

# Bayesian inversion of surface-wave data for radial and azimuthal shear-wave anisotropy, with applications to central Mongolia and west-central Italy

Matteo Ravenna<sup>1,2</sup> and Sergei Lebedev<sup>1</sup>

<sup>1</sup>*School of Cosmic Physics, Geophysics Section, Dublin Institute for Advanced Studies, 5 Merrion Square, Dublin 2, Ireland. E-mail: [mravenna@cp.dias.ie](mailto:mravenna@cp.dias.ie)*

<sup>2</sup>*School of Earth Sciences, University College Dublin, Belfield, Dublin 4, Ireland*

Accepted 2017 November 16. Received 2017 November 9; in original form 2017 July 6

## SUMMARY

Seismic anisotropy provides important information on the deformation history of the Earth's interior. Rayleigh and Love surface-waves are sensitive to and can be used to determine both radial and azimuthal shear-wave anisotropies at depth, but parameter trade-offs give rise to substantial model non-uniqueness.

Here, we explore the trade-offs between isotropic and anisotropic structure parameters and present a suite of methods for the inversion of surface-wave, phase-velocity curves for radial and azimuthal anisotropies. One Markov chain Monte Carlo (MCMC) implementation inverts Rayleigh and Love dispersion curves for a radially anisotropic shear velocity profile of the crust and upper mantle. Another MCMC implementation inverts Rayleigh phase velocities and their azimuthal anisotropy for profiles of vertically polarized shear velocity and its depth-dependent azimuthal anisotropy. The azimuthal anisotropy inversion is fully non-linear, with the forward problem solved numerically at different azimuths for every model realization, which ensures that any linearization biases are avoided. The computations are performed in parallel, in order to reduce the computing time. The often challenging issue of data noise estimation is addressed by means of a Hierarchical Bayesian approach, with the variance of the noise treated as an unknown during the radial anisotropy inversion. In addition to the MCMC inversions, we also present faster, non-linear gradient-search inversions for the same anisotropic structure. The results of the two approaches are mutually consistent; the advantage of the MCMC inversions is that they provide a measure of uncertainty of the models.

Applying the method to broad-band data from the Baikal-central Mongolia region, we determine radial anisotropy from the crust down to the transition-zone depths. Robust negative anisotropy ( $V_{sh} < V_{sv}$ ) in the asthenosphere, at 100–300 km depths, presents strong new evidence for a vertical component of asthenospheric flow. This is consistent with an upward flow from below the thick lithosphere of the Siberian Craton to below the thinner lithosphere of central Mongolia, likely to give rise to decompression melting and the scattered, sporadic volcanism observed in the Baikal Rift area, as proposed previously. Inversion of phase-velocity data from west-central Italy for azimuthal anisotropy reveals a clear change in the shear-wave fast-propagation direction at 70–100 km depths, near the lithosphere–asthenosphere boundary. The orientation of the fabric in the lithosphere is roughly E–W, parallel to the direction of stretching over the last 10 m.y. The orientation of the fabric in the asthenosphere is NW–SE, matching the fast directions inferred from shear-wave splitting and probably indicating the direction of the asthenospheric flow.

**Key words:** Inverse theory; Probability distributions; Computational seismology; Seismic anisotropy; Seismic tomography; Surface waves and free oscillations.

## 1 INTRODUCTION

Seismic anisotropy—the directional and polarization dependence of seismic wave speeds—yields essential constraints on the structure

and dynamics of the Earth's crust and mantle. The current growth in the quantity of available seismic data offers unprecedented amounts of observational evidence on anisotropy distributions. However, the amplitude of anisotropy within the Earth is normally much smaller

than that of the variations in isotropic wave speeds. It is thus an important outstanding challenge to develop methods that can resolve the trade-offs between the isotropic and anisotropic heterogeneities and constrain robust models of anisotropy.

Seismic anisotropy observed in the Earth at different depths and at different spatial scales may have a number of origins. In the upper mantle, lattice or crystallographic preferential orientation of olivine (LPO and CPO) develops due to mantle deformation (e.g. Montagner & Anderson 1989), with the directions of fast seismic-wave propagation expected to align parallel to the mantle flow for large strain by progressive simple shear (Ribe 1989; Zhang & Karato 1995; Long & Becker 2010; Becker *et al.* 2014). In the crust, the crystallographic preferential orientation of crustal minerals also gives rise to anisotropy (e.g. Weiss *et al.* 1999; Endrun *et al.* 2008; Tatham *et al.* 2008; Xie *et al.* 2015).

In the upper crust, another cause of anisotropy is aligned cracks (with or without fluid inclusions), aligned parallel to the compressional component of the tectonic stress in the region (Nur 1971). This is one of the mechanisms referred to as shape-preferred orientation (SPO) mechanisms (Crampin 1984; Babuska & Cara 1991). They result in ‘extrinsic’ anisotropy (Wang *et al.* 2013), not related to the orientation of crystals, in contrast to the ‘intrinsic’ anisotropy, produced by the LPO (CPO) mechanisms. Furthermore, a purely isotropic layered medium with layer thickness much smaller than the seismic wavelength, for example, a sedimentary deposit with fine layers or material with layers of partial melt (Kawakatsu *et al.* 2009), can be seismically equivalent to a transversely isotropic medium (Backus 1962; Bodin *et al.* 2015).

The evolution of anisotropy in a rock is dependent on the temperature and pressure that it is subjected to. Preferred orientation of minerals can develop at high temperatures, when the rock is undergoing pervasive deformation, and then preserved intact at a lower temperature, when the rock is stronger and not deforming significantly. Therefore, regions deformed at high temperatures and then cooled have their anisotropy ‘frozen in’ from past tectonic episodes (Silver & Chan 1988; Deschamps *et al.* 2008b), with the anisotropy reflecting the past and not current deformation.

Seismic surface waves are particularly sensitive to the anisotropic structure of the crust and upper mantle. Regional- and global-scale tomography studies have used growing volumes of surface-wave data to constrain increasingly detailed anisotropic models (Trampert & Woodhouse 1996; Debayle & Kennet 2000; Simons *et al.* 2002; Gung *et al.* 2003; Debayle *et al.* 2005; Panning & Romanowicz 2006; Zhang *et al.* 2009; Yuan & Romanowicz 2010; Becker *et al.* 2012; Polat *et al.* 2012; Zhu & Tromp 2013; Chang *et al.* 2014, 2015; Fry *et al.* 2014; Yuan & Beghein 2014; Debayle *et al.* 2016; Schaeffer *et al.* 2016), which provided new insights into the dynamics of the upper mantle and crust and even the absolute motions of tectonic plates (Becker *et al.* 2015). Many of the methods used for the inversion of surface-wave data rely on waveform fitting, based on optimization techniques that minimize the difference between synthetic and observed waveforms (e.g. Woodhouse & Dziewonski 1984; Snieder 1988; Nolet 1990; Lebedev *et al.* 2005; Lebedev & Van der Hilst 2008). Most global and regional anisotropic models have been derived using asymptotic and ray-based approaches (Debayle *et al.* 2005; Debayle & Ricard 2012; Schaeffer & Lebedev 2013), which are computationally less expensive than numerical wavefield simulation approaches (Fichtner *et al.* 2010; Lekic & Romanowicz 2011; Zhu *et al.* 2012; French *et al.* 2013).

The inversions of Rayleigh- and Love-wave phase and group velocities for azimuthal and radial anisotropies (Montagner &

Tanimoto 1990, 1991; Trampert & Woodhouse 1996; Ekstrom *et al.* 1997; Ekstrom & Dziewonski 1998; Ritzwoller *et al.* 2002) is an approach alternative to waveform fitting. It has been used since the early observations of the discrepancy between Rayleigh and Love dispersion curves (Anderson 1961; Harkrider & Anderson 1962; McEvelly 1964; Anderson 1967).

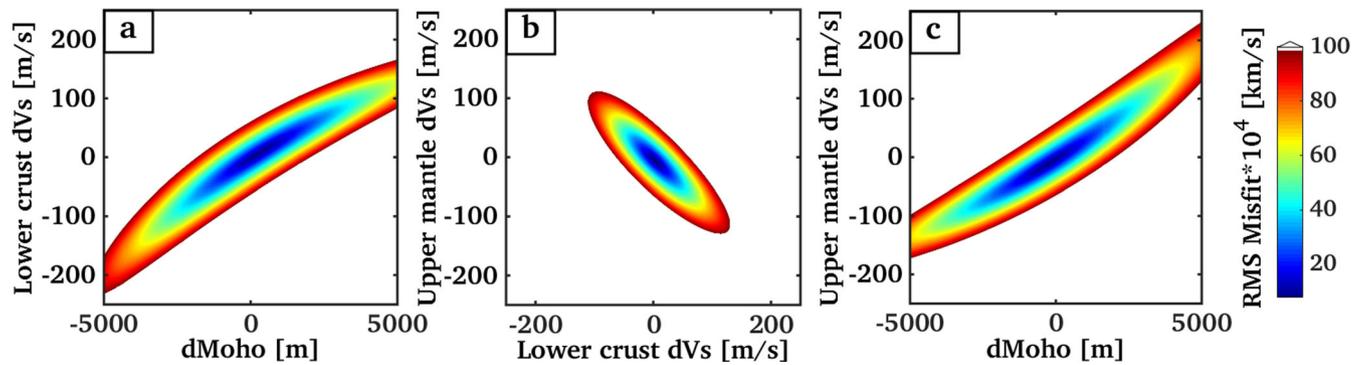
Inversions of surface-wave dispersion measurements provide constraints on the depth dependence of both radial anisotropy (Lebedev *et al.* 2006; Nettles & Dziewonski 2008; Visser *et al.* 2008; Ferreira *et al.* 2010; Huang *et al.* 2010; Agius & Lebedev, 2013, 2014; Auer *et al.* 2014; Chang *et al.* 2014, 2015) and azimuthal anisotropy (Yao *et al.* 2010; Roux *et al.* 2011; Darbyshire *et al.* 2013; Sodoudi *et al.* 2013; Yuan & Beghein 2013; Beghein *et al.* 2014; Burgos *et al.* 2014), at global and regional scales.

Most of the techniques used in inversions for anisotropy are based on linear approximations of the relation linking Earth’s model parameter to surface-wave phase velocities. Also, all inversion approaches face the problem of the severe non-uniqueness of solutions, which is especially challenging for the anisotropic, compared to isotropic, structure.

The Monte Carlo method provides a way to address and quantify the non-uniqueness, using exact, numerical solutions of the forward problem and direct sampling of the parameter space. The result of the inversion is not a single best-fitting model, but a probabilistic ensemble of models that fit the data (almost) equally well. This method has already been successfully applied to the inversion of dispersion curves, sometimes jointly with other data types, for anisotropic models of the Earth (Shapiro & Ritzwoller 2002; Khan *et al.* 2011; Drilleau *et al.* 2013; Xie *et al.* 2013, 2015; Bodin *et al.* 2016).

In this paper, we develop a Markov chain Monte Carlo (MCMC) inversion of Rayleigh- and Love-wave dispersion data for 1-D, radially and azimuthally anisotropic shear velocity profiles, using a fully non-linear model space search approach. One key innovation of this study is the fully non-linear approach for the inversion of azimuthally anisotropic dispersion curves: the computation of the synthetic azimuthally anisotropic phase velocities is performed in parallel and separately at different azimuths, without using sensitivity kernels for a model that is averaged azimuthally (Maupin 1985; Montagner & Nataf 1986). Furthermore, we also address the important issue of data noise estimation, which has a direct impact on the structure and non-uniqueness of the resulting Earth models, by using a Hierarchical Bayesian approach in the inversions for radial anisotropy. This allows the variance of data noise to be treated as an unknown during the inversion, so that it is possible to let the data infer its degree of uncertainty. This approach has previously been applied in the inversion of surface-wave data jointly with other data types (Bodin *et al.* 2012). A new element presented here is the parametrization of the variance of data noise as a function of period.

