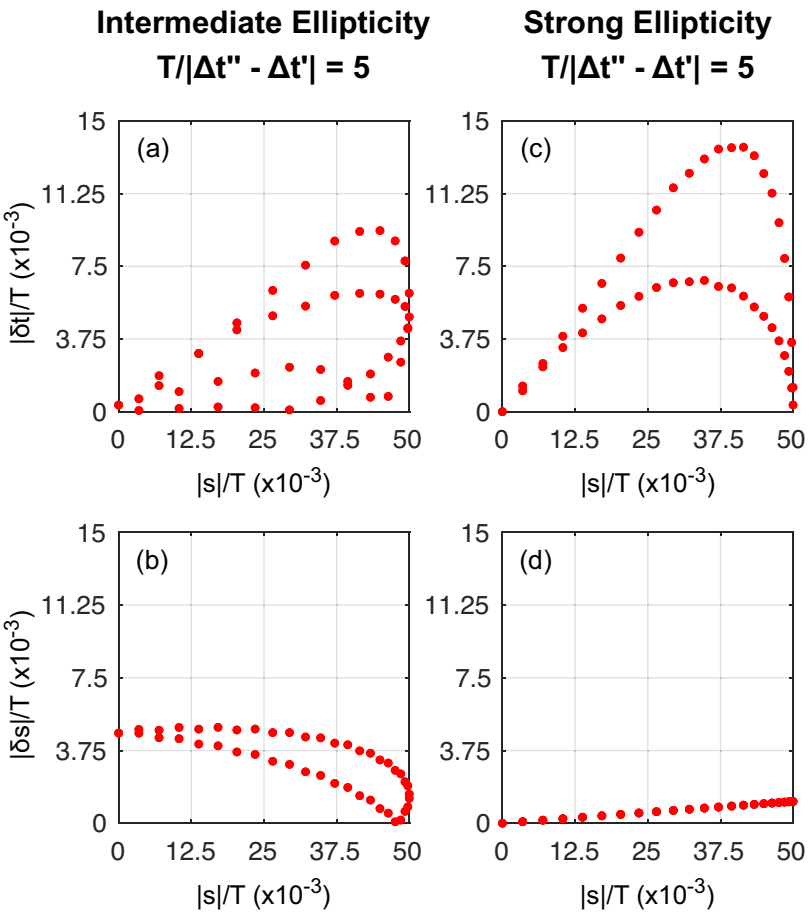

Figure B7. Absolute errors between the predicted and measured (top row) delay times and (bottom row) splitting intensities shown in Fig. B6. Errors in splitting intensity are computed after removing the mean offset between the measured and predicted curves (Figs B6c and d). Errors are plotted as a function of the predicted splitting intensity (eq. 7).