In the following, we first explore the model spaces associated with the anisotropy inversions, in order to analyse the correlations between anisotropic and isotropic model parameters (Section 2). The analysis of these correlations provides guidance for the implementation of the MCMC algorithms, as well as a better understanding of the inversion results. We then present a suite of independent inversion programs tackling the phase-velocity inversion problem with two different approaches: in Section 3 we introduce a fast, non-linear gradient-search inversion algorithm for obtaining single, best-fitting, anisotropic shear velocity profiles; in Section 4 we detail the implementation of our MCMC inversion algorithms in the Bayesian Framework. Next, we demonstrate the resolving power of the MCMC algorithms when applied to the inversion of synthetic



**Figure 1.** The rms misfit surfaces for the isotropic-average  $V_s$  (at different depths) and the Moho depth. Perturbations have been applied to an isotropic continental model with a Moho depth of 46 km. The rms misfit between the reference and perturbed phase-velocity curves is computed over their entire length (9–333 s for Rayleigh and 8–341 s for Love). (a) The misfit as a function of variations of the Moho depth and the  $S$ -wave velocity in the lower crust (between 27 km and the Moho). (b) The misfit as a function of variations of the  $S$ -wave velocity in the lower crust and in the uppermost mantle (between the Moho and a 76 km depth). (c) The misfit as a function of variations of the Moho depth and  $S$ -wave velocity in the uppermost mantle.

and real fundamental-mode phase-velocity data sets. The data are from two regions associated with important open questions on the dynamics of the lithosphere and intraplate volcanism, the Baikal-Mongolia region (Section 5) and Tuscany (Section 6). We briefly discuss the implications of our results for the depth and evolution of the lithosphere–asthenosphere boundary (LAB) and the lithospheric dynamic processes.

## 2 THE MODEL SPACE

The non-uniqueness of the shear-wave velocity profiles constrained by surface-wave inversion is characterized by trade-offs between  $V_s$  in neighbouring depth ranges as well as between depths of discontinuities (e.g. the Moho) and  $V_s$  in the adjacent depth ranges. The trade-offs in isotropic-average  $V_s$  have been investigated in a number of studies (e.g. Bodin *et al.* 2012; Lebedev *et al.* 2013), but the trade-offs between anisotropic and isotropic parameters are less well understood, even though their existence is well known (Levshin & Ratnikova 1984).

In this section, we explore the trade-offs between the following model parameters: the Moho depth and the isotropic  $V_s$  in the lower crust and in the upper mantle (Fig. 1; these trade-offs are well known but are important to include for comparison with the anisotropy trade-offs); radial anisotropy ( $V_{sh}-V_{sv}$ ) and isotropic  $V_s$  in the lower crust and in the upper mantle and the Moho depth (Fig. 2); and the fast-propagation azimuths of the  $2\varphi$  components of azimuthal anisotropy in the lower crust and in the upper mantle (Fig. 3). To explore the model space, we performed several 2-D grid searches. In each run of the grid search, we perturbed two parameters of a reference model over ranges of values, leaving unchanged the other parameters, and we evaluated the root-mean-square (rms) misfit between the synthetic phase velocities computed for the reference model and for the perturbed models.

In an isotropic medium, surface-wave phase velocities are mainly sensitive to the (isotropic) shear speed  $V_s$ , varying as a function of depth for a 1-D Earth model. (Phase-velocity sensitivity to compressional speed  $V_p$  is also not negligible but is smaller).

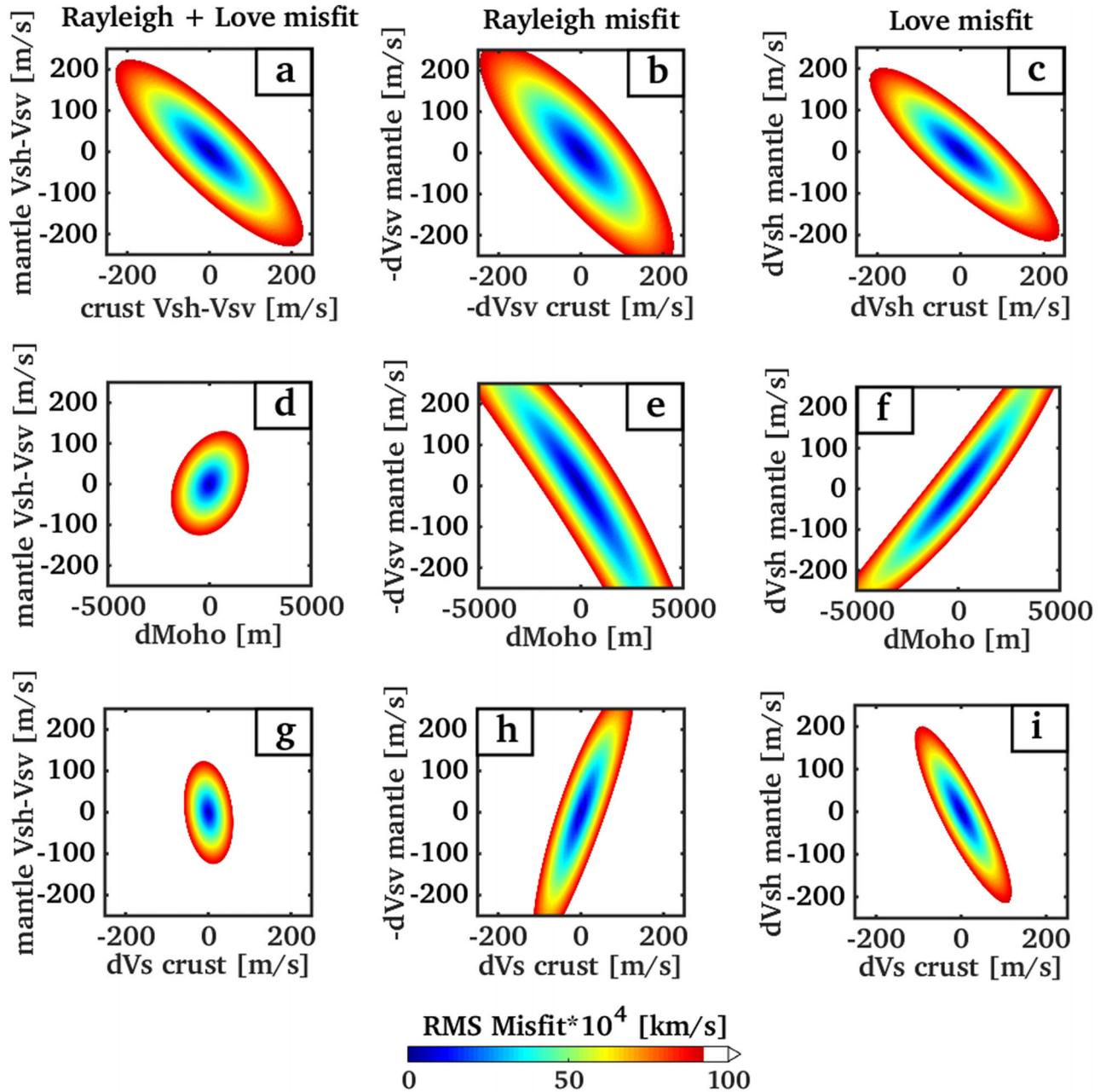
To compute the misfit functions shown in Fig. 1, we generated synthetic phase velocities (Section 3.1) for a reference, isotropic model and then applied perturbations to the Moho depth, the isotropic shear speeds in the lower crust and the isotropic shear speeds in the upper mantle just below the Moho, while leaving all

the other parameters unchanged. The reference model is a modified AK135 (Kennett *et al.* 1995) and has three layers in the crust and a Moho depth of 46 km; perturbations are controlled by boxcar-shaped basis functions in the lower crust (from 27 to 46 km) and in the uppermost mantle (from 46 to 76 km, Bartzsch *et al.* 2011; Agius & Lebedev 2013). The misfit is defined as the rms difference between the reference and perturbed Rayleigh and Love phase velocities, integrated over their entire frequency range. As seen in previous studies (e.g. Bodin *et al.* 2012; Lebedev *et al.* 2013), the valleys of low misfit clearly indicate the presence of trade-offs between the Moho depth and the shear speeds: a deep Moho correlates with high lower crustal and upper-mantle velocities, and a shallow Moho correlates with low lower crustal and upper-mantle velocities. We also noted an anti-correlation between lower crustal and upper-mantle velocities.

In a transversely isotropic (radially anisotropic) medium, Rayleigh and Love surface-wave phase velocities are mainly sensitive to the vertically and horizontally polarized shear speeds  $V_{sv}$  and  $V_{sh}$ , respectively. It has been reported that  $V_s$  radial anisotropy mapped in the mantle in global tomographic inversions of long-period surface-wave data can be affected by the assumed isotropic shear velocity structure of the crust (Ferreira *et al.* 2010). Chang & Ferreira (2017) reported that including shorter period group-velocity data with periods down to 20 s or below helped to avoid the leakage of crustal structure into the mantle radial anisotropy structure. Here, we explore the parameter space in order to see the basic trade-offs between isotropic and radially anisotropic shear velocity parameters of the lower crust and the upper mantle.

We generated synthetic phase velocities for a reference, isotropic model and applied perturbations to the Moho depth, radial anisotropy ( $V_{sh}-V_{sv}$ ) of the lower crust, radial anisotropy ( $V_{sh}-V_{sv}$ ) of the upper mantle just below the Moho and the isotropic  $V_s$  of the lower crust, while leaving all other parameters unchanged. The reference model is a modified AK135, with three layers in the crust and a Moho depth of 46 km; perturbations are controlled by boxcar-shaped basis functions both in the lower crust (27–46 km) and in the uppermost mantle (46–76 km). The misfits (Fig. 2) have been defined, as in the previous grid search, as the rms misfits between the reference and perturbed Rayleigh and Love phase-velocity curves, integrated over their lengths.

The low-misfit valleys indicate an anti-correlation between radial anisotropy in the lower crust and in the uppermost mantle (Fig. 2a). In contrast, the trade-offs between the Moho depth and

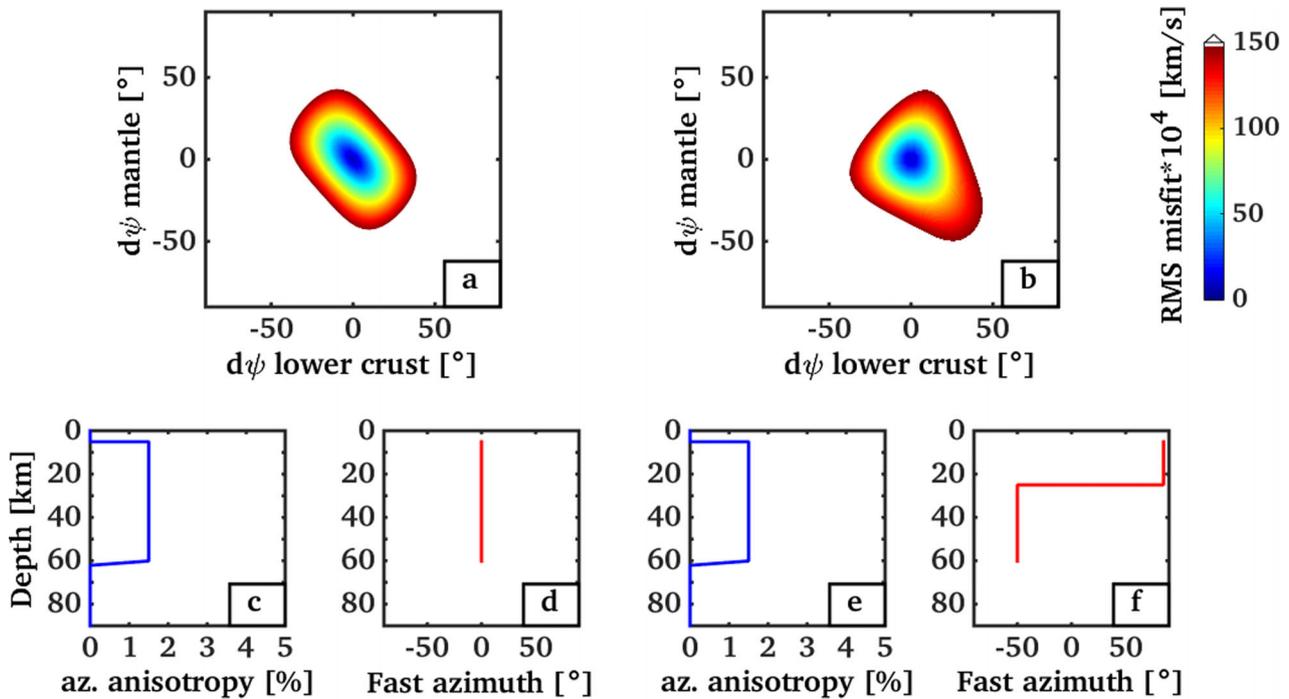


**Figure 2.** The rms misfit surfaces for radial anisotropy ( $V_{sh}-V_{sv}$ ),  $V_s$ ,  $V_{sv}$ ,  $V_{sh}$  (at different depths), and the Moho depth. Model perturbations have been applied to an isotropic continental model with a Moho depth of 46 km. The rms misfit between the reference and perturbed phase-velocity curves is computed over their entire length (9–333 s for Rayleigh and 8–341 s for Love). (a) The misfit dependence on radial anisotropy in the lower crust (between 27 km and the Moho) and in the uppermost mantle (between the Moho and the 76 km depth). (d) The misfit dependence on the Moho depth and radial anisotropy in the uppermost mantle. (g) The misfit dependence on isotropic  $V_s$  in the lower crust and radial anisotropy in the uppermost mantle. (b), (e) and (h) Same as (a), (d) and (g), but displaying only the effect of  $V_{sv}$  perturbations on the Rayleigh phase-velocity misfit. (c), (f) and (i) Same as (a), (d) and (g), but displaying only the effect of  $V_{sh}$  perturbations on the Love phase-velocity misfit.

radial anisotropy in the uppermost mantle (Fig. 2d) and between isotropic average  $V_s$  in the lower crust and radial anisotropy in the uppermost mantle (Fig. 2g) are small. This is because the trade-offs of  $V_{sv}$  and  $V_{sh}$  with the Moho depth, which we map, for reference, by computing misfits over the length of the Rayleigh (Figs 2b, e and h) and Love (Figs 2, c, f and i) phase-velocity curves, respectively, cancel each other out to a large extent, as the comparison of the different panels shows.

This suggests that phase-velocity data including periods down to 7–8 s, as in our tests and inversions, resolves the potential trade-offs between crustal structure and the radial anisotropy in the mantle.

The variations of the  $2\varphi$  azimuthal anisotropy with depth have important implications for the dynamics and evolution of the lithosphere; they can be determined from Rayleigh-wave azimuthal anisotropy (e.g. Deschamps *et al.* 2008a; Yao *et al.* 2010; Yuan & Romanowicz 2010; Agius & Lebedev 2017).



**Figure 3.** The rms misfit surfaces between the fast-propagation azimuths of azimuthally anisotropic shear-waves at different depths. (c)–(f) Perturbations have been applied to two different azimuthally anisotropic, continental models, characterized by different layering of azimuthal anisotropy. The rms misfit between the reference and perturbed Rayleigh-wave phase-velocity curves is computed over their entire length (5–95 s) over nine different, equally spaced azimuths. The azimuthal anisotropy in the model is purely  $2\phi$ . (a) The misfit dependence on the fast-propagation azimuth in the lower crust (between 14 km and the Moho) and in the uppermost mantle (between the Moho and 60 km), for the reference model with azimuthally anisotropic parameters shown in (c) and (d). (b) The misfit for the fast-propagation azimuth in the lower crust (between 14 km and the Moho) and in the uppermost mantle (between the Moho and 60 km), for the reference model shown in (e) and (f).

It is important to evaluate the trade-offs of the azimuthal anisotropy parameters at different depths within a model. We now map the trade-offs between the fast-propagation azimuth in the lower crust and in the upper mantle, in order to explore how much they may affect the azimuthal anisotropy models, including the profiles sampled in our MCMC azimuthally anisotropic inversions below.

To compute the misfit functions in each of the two tests shown in Fig. 3, we generated azimuth-specific synthetic Rayleigh-wave phase-velocity curves at nine equally spaced azimuths (between  $-90^\circ$  and  $70^\circ$ , at a  $20^\circ$  interval). The reference models, characterized by different amplitude and fast-propagation azimuth of azimuthal anisotropy, are shown in Figs 3(c)–(f). We applied perturbations to the fast-propagation azimuths in the lower crust and in the uppermost mantle, while leaving the rest of the models unchanged. Both reference models have a Moho depth of 25 km. The perturbations are controlled by boxcar-shaped basis functions, spanning 14–25 km depths in the lower crust and 25–60 km depths in the upper mantle. The rms misfit between the reference and perturbed Rayleigh phase velocities is integrated over the length of the dispersion curves (5–95 s) and over the nine curves at the different azimuths.

The difference of the low-misfit valleys in the two examples in Fig. 3 shows that the trade-off between fast-propagation azimuths in adjacent depth ranges depends on the relative values of the fast-propagation azimuths. In particular, we observe a more pronounced anti-correlation between the fast-propagation azimuth in the lower crust and in the uppermost mantle when the fast azimuth of the reference model is constant in the entire depth range, compared to

when it changes at the Moho, the boundary of the two anisotropic layers in these test models.

### 3 GRADIENT-SEARCH INVERSION

This section describes the inversion of phase-velocity curves for anisotropic structure using a non-linear gradient-search algorithm. This is a fast and robust inversion technique able to retrieve an accurate best-fitting Earth model from phase-velocity curves. Comparison of the models with those obtained by the MCMC inversions, described in the following sections, helps us to cross-validate the different inversion approaches. The gradient-search inversions can also be used for fast estimates of frequency-dependent errors in the measurements (Section 3.2). We have also used this approach while assessing how much it is possible to decimate (resample at a lower sampling rate) the dispersion curves without losing resolution.

#### 3.1 Physical parameters and forward problem

For a vertically transverse isotropic (VTI) medium (i.e. medium with hexagonal symmetry around a vertical axis), the fourth-order elastic tensor is specified by five moduli (Love 1927), related to compressional ( $V_{pv}$ ,  $V_{ph}$ ) and shear ( $V_{sv}$ ,  $V_{sh}$ ) wave speeds.

Rayleigh-wave phase velocities are mainly sensitive to vertically polarized shear-wave velocities ( $V_{sv}$ ), and are also sensitive, to a lesser extent, to compressional velocities. Love-wave phase velocities are insensitive to compressional velocities

and are mainly sensitive to horizontally polarized shear-wave velocities ( $V_{sh}$ ).

In our inversion schemes, when we invert jointly the Rayleigh and Love fundamental-mode dispersion curves, we choose as parameters the isotropic-average shear speed  $V_s = \frac{V_{sv} + V_{sh}}{2}$  and radial anisotropy  $R_{ani} = \frac{V_{sh} - V_{sv}}{V_s}$  as functions of depth.

The computation of the synthetic phase-velocity curves of the fundamental-mode Rayleigh and Love waves is performed using the MINEOS modes code (Masters *et al.* 2007, <http://geodynamics.org/cig/software/mineos>), adapted for the traveling wave decomposition (Nolet 1990, 2008) and streamlined for fast applications within inversions (e.g. Agius & Lebedev 2013), in the approximation of a self-gravitating, spherically symmetric, transversely isotropic, attenuating Earth. To speed up the calculations in our inversions, we incorporate MINEOS as a function into the inversion programs. Important features of MINEOS are its accuracy, robustness and flexibility regarding the input 1-D Earth models, with no *a priori* restrictions on their parametrization. The models are parametrized using radial knots with values of the depth, density, vertically and horizontally polarized compressional and shear velocities,  $Q$  factors and the parameter  $\eta$  that controls the variation of seismic velocities away from the symmetry axis in the presence of anisotropy.

Although surface-wave phase velocities are sensitive to compressional wave velocities, density ( $\rho$ ) and  $\eta$  (e.g. Beghein & Trampert 2004; Beghein *et al.* 2006), these parameters are difficult to resolve in an inversion as independent variables due to their strong trade-offs with  $V_{sv}$  and  $V_{sh}$ , which the phase velocities are most sensitive to. We use scaling relations to calculate the variations in  $V_{pv}$  and  $V_{ph}$ . To account for radial anisotropy in  $V_p$ , the compressional speeds are recomputed from the shear speeds assuming an isotropic bulk modulus ( $K_v = K_h = K$ ), so that

$$K = \rho(V_p^2 - (4/3)V_s^2), \quad (1)$$

and a VTI shear modulus ( $\mu_v \neq \mu_h$ ). In this case, using the isotropic relations between shear speeds and elastic parameters along the vertical and horizontal directions, we have:

$$V_{pv}^2 = \frac{K}{\rho} + (4/3)V_{sv}^2; \quad (2)$$

$$V_{ph}^2 = \frac{K}{\rho} + (4/3)V_{sh}^2. \quad (3)$$

Therefore, substituting eq. (1) in eqs (2) and (3), we have:

$$V_{pv} = \sqrt{V_p^2 + 4/3(V_{sv}^2 - V_s^2)}; \quad (4)$$

$$V_{ph} = \sqrt{V_p^2 + 4/3(V_{sh}^2 - V_s^2)}, \quad (5)$$

where  $V_p$  (the isotropic-average compressional speed) is recomputed from  $V_s$  using reference  $\frac{V_p}{V_s}$  ratios, taken from CRUST 2.0 (Bassin *et al.* 2000) in the crust and AK135 in the mantle. Density and the compressional and shear  $Q$  factors are fixed at the reference values (CRUST 2.0 in the crust and AK135 in the mantle), and the parameter  $\eta$  is set to 1.

In the case of small azimuthal anisotropy of the medium, the azimuthally anisotropic Rayleigh and Love phase velocities can be expressed using perturbations to the isotropic-average phase velocity  $C_0(T)$  (Smith & Dahlen 1973):

$$C(T, \varphi) = C_0(T) + C_1(T) \cos(2\varphi) + C_2(T) \sin(2\varphi) \\ + C_3(T) \cos(4\varphi) + C_4(T) \sin(4\varphi), \quad (6)$$

where  $\varphi$  is the azimuth of the propagation direction (positive from north to east) and  $C_1(T)$ ,  $C_2(T)$ ,  $C_3(T)$ ,  $C_4(T)$  are the azimuthal anisotropy coefficients. Montagner & Nataf (1986) showed that for Rayleigh-wave phase velocities at periods sampling the upper mantle, the  $2\varphi$  terms are dominant, with the  $4\varphi$  ones small.

In order to invert azimuthally anisotropic phase velocities for seismic-velocity anisotropy as a function of depth, the relationship between phase velocities and the elastic parameters needs to be established. One common approach is the integration over depth of the linearized relations between the two (Maupin 1985; Montagner & Nataf 1986), with  $C_1(T)$  and  $C_2(T)$  linearized around a model that is averaged azimuthally. The sensitivity kernels (partial derivatives of phase velocities with respect to the elastic parameters) can be computed using the theory for a transversely isotropic model (Takeuchi & Saito 1972).

In our inversion schemes for azimuthal anisotropy, we treat the problem fully non-linearly, which is more numerically expensive but avoids potential biases due to linearization. We parametrize our model in terms of  $V_{sv}^0(d)$  (isotropic average, vertically polarized shear velocities as a function of depth  $d$ ) and  $A_1(d)$  and  $A_2(d)$  (azimuthal anisotropy coefficients) and define the velocity of a vertically polarized, horizontally propagating  $SV$  wave at each depth (Crampin 1984; Montagner & Nataf 1986; Lévêque *et al.* 1998) as

$$V_{sv}(d, \varphi) = V_{sv}^0(d) + A_1(d) \cos(2\varphi) + A_2(d) \sin(2\varphi). \quad (7)$$

We compute the fundamental-mode Rayleigh-wave phase-velocity curves for a given azimuth using MINEOS, with the azimuth-specific 1-D shear velocity profile determined through eq. (7).

The compressional speeds are recomputed, as previously, assuming an isotropic bulk modulus and azimuthally anisotropic shear modulus:

$$V_{pv}(\varphi) = \sqrt{(V_{pv}^0)^2 + 4/3(V_{sv}(\varphi))^2 - (V_{sv}^0)^2}. \quad (8)$$

Once the parameters  $A_1(d)$  and  $A_2(d)$  are determined by an inversion, we compute the depth-dependent fast-propagation azimuth  $\Psi$  and the amplitude of anisotropy  $A$  as

$$\Psi(d) = 0.5^* \arctan(A_2(d)/A_1(d)); \quad (9)$$

$$A(d) = \sqrt{A_1(d)^2 + A_2(d)^2}. \quad (10)$$

### 3.2 Inversion for radial anisotropy and isotropic-average $V_s$

The gradient-search algorithm we use for the inversion of Rayleigh- and Love-wave dispersion curves for radial anisotropy has been developed and extensively tested previously (Meier *et al.* 2004; Lebedev *et al.* 2006; Endrun *et al.* 2008; Erduran *et al.* 2008; Agius & Lebedev 2014). Here, we updated the technique to make all phase-velocity calculations with MINEOS fully radially anisotropic and to incorporate MINEOS into the code as a function, in order to speed up the calculations. We invert Rayleigh- and Love-wave dispersion curves for 1-D profiles of the isotropic-average shear speed  $V_s$  as a function of depth and radial anisotropy  $R_{ani}$  as a function of depth. The Levenberg–Marquardt algorithm we use is a Gauss–Newton algorithm that iteratively perturbs the model parameters so as to minimize a misfit function in a non-linear least-squares approach; it is more robust than a conventional Gauss–Newton algorithm as it converges to the true best-fitting solution even if the initial guess is very far from it. The misfit function minimized

by the optimization procedure quantifies the difference between the observed and synthetic phase velocities:

$$F \propto \sum_{i=1}^N \left( \frac{d_i^R - C(\mathbf{V}_S, \mathbf{R}_{\text{ani}})_i^R}{\sigma_i^R} \right)^2 + \sum_{i=1}^N \left( \frac{d_i^L - C(\mathbf{V}_S, \mathbf{R}_{\text{ani}})_i^L}{\sigma_i^L} \right)^2, \quad (11)$$

where  $d_i^R$  and  $d_i^L$  are the observed Rayleigh and Love phase velocities at the  $i$ th period  $T_i$ ,  $C(\mathbf{V}_S, \mathbf{R}_{\text{ani}})_i^R$ ,  $C(\mathbf{V}_S, \mathbf{R}_{\text{ani}})_i^L$  are the synthetic Rayleigh and Love phase velocities computed for a model with isotropic-average shear speed  $\mathbf{V}_S$  and anisotropy  $\mathbf{R}_{\text{ani}}$  at the period  $T_i$ , and  $\sigma_i^R$ , and  $\sigma_i^L$  are the estimated errors of the Rayleigh and Love measured phase velocities.

Damping coefficients for each of the model parameters are also added to the misfit function, to control the smoothness of the resulting model. This is necessary in order to prevent the inversion from overfitting (noisy) data with unrealistic oscillatory models.

Perturbations in the model parameters are controlled by basis functions, boxcar-shaped in the crust and triangular in the mantle (e.g. Bartsch *et al.* 2011; Agius & Lebedev 2013). The depth of the discontinuities in the crust (including the Moho depth) is also parameters of the inversion and are allowed to vary. The radially anisotropic  $P$ -wave velocity is recomputed during the inversion from the radially anisotropic  $S$ -wave velocity (Section 3.1).

We used the gradient-search inversion algorithm to estimate errors of the observed Rayleigh and Love phase-velocity curves; these estimates were then used in the computation of the likelihood function defined in the MCMC algorithm (Section 4.4). The errors were estimated as the difference between the phase velocities computed for the best-fitting, weakly regularized model and the phase velocities observed, and then evaluating the envelope of the resulting misfit function. This estimation is effective for random noise or for errors with substantial variability with frequency. Such variability stands out as noise because it cannot be reproduced by any synthetic phase-velocity curve (phase-velocity curves are always smooth, because phase velocities depend on integrals of shear and compressional velocities in broad depth ranges within the Earth and because these depth ranges change with frequency very gradually). This error estimation method is less effective for measurement biases that are nearly constant over broad period ranges: such biases in the phase-velocity curves could, possibly, be reproduced, at least to some extent, by (artificial) anomalies in Earth models. We thus consider these error estimates to be lower estimates.

We also used the gradient-search inversion to determine to what extent a phase-velocity curve can be decimated (resampled with fewer samples) without a decrease in the accuracy of the inversion. The decimation is desirable in order to reduce the substantial computational cost of the MCMC inversions. In the presence of noise, however, choosing a small subset of samples can bias the models produced by the inversion of the data. We tested this by means of a series of gradient-search inversions of the Rayleigh and Love phase-velocity curves measured in the Baikal-Mongolia region (the data inverted in Section 5). The period ranges of the curves are 8–341 and 9–333 s and the number of samples in the full sample sets are 240 and 236 for the Love- and Rayleigh-wave curves, respectively. The sample spacing increases with period logarithmically, for the structural sensitivity along the length of the broad-band curves to be balanced, approximately. For each inversion in our test series, these original curves were resampled at a lower rate, and the resampled curve was inverted. Then, an rms misfit was computed as the difference between the best-fitting phase velocities resulting from the inversion of the original curves and the best-fitting phase ve-

locities (interpolated where necessary) resulting from the inversion of the resampled curves. As can be seen in Fig. 4, for a resampling factor equal to 0.17 (i.e. 17 per cent of samples are kept), the rms misfit is small but clearly non-zero; 1.6 and 2.17 m s<sup>-1</sup> for the Love- and Rayleigh-wave curves, respectively. Guided by these tests, we choose a resampling factor of 0.5 and use it in our MCMC inversions.

### 3.3 Inversion for azimuthal anisotropy

We also implemented a gradient-search inversion of the azimuthal anisotropy of surface-wave phase velocities for 1-D profiles of depth-dependent azimuthal anisotropy of shear-wave velocities.

The parameters of the inversion span the 1-D profile of the isotropic-average, vertically polarized shear velocity  $\mathbf{V}_{\text{sv}}^0$  in the crust and upper mantle and the  $2\varphi$  components of its azimuthal anisotropy,  $\mathbf{A}_1$  and  $\mathbf{A}_2$ . The data are the isotropic-average phase velocities of Rayleigh waves and their  $2\varphi$  azimuthal anisotropy as a function of period, defined by  $C_1(T)$  and  $C_2(T)$  (eq. 6, Adam & Lebedev 2012). (The same approach can be applied for the inversion of anisotropic Love-wave phase velocities for horizontally polarized shear velocity  $\mathbf{V}_{\text{sh}}^0$  and the  $4\varphi$  components of azimuthal anisotropy.)

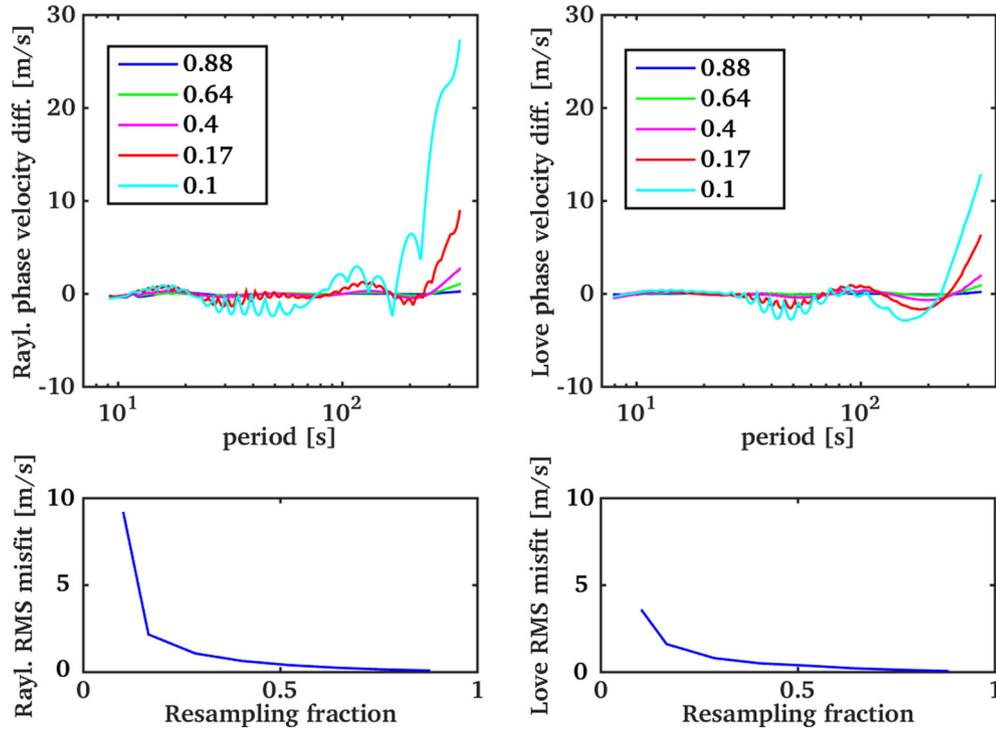
The data-synthetic misfit is computed over a set of nine azimuth-specific, Rayleigh-wave, phase-velocity curves at nine equally spaced azimuths, between  $-90^\circ$  and  $70^\circ$  at a  $20^\circ$  interval (Adam 2013; Adam & Lebedev 2012). The ‘observed’ phase-velocity curve at each azimuth is computed using  $C_0(T)$ ,  $C_1(T)$ ,  $C_2(T)$  (eq. 6). For every model tested in the course of the inversion, an azimuth-specific Earth model is computed for each azimuth using  $\mathbf{V}_{\text{sv}}^0$ ,  $\mathbf{A}_1$ ,  $\mathbf{A}_2$  (with  $V_p$  values modified accordingly, as described in Section 3.1) and phase-velocity curves are computed for each azimuth using transversely isotropic MINEOS.

The inversion uses the Levenberg–Marquardt, non-linear, least-squares approach. The misfit function minimized by the optimization procedure is proportional to the sum of the misfits computed at each azimuth:

$$F \propto \sum_{i=1}^N \left( \frac{(d_i^R)_{\varphi_1} - (C(\mathbf{V}_{\text{sv}}^0, \mathbf{A}_1, \mathbf{A}_2)_i^R)_{\varphi_1}}{(\sigma_i^R)_{\varphi_1}} \right)^2 + \sum_{i=1}^N \left( \frac{(d_i^R)_{\varphi_2} - (C(\mathbf{V}_{\text{sv}}^0, \mathbf{A}_1, \mathbf{A}_2)_i^R)_{\varphi_2}}{(\sigma_i^R)_{\varphi_2}} \right)^2 + \dots + \sum_{i=1}^N \left( \frac{(d_i^R)_{\varphi_n} - (C(\mathbf{V}_{\text{sv}}^0, \mathbf{A}_1, \mathbf{A}_2)_i^R)_{\varphi_n}}{(\sigma_i^R)_{\varphi_n}} \right)^2, \quad (12)$$

where  $(d_i^R)_{\varphi_j}$  and  $(C(\mathbf{V}_{\text{sv}}^0, \mathbf{A}_1, \mathbf{A}_2)_i^R)_{\varphi_j}$  are the Rayleigh dispersion curves observed and computed with the MINEOS code for a given model at period  $T_i$  and azimuth  $\varphi_j$  ( $j = 1, \dots, n$ ), respectively.  $(\sigma_i^R)_{\varphi_j}$  are the estimated errors on Rayleigh measured phase velocities at period  $T_i$  and azimuth  $\varphi_j$  ( $j = 1, \dots, n$ ).

As in the radial anisotropy inversion described in the previous section, perturbations in the model parameters ( $\mathbf{V}_{\text{sv}}^0$ ,  $\mathbf{A}_1$ ,  $\mathbf{A}_2$ ) are controlled by basis functions, boxcar-shaped in the crust and triangular in the mantle, with additional parameters for the depth of the discontinuities in the crust, including the Moho. Damping coefficients for each of the model parameters are also added to the misfit function to control the smoothness of the resulting model. It has been verified previously (Adam 2013) that there are no



**Figure 4.** Tests on the effect of the decimation (downsampling) of densely sampled phase-velocity curves on the results of their inversion. Top: the difference between the Rayleigh (left) and Love (right) phase velocities computed for the best-fitting models from gradient-search inversions of the original phase-velocity curves and the ones decimated by a factor specified (e.g. 0.1 means that 10 per cent of the samples were kept, and 0.88 means that 88 per cent of the samples were kept). The full number of samples for these data, measured in central Mongolia, is 236 for Rayleigh and 240 for Love-wave dispersion curves. Bottom: Rayleigh (left) and Love (right) rms misfit for the net curve differences shown above, plotted as a function of the resampling factor. Even though errors in the data are small, they give rise to increasingly significant differences as the number of samples decreases.

significant trade-offs between isotropic ( $\mathbf{V}_{sv}^0$ ) and anisotropic ( $A_1, A_2$ ) parameters.

The algorithm can be used for fast inversion of anisotropic phase velocities for depth-dependent anisotropy. Here, we also used it for cross-validation with the MCMC azimuthally anisotropic inversions (Section 6.2).

## 4 MCMC INVERSION

In this section, we present our implementation of the MCMC algorithm for the inversion of surface-wave anisotropy. The MCMC approach is effective with non-linear inverse problems because it uses a direct sampling of the parameter space, in a Bayesian statistical framework. Furthermore, it provides a way to quantify model non-uniqueness, as not just one best-fit model, but a whole ensemble of models that fit the data is determined. Inaccurate *a priori* assessment of data errors can affect the inversion results. We thus use, where possible, a Hierarchical Bayesian approach, in which the variance of data noise itself is treated as an unknown in the inversion.

### 4.1 Bayesian inference

Most of the Monte Carlo sampling techniques are based on a probabilistic formulation of the inverse problem. The aim of the sampling process is to perform an iterative random search in the solution space. To this end, at each iteration of the sampling procedure, a proposal probability density function is used to perturb the model

parameters, in order to generate a new proposed sample. The sampling algorithm can be designed to ensure that the generated sequence of models converges to a target distribution of the model space,  $\pi(\mathbf{x})$ , which approximates the solution space (Gilks *et al.* 1996; Denison *et al.* 2002). If Bayesian Inference is used, as in our specific case, the target distribution is seen as a posterior probability distribution of the model parameters  $\mathbf{x}$  conditioned on the observed data  $\mathbf{d}$  through a likelihood function,  $p(\mathbf{d}|\mathbf{x})$ , and the prior information on the model parameters,  $p(\mathbf{x})$ , via Bayes' rule (Bayes 1763):

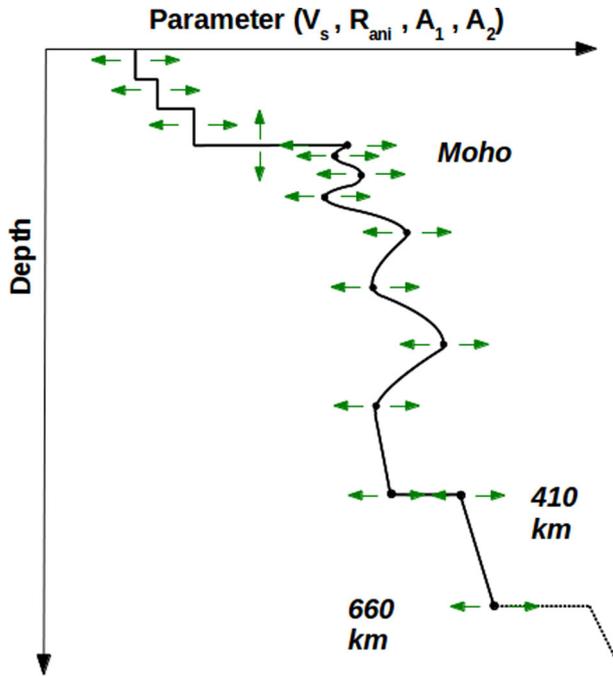
$$p(\mathbf{x}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{x})p(\mathbf{x}). \quad (13)$$

The likelihood function quantifies the probability of obtaining the observed data  $\mathbf{d}$  given the model  $\mathbf{x}$ , by measuring the match between the observed data  $\mathbf{d}$  and the data predicted by the model. Its general expression is:

$$p(\mathbf{d}|\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |C_e|}} \exp \frac{-(g(\mathbf{x}) - \mathbf{d})^T C_e^{-1} (g(\mathbf{x}) - \mathbf{d})}{2}, \quad (14)$$

where  $g(\mathbf{x})$  are the data estimated for the model  $\mathbf{x}$ ,  $C_e^{-1}$  is the inverse of the data noise covariance matrix,  $|C_e|$  is its determinant and  $n$  is the size of the data vector.

The prior probability distribution captures qualitative and quantitative prior knowledge regarding the model parameters; it provides one of the ways to deal with non-uniqueness by setting constraints (boundaries, hard or soft) on the model space sampling.



**Figure 5.** Schematic representation of the model parametrization. The crust is parametrized with stepwise layers (boxcar basis functions), and the mantle is parametrized with piecewise cubic Hermite spline polynomials based on control points that are chosen to be more dense in the uppermost mantle depth range. In the inversions of broad-band phase-velocity curves, the parameters at 410 and 660 km are also allowed to vary and the interpolation at depths between them is linear. The Moho depth is a parameter of the inversion. This parametrization is used for any kind of parameter, both in the radial anisotropy inversions and in the azimuthal anisotropy inversions.

#### 4.2 Data uncertainty quantification: Hierarchical Bayes

An extended formulation of the Bayesian theory is the Hierarchical Bayes approach (Gelman *et al.* 2004). This approach takes into account the possibility that questionable and subjective choices were made on the settings of the likelihood function and prior distribution. This is achieved by defining an ensemble of ‘hyperparameters’ that are treated as unknowns in the inversion. In this way, it is possible to let the data infer the values of parameters that are poorly known *a priori*, relying on uninformative *a priori* assumptions (hyperprior distributions, Malinverno & Briggs 2004).

In this study, in the inversions for radial anisotropy, we treat the variance of data noise as an unknown parameter; by doing so it is possible to let the data infer its degree of uncertainty (Bodin *et al.* 2012).

#### 4.3 Parametrization

Considering that surface-wave dispersion curves show high sensitivity to shear-wave velocities in broad depth intervals but low sensitivity to sharp discontinuities, we chose different parametrizations for the 1-D structure of the Earth’s crust and mantle, taking into account the strong variations in elastic parameters in the crust and the normally smoother variations in the mantle.

We parametrize the crust with three layers with constant elastic properties within each and with stepwise changes at the discontinuities between the layers (Fig. 5). The number and location of the intra-crustal discontinuities is fixed at 2 (upper crust-middle crust and middle crust-lower crust), whereas the Moho depth is allowed to vary during the inversion. The mantle is parametrized using

piecewise cubic Hermite spline polynomials, based on a number of control points. We also impose the presence of the 410 and 660 km discontinuities.

When we apply the algorithm to broad-band data with sensitivity to the transition-zone structure, we vary the values of the parameters at these discontinuities (seismic velocities and anisotropy just above and just below the discontinuity) and interpolate linearly between them. Below 660 km, the model is the isotropic AK135.

The locations of the control points in the mantle are fixed in each strand of the inversion (each Markov chain) and are spaced denser in the uppermost mantle, where we have more sensitivity. In different chains, the spline control points are different. The combination of the results of all the chains then minimizes the dependency of the final result on the parametrization. The same parametrization is used for the profiles of the isotropic-average shear-wave velocity, radial anisotropy and or azimuthal anisotropy coefficients (Fig. 5).

#### 4.4 Inversion for radial anisotropy and isotropic-average $V_s$

The first implementation of the algorithm is designed to determine the 1-D radially anisotropic, shear velocity structure of the crust and upper mantle using Rayleigh- and Love-wave phase-velocity curves. Such curves would normally be either a direct result of interstation dispersion measurements (e.g. Lebedev *et al.* 2006; Endrun *et al.* 2008; Agius & Lebedev 2014) or extracted from sets of phase-velocity maps at different periods, pre-computed using many source-station or station-station phase-velocity measurements (e.g., Shapiro & Ritzwoller 2002; Calò *et al.* 2016).

The quantitative comparison of the observed data and the synthetic data predicted by a given model is performed in the Bayesian inference approach, using a likelihood function (14). Because we invert Rayleigh and Love dispersion measurements jointly, the data vector  $\mathbf{d}$  and covariance matrix  $\mathbf{C}_e$  in eq. (14) comprise the data vectors and covariance matrices for the two wave types:

$$\mathbf{d} = [\mathbf{d}^R, \mathbf{d}^L], \quad (15)$$

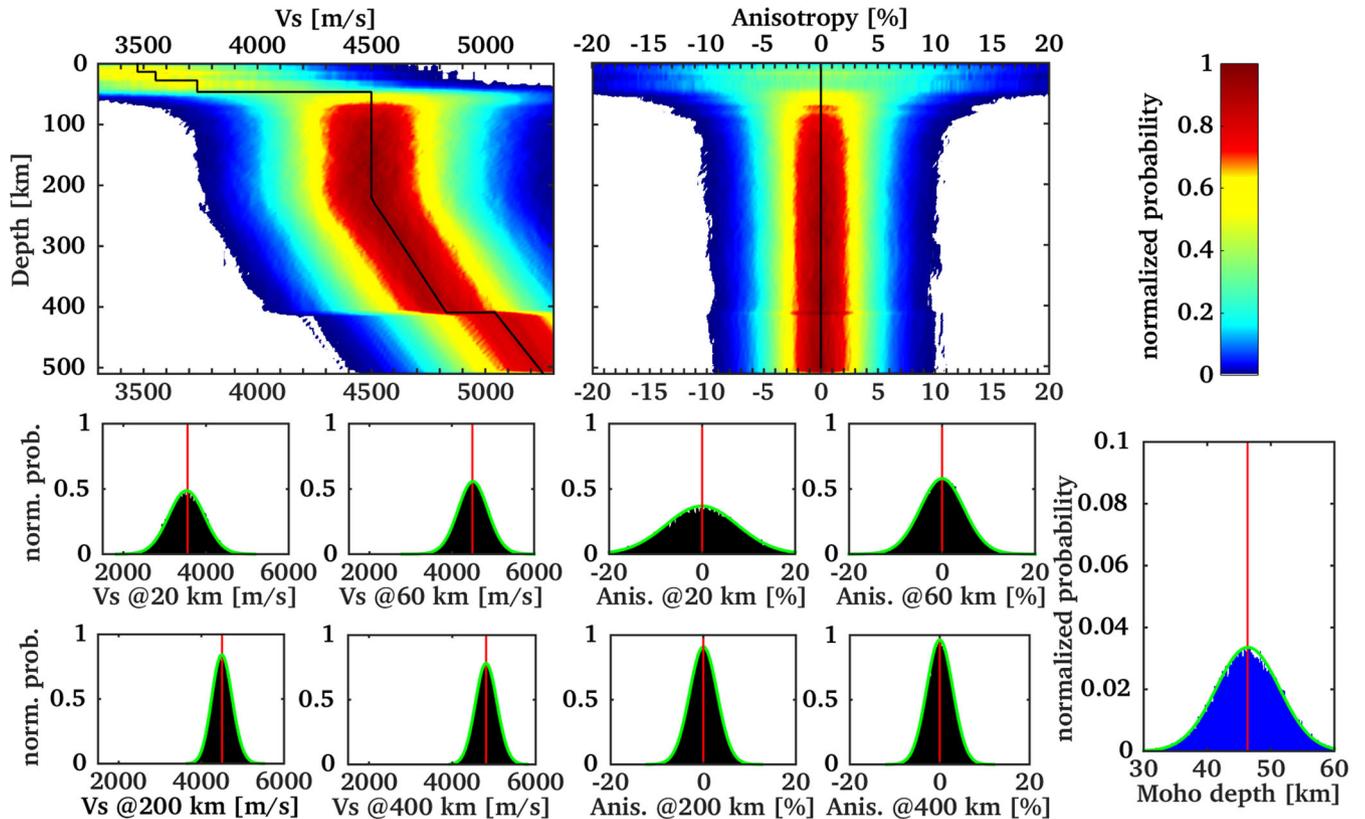
$$\mathbf{C}_e = \begin{bmatrix} \mathbf{C}_e^R & 0 \\ 0 & \mathbf{C}_e^L \end{bmatrix}. \quad (16)$$

With the assumption of uncorrelated measurement errors, the covariance matrices in eq. (16) are diagonal, and the total log-likelihood function [neglecting some constant terms in eq. (14)] can be written as the sum of Rayleigh and Love terms:

$$\log(p(\mathbf{d}|\mathbf{x})) \propto (\mathbf{d} - g(\mathbf{x}))^T \mathbf{C}_e^{-1} (\mathbf{d} - g(\mathbf{x})) = \sum_{i=1}^N \left( \frac{d_i^R - g(\mathbf{x})_i^R}{\sigma_i^R} \right)^2 + \sum_{i=1}^N \left( \frac{d_i^L - g(\mathbf{x})_i^L}{\sigma_i^L} \right)^2 \quad (17)$$

where  $d_i^R$ ,  $d_i^L$ ,  $g(\mathbf{x})_i^R$  and  $g(\mathbf{x})_i^L$  are the Rayleigh and Love dispersion curves observed and predicted for a given model  $\mathbf{x}$  at period  $T_i$ , and  $\sigma_i^R$ , and  $\sigma_i^L$  are the standard deviations of Rayleigh and Love data noise at period  $T_i$ . The sums are computed over the entire period range of the dispersion curve.

We model the noise with a function that varies with period, both for Rayleigh and Love dispersion curves. The frequency-dependency function is fixed during the inversion but is scaled by a multiplication factor that we allow to vary during the inversion,



**Figure 6.** Marginal prior probability distributions for isotropic-average shear velocity ( $V_s$ ), radial anisotropy and Moho depth, used in the radial anisotropy inversions of broad-band synthetic and real phase velocities (Section 5). The Gaussian shape of the distributions at different depths results from the combination of the Gaussian prior distribution set on parameters at each depth and the Gaussian prior constraint on neighbouring parameters. The reference models chosen for  $V_s$  and anisotropy are shown in black (top figures), whereas the values of the reference models at given depths are shown in red (bottom figures). Green lines represent normal fits of the sampled distributions.

treating it as an unknown hyperparameter, as in the Hierarchical Bayes approach:

$$\sigma^R(t) = h_R f(t), \quad (18)$$

$$\sigma^L(t) = h_L g(t). \quad (19)$$

The functions  $f(t)$  and  $g(t)$ , describing the dependencies of Rayleigh and Love noise standard deviations with period, are computed as envelopes of the phase-velocity differences between the observed dispersion curves and the ones computed for the best-fitting model from a weakly regularized non-linear gradient-search inversion (Section 3.1).

The prior distributions that we associate with the model parameters are based on reasonable assumptions regarding the models. Here, the model parameters are  $V_s$  (isotropic shear velocity as a function of depth),  $\mathbf{R}_{\text{ani}}$  (radial anisotropy as a function of depth) and the Moho depth. For the isotropic shear velocities, we choose at each depth a Gaussian prior distribution centred at the reference model. The reference crustal model varies depending on the region of study and is taken from the 3-D global crustal model CRUST 2; the reference model for the mantle is AK135. For the radial anisotropy parameters, although the 1-D, anisotropic global reference model PREM (Preliminary Reference Earth Model; Dziewonski & Anderson 1981) exhibits positive radial anisotropy in the upper mantle, we know that negative anisotropy in the crust and upper mantle is also present in some regions of the Earth. Therefore, we do not introduce *a priori* structure and set, at each depth, a Gaussian prior distribution for radial anisotropy that is centred at 0.

For the Moho depth, we set a Gaussian prior distribution centred at a reference value taken from CRUST 2, or from a regional study if available. For the hyperparameters  $h_R$ , and  $h_L$ , we set a uniform prior distribution over the interval 0–10.

In order to discourage unrealistic oscillatory models (caused by the overfitting of noisy data), we set a Gaussian prior distribution on the differences between seismic velocities in neighbouring layers. A valid alternative to the use of such smoothing constraints would be a trans-dimensional Bayesian approach, where the number of parameters changes during the inversion, and overcomplicated models are penalized because of ‘natural parsimony’ (Denison *et al.* 2002; Malinverno 2002; Sambridge *et al.* 2006; Piana Agostinetti & Malinverno 2010). The advantage of our simple approach is that it may retrieve better the subtle variations of seismic-velocity anomaly with depth in the upper mantle.

The complete mathematical form of our prior distribution is detailed in the Supporting Information (Supplement A). To prevent the posterior distribution being influenced by the form of the prior, the standard deviations of the Gaussian prior distributions are set at relatively large values in all the inversions. Fig. 6 shows an example of the prior distributions used in real data inversions. Even if a Gaussian distribution is centred around a reference value, it does not have rigid bounds, so that every parameter value in the model space has a non-zero probability of being sampled.

The implementation of the algorithm follows the rules of the Metropolis–Hastings algorithm (Metropolis *et al.* 1953; Hastings 1970), based on ‘Markov chain’ sampling. At each iteration of the inversion algorithm, the values of the parameters of the updated

model depend only on the values of the parameters of the previous model. After randomly selecting an initial model, the algorithm proceeds iteratively. Each step of the Markov chain is divided into three stages:

(1) Propose a new model by drawing from a probability distribution  $q(\mathbf{x}'|\mathbf{x})$  that perturbs the current model  $\mathbf{x}$ . This is done by choosing at random between one of the six possible moves:

(i) *Change crustal  $V_S$* : perturb isotropic velocities of all crustal layers according to a multidimensional Gaussian distribution centred on the current velocity values.

(ii) *Change crustal  $R_{ani}$* : change radial anisotropy values of all crustal layers according to a multidimensional Gaussian distribution centred on the current radial anisotropy values.

(iii) *Change mantle  $V_S$* : update isotropic velocities of all control points in the mantle according to a multidimensional Gaussian distribution centred on the current velocity values.

(iv) *Change mantle  $R_{ani}$* : change radial anisotropy values of all control points in the mantle according to a multidimensional Gaussian distribution centred on the current radial anisotropy values.

(v) *Move Moho depth*: perturb the depth of the Moho interface according to a Gaussian probability distribution centred at the current Moho depth value.

(vi) *Change one noise parameter*: randomly select one of the two hyperparameters ( $h_R, h_L$ ) and perturb its value according to a Gaussian distribution centred on the current hyperparameter value.

(2) Compute synthetic fundamental-mode phase velocities for the proposed model and build the likelihood function for the proposed model (17) by comparing synthetic and observed phase velocities.

(3) Compute acceptance probability for the proposed model (Hastings, 1970) using Bayes' rule (eq. 13). If the proposed model is accepted then it is added to the posterior distribution, otherwise it is discarded and the current model is retained for the next step.

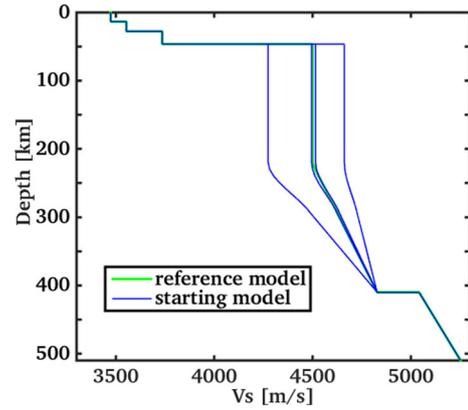
The detailed forms of the proposal distributions used to perturb the model parameters are reported in Supplement B. The expression and the calculation of the acceptance probabilities of a proposed model are reported in Supplement C.

After an initial 'burn-in' period, the models generated by the chain are distributed according to the posterior probability distribution (Mosegaard & Tarantola 1995; Mosegaard *et al.* 1997; Gallagher *et al.* 2009). In our inversions, we usually discard the first  $1 \times 10^5$  models as burn-in steps, and every 100th model visited is selected for the posterior distribution ensemble.

To speed up the computations, we parallelized the code in order to compute the synthetic phase velocities on different processors. Phase velocities in different period ranges are computed at the same time on 24 parallel CPU cores, with the parallelization implemented using the FORTRAN MPI libraries. Posterior inference is made using an ensemble obtained by merging the posterior distributions of four chains (each made of  $10^6$  iterations) running independently for different starting models. Fig. 7 shows examples of different starting velocity models, each of them isotropic.

#### 4.5 Inversion for azimuthal anisotropy

We developed another implementation of our algorithm for the inversion of phase-velocity anisotropy for the azimuthal anisotropy of shear-wave velocities (7) as a function of depth. The data are



**Figure 7.** Example of four different, randomly chosen, shear velocity models (blue lines) set as starting models in the radial anisotropy inversions of broad-band synthetic and real phase velocities (Section 5). The models are isotropic ( $V_{sv} = V_{sh} = V_s$ ) and share common features such as a flat velocity profile from the Moho to 220 km and a positive velocity gradient below.

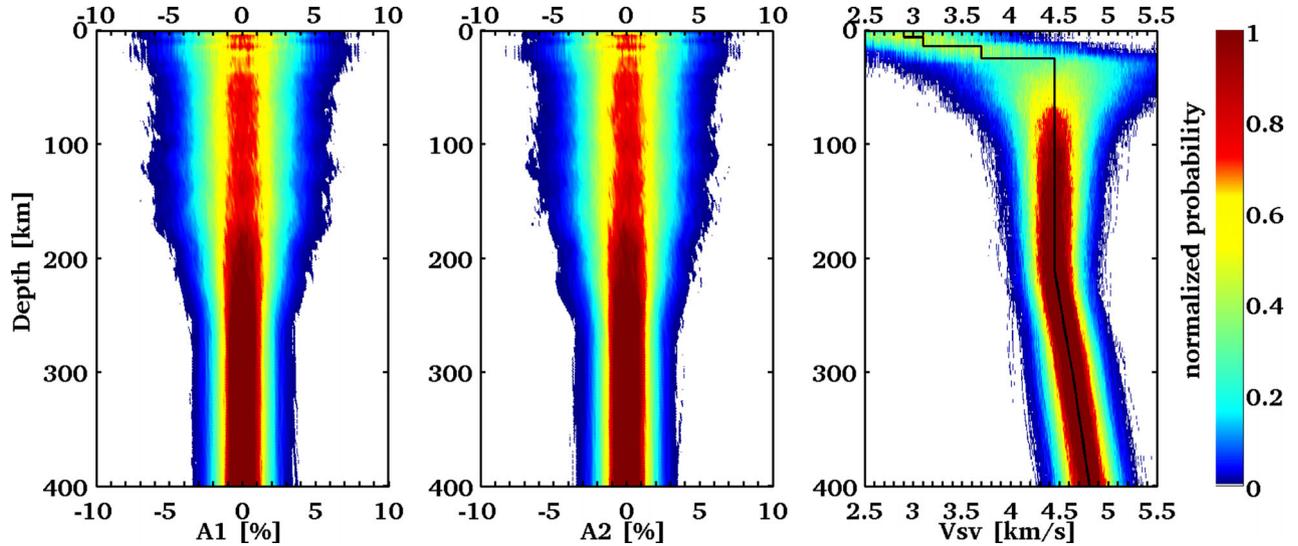
isotropic-average phase velocities and their azimuthal anisotropy parameters as a function of period. These can be determined by anisotropic phase-velocity tomography at a range of periods (e.g. Endrun *et al.* 2008; Lin *et al.* 2011) or by inversion of array data for region-average anisotropy of phase velocities (e.g. Pedersen *et al.* 2006; Adam & Lebedev 2012; Agius & Lebedev 2017).

The approach used for the implementation is based on a simultaneous inversion of Rayleigh-wave dispersion curves observed at a number of equally spaced azimuths (Section 3.3). In the Bayesian framework, the full data vector and covariance matrix comprise the azimuth-specific data vectors and covariance matrices, as in eqs (15) and (16). In case of  $n$  azimuths, the total log-likelihood function can be written as the sum of all the contributions at different azimuths:

$$\log(p(\mathbf{d}|\mathbf{x})) \propto \sum_{i=1}^N \left( \frac{(d_i^R)_{\varphi_1} - (g(\mathbf{x})_i^R)_{\varphi_1}}{(\sigma_i^R)_{\varphi_1}} \right)^2 + \sum_{i=1}^N \left( \frac{(d_i^R)_{\varphi_2} - (g(\mathbf{x})_i^R)_{\varphi_2}}{(\sigma_i^R)_{\varphi_2}} \right)^2 + \dots + \sum_{i=1}^N \left( \frac{(d_i^R)_{\varphi_n} - (g(\mathbf{x})_i^R)_{\varphi_n}}{(\sigma_i^R)_{\varphi_n}} \right)^2, \quad (20)$$

where  $(d_i^R)_{\varphi_j}$  and  $(g(\mathbf{x})_i^R)_{\varphi_j}$  are the Rayleigh dispersion curves observed and computed for a model  $\mathbf{x}$  at a period  $T_i$  and azimuth  $\varphi_j$  ( $j = 1, \dots, n$ ), and  $(\sigma_i^R)_{\varphi_j}$  are the standard deviations of Rayleigh data noise at the period  $T_i$  and azimuth  $\varphi_j$  ( $j = 1, \dots, n$ ), with the assumption of uncorrelated noise. In this case, we do not treat errors of each azimuth-specific dispersion curve (which, generally, can be different, given the actual orientation of anisotropy at the location) as unknowns in the inversion, because this would require the introduction of a high number of hyperparameters (at least one for each azimuth). Instead, we estimate the error from the difference between the observed dispersion curves at each azimuth and the synthetics computed for a best-fit model using the anisotropic gradient-search inversion described in Section 3.2.

The parameters of the MCMC inversion algorithm are  $\mathbf{V}_{sv}^0$ ,  $\mathbf{A}_1$  and  $\mathbf{A}_2$  (azimuthally isotropic vertically polarized shear velocities and azimuthal perturbations, considered here as vectors because they are functions of depth and have different values at the control points at different depths in our inversion set up), and the Moho depth.



**Figure 8.** Marginal prior probability distributions for  $2\varphi$  azimuthal perturbations ( $A_1$ ,  $A_2$ ) and azimuthally isotropic shear velocity ( $V_{sv}^0$ ), used in the inversions of synthetic and real phase velocities of Section 6. The reference model chosen for shear velocities is shown in black (right), whereas the reference model chosen for  $A_1$  and  $A_2$  is equal to 0 in the entire depth range. The distributions are normally distributed around the reference models, with a standard deviation that decreases with increasing depths.

For  $V_{sv}^0$ , we use at each depth a Gaussian prior distribution that is centred at the reference model (Fig. 8, right). A Gaussian prior is also applied to the difference between velocities within neighbouring depth ranges.

For  $A_1$  and  $A_2$  ( $2\varphi$  azimuthal perturbations), we choose at each depth a Gaussian prior distribution centred at 0 and a Gaussian prior distribution on the differences between parameters at neighbouring depths (Fig. 8, left and centre).

For the Moho depth, we set a Gaussian prior distribution centred at the reference crustal model (see Supplement A for the complete mathematical form of the prior distribution).

Each step of the Markov chain comprises three stages:

(1) Propose a new model by drawing from a probability distribution  $q(\mathbf{x}'|\mathbf{x})$  that perturbs the current model  $\mathbf{x}$ ; this is done by choosing at random between one of the seven possible moves:

(i) *Change crustal  $V_{sv}^0$* : change azimuthally isotropic velocity values of all layers according to a multidimensional Gaussian distribution centred on current velocity values.

(ii) *Change crustal  $A_1$* : update azimuthal perturbations  $A_1$  of all crustal layers according to a multidimensional Gaussian distribution centred on the current  $A_1$  values.

(iii) *Change crustal  $A_2$* : update azimuthal components  $A_2$  of all layers according to a multidimensional Gaussian distribution centred on the current  $A_2$  values.

(iv) *Change mantle  $V_{sv}^0$* : change azimuthally isotropic velocity values of all control points in the mantle according to a multidimensional Gaussian distribution centred on the current velocity values.

(v) *Change mantle  $A_1$* : update azimuthal components  $A_1$  of all control points in the mantle according to a multidimensional Gaussian distribution centred on the current  $A_1$  values.

(vi) *Change mantle  $A_2$* : update azimuthal perturbations  $A_2$  of all control points in the mantle according to a multidimensional Gaussian distribution centred on the current  $A_2$  values.

(vii) *Move Moho depth*: perturb the depth of the Moho interface according to a Gaussian probability distribution centred at the current Moho depth value.

(2) Compute synthetic phase velocities at the different azimuths for the proposed model. To perform the computation, eq. (7) is used to derive the azimuth-dependent shear velocities from  $V_{sv}^0$ ,  $A_1$  and  $A_2$ ; then they are given as inputs to the MINEOS forward solver to compute Rayleigh-wave fundamental-mode dispersion curves at each azimuth. The likelihood function is then obtained from eq. (20).

(3) Compute acceptance probability for the proposed model. If the proposed model is accepted, then it is added to the posterior distribution, otherwise it is discarded and the current model is retained for the next step. (The detailed forms of the proposal distributions and the calculation of the acceptance probabilities of a proposed model are reported in Supplements B and C.)

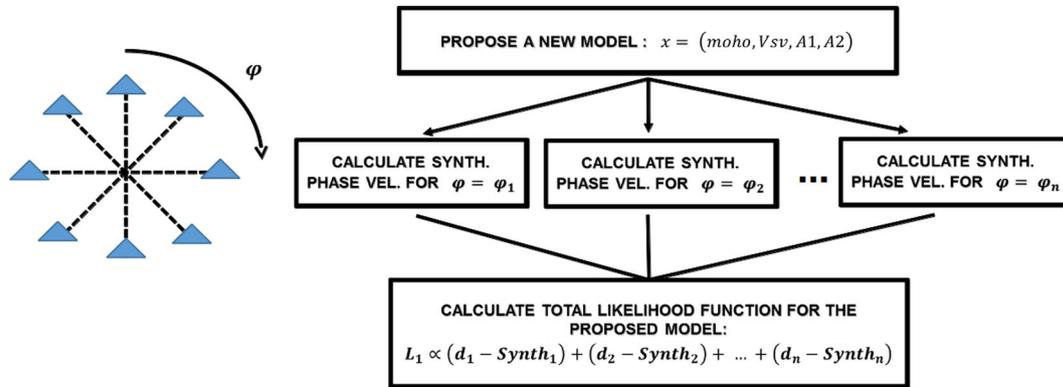
To speed up the computations, we parallelized the code and computed synthetic phase velocities at different azimuths on different processors at the same time (the parallelization has been implemented using the FORTRAN MPI libraries). The workflow of the algorithm is shown in the simple diagram in Fig. 9.

In our inversions of the azimuthally anisotropic phase velocities from west-central Italy (Section 6.2), we compute at the same time synthetic phase velocities at nine different azimuths on nine parallel CPU cores.

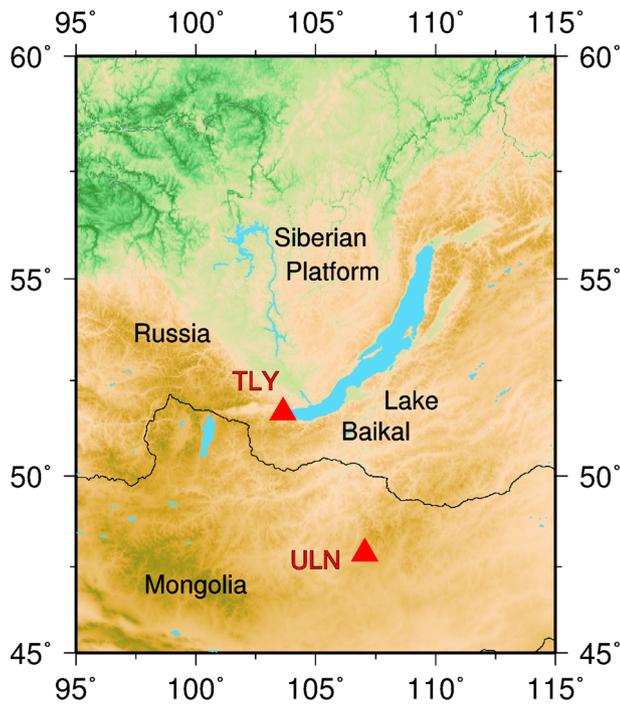
## 5 RADIAL ANISOTROPY: MONGOLIA

We now test our radial anisotropy inversion on the dispersion curves derived previously (Lebedev *et al.* 2006) from cross-correlation and waveform-inversion measurements using the Global Seismographic Network (GSN) stations Talaya (TLY) in Russia and Ulaanbaatar (ULN) in Mongolia (Fig. 10). The Rayleigh- and Love-wave, fundamental-mode phase-velocity curves have very wide period ranges: 9–333 s for Rayleigh and 8–341 s for Love waves.

We first resample the dispersion curves at a logarithmic sample spacing so as to give more equal weight to the structural sensitivity of the data at different frequencies and then decimate the dispersion curves following the procedure described in Section 3.1, in order to reduce the computational time of the inversions. We choose a resampling factor of 0.5, reducing the number of points of the



**Figure 9.** Schematic representation illustrating the implementation of the MCMC algorithm used to resolve  $2\varphi$  azimuthal anisotropy. In this sketch (left), station pairs are located at a number of different azimuths. When a new model is proposed (right), the computation of the synthetic dispersion curves predicted at each azimuth is performed in parallel. The total log-likelihood function is defined as the sum of the log-likelihood functions computed for different azimuths.

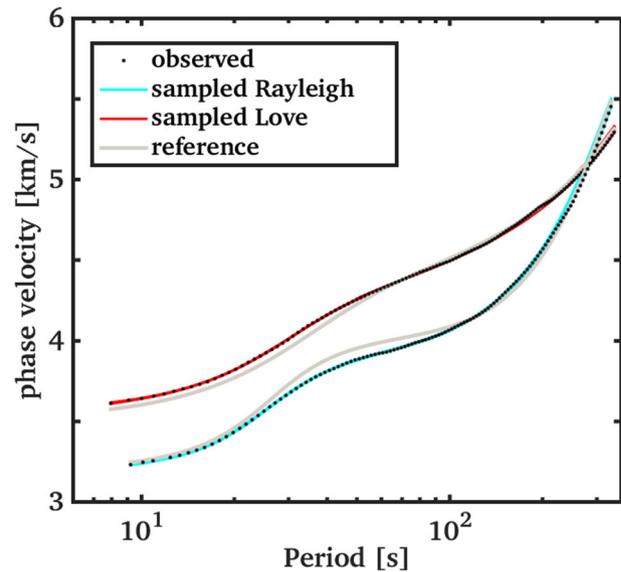


**Figure 10.** Baikal-central Mongolia region. The Rayleigh- and Love-wave dispersion curves used in our inversions are averages along the corridor between the GSN seismic stations Talaya (TLY) and Ulaanbaatar (ULN), depicted as red triangles.

dispersion curves from 240 and 236 to 120 and 118 for Love and Rayleigh curves, respectively. The resulting phase-velocity curves are shown in Fig. 11 (black dots).

We allow the amplitude of the data noise to vary during the inversions, the variations controlled by the hyperparameters  $h_R$  and  $h_L$ . The functions describing the dependency of Rayleigh and Love standard deviations with period were derived using a moderately damped gradient-search inversion of the averaged phase-velocity curves (Section 3.1). They are characterized by monotonically increases for both the Rayleigh and Love dispersion curves.

The prior distributions chosen for the isotropic-average shear speed, radial anisotropy and Moho depth are as in Fig. 6. The reference model for the isotropic-average shear velocities in the crust is a three-layered crustal model for the region taken from CRUST2.0. Perturbations in both the isotropic-average shear speed and radial



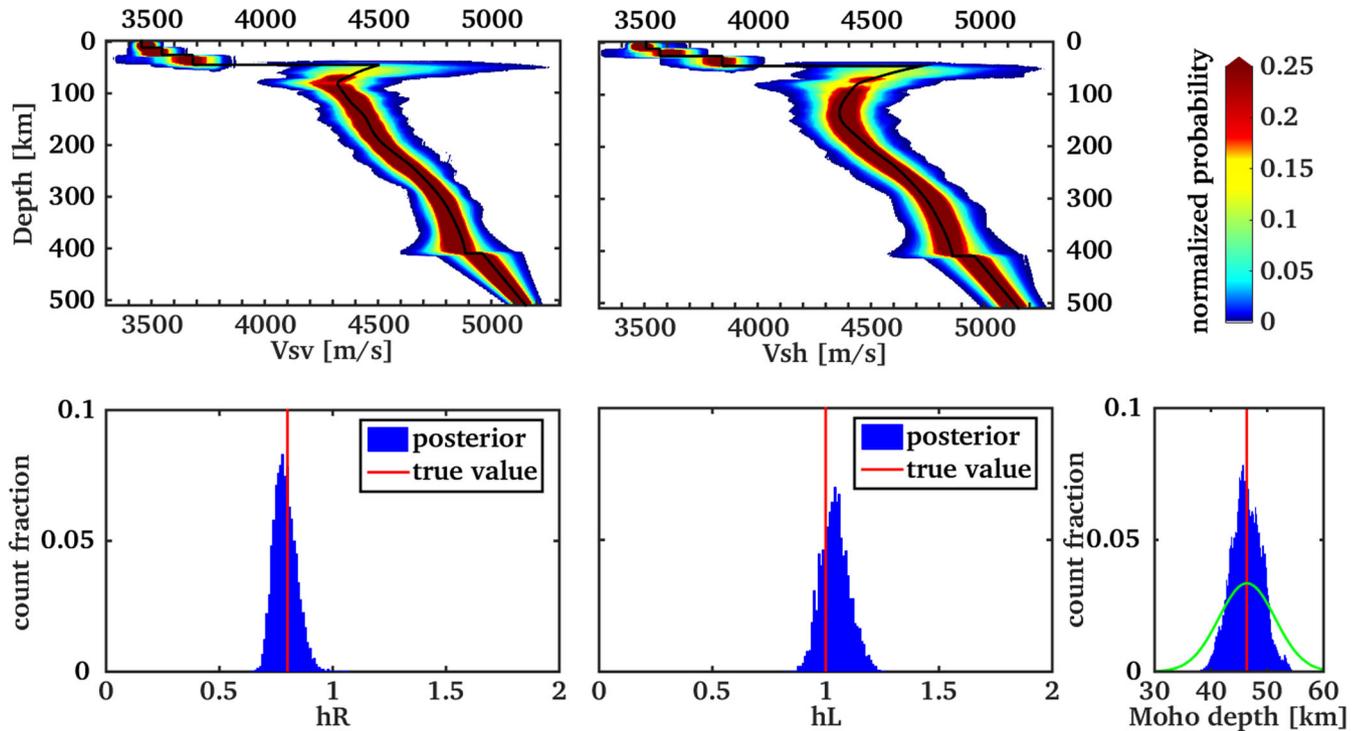
**Figure 11.** Rayleigh and Love, broad-band, phase-velocity curves observed and sampled in real-data inversion for radial anisotropy in the Baikal-Mongolia region. The observed data are plotted as black dots. The blue (Rayleigh) and red (Love) lines represent the dispersion curves corresponding to the sampled models of the posterior distribution. The synthetic dispersion curves of the reference model used in the inversion (black lines in Fig. 6) are shown with thick grey lines.

anisotropy are parametrized using three crustal layers (boxcar basis functions) and 11 control points in the mantle (splines).

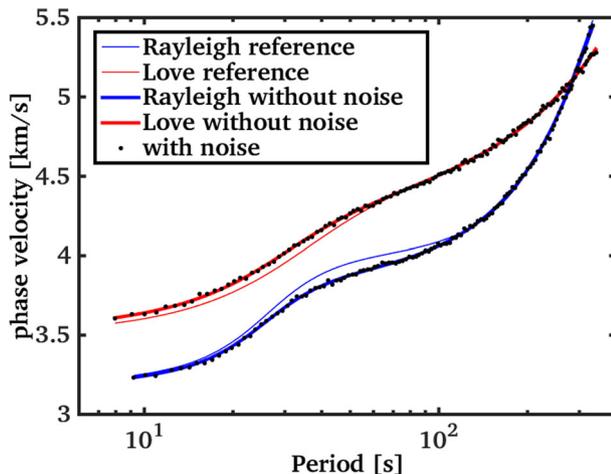
### 5.1 Synthetic tests

We first test the radial anisotropy algorithm with synthetic data computed for a radially anisotropic model produced by a gradient-search inversion of the data with moderate damping. This model (black line in Fig. 12) has a relatively thick crust ( $\sim 46$  km Moho depth) with three layers and discontinuities at  $\sim 13$  and  $\sim 27$  km depth; below the Moho the lithosphere presents strong radial anisotropy ( $V_{sh} > V_{sv}$ ) and a seismic velocity decrease down to  $\sim 80$  km depth. The synthetic Rayleigh- and Love-wave dispersion curves are calculated from the model with MINEOS (Section 3.1).

Uncorrelated random noise with known variance was added to the synthetic dispersion curves. Fig. 13 shows the synthetic Rayleigh- and Love-wave dispersion curves used in the inversions: the thick



**Figure 12.** Marginal posterior probability distributions for  $V_{sv}$ ,  $V_{sh}$ , noise hyperparameters ( $h_R$ ,  $h_L$ ) and the Moho depth, resulting from the inversion of synthetic data with added uncorrelated noise. The true synthetic model is shown in black (upper frames), and the true values of the hyperparameters and the Moho depth are shown in red (bottom figures). The marginal prior probability on the Moho depth is shown as a green line superimposed on the marginal posterior distribution of the sampled Moho depths (bottom right).



**Figure 13.** Rayleigh and Love broad-band synthetic phase-velocity curves used in the synthetic inversions for radial anisotropy. The thick blue (Rayleigh) and red (Love) lines represent the dispersion curves computed from a known radially anisotropic continental model with no noise added, whereas the black dots represent the same curves but with added uncorrelated noise. The synthetic dispersion curves of the reference model used in the inversions (black lines in Fig. 6) are shown in thin blue (Rayleigh) and red (Love).

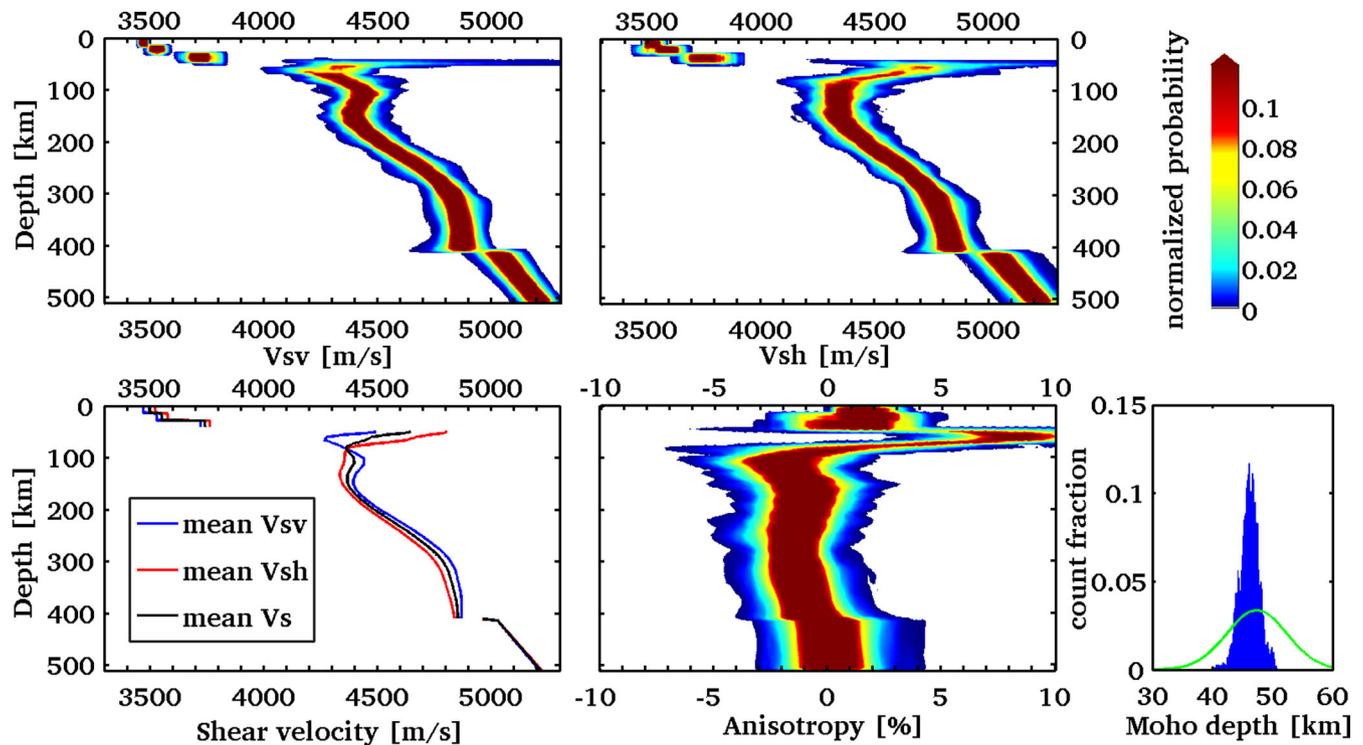
blue (Rayleigh) and red (Love) lines are the synthetic curves with no noise added; the black dots show the synthetic curves with added uncorrelated noise, generated according to a diagonal covariance matrix  $C_c$  with a standard deviation that increases with period from 0.2 to 0.3 per cent of the synthetic phase-velocity values, for both Rayleigh and Love curves.

During the inversion, the parameters of the radially anisotropic shear speed structure perturb isotropic-average  $V_s$  and radial anisotropy in three crustal layers and at 11 control points in the mantle. The Moho depth is also a parameter of the inversion. The control points in the mantle (Fig. 5) are fixed at chosen depths and are denser in the uppermost mantle depth range, where the depth resolution of surface waves is higher. Multiple Markov chains are run, with control points at different locations in the depth range between the Moho depth and 410 km, in order to minimize the dependency of the inversion results from the parametrization.

Two hyperparameters are also treated as unknowns, as in the Hierarchical Bayes framework (Section 4.4); these are  $h_R$  and  $h_L$  that multiply our estimates of the Rayleigh and Love noise standard deviations ( $\sigma^R(t) = h_R f(t)$ ,  $\sigma^L(t) = h_L g(t)$ ). The ‘true’ values chosen for these tests are  $h_R = 0.8$  and  $h_L = 1$ , with  $f(t)$  and  $g(t)$  monotonically increasing with period.

The results of the inversion are shown in Fig. 12. The profiles were obtained by merging the posterior probability distributions of four different Markov chains, each consisting of  $1.5 \times 10^6$  iterations and starting with a different random model (Fig. 7). The results demonstrate the ability of the inversion algorithm to resolve the radially anisotropic shear velocity structure of the crust and upper mantle. The profiles capture the smooth mantle structure, and resolve the sharp gradient in  $V_s$  and anisotropy in the uppermost mantle.

The results also show that the broad period range covered by the synthetic phase velocities allow us to resolve both  $V_s$  and anisotropy down to 410 km and below, despite our including only the fundamental mode in the computation of the forward problem. The posterior distribution of the sampled Moho depths is peaked around the true model (red line).



**Figure 14.** Marginal posterior probability distributions for  $V_{sv}$ ,  $V_{sh}$ , radial anisotropy and the Moho depth, resulting from the inversion of the dispersion curves from Baikal-central Mongolia. The posterior mean models of the  $V_{sv}$ ,  $V_{sh}$  and  $V_s$  distributions are plotted with blue, red and black lines, respectively (bottom left). The marginal prior probability on the Moho depth is shown as a green line superimposed on the marginal posterior distribution of the sampled Moho depths (bottom right).

Fig. 12 shows how a Hierarchical Bayesian procedure can remedy insufficient prior knowledge of the noise in the data. Here, we only assumed that the dependence of Rayleigh and Love standard deviations on period was described by the monotonically increasing functions  $f(T)$  and  $g(T)$ , which were kept fixed during the inversions. However, we varied the magnitude of data noise during the inversions by means of the hyperparameters  $h_R$  and  $h_L$ , and we were able to retrieve their ‘true’ values (0.8 and 1).

## 5.2 Real-data inversions

We now apply the inversion to the measured dispersion curves. The resulting profiles (Fig. 14) are obtained by merging the posterior probability distributions of four different Markov chains, each consisting of  $1.5 \times 10^6$  iterations. The profiles represent the 1-D, radially anisotropic shear velocity structure averaged along the corridor between the two stations in the Baikal-Mongolia region. The phase-velocity curves corresponding to the sampled velocity models (the models selected for the posterior distribution ensembles) are shown in Fig. 11 (blue lines for Rayleigh and red lines for Love).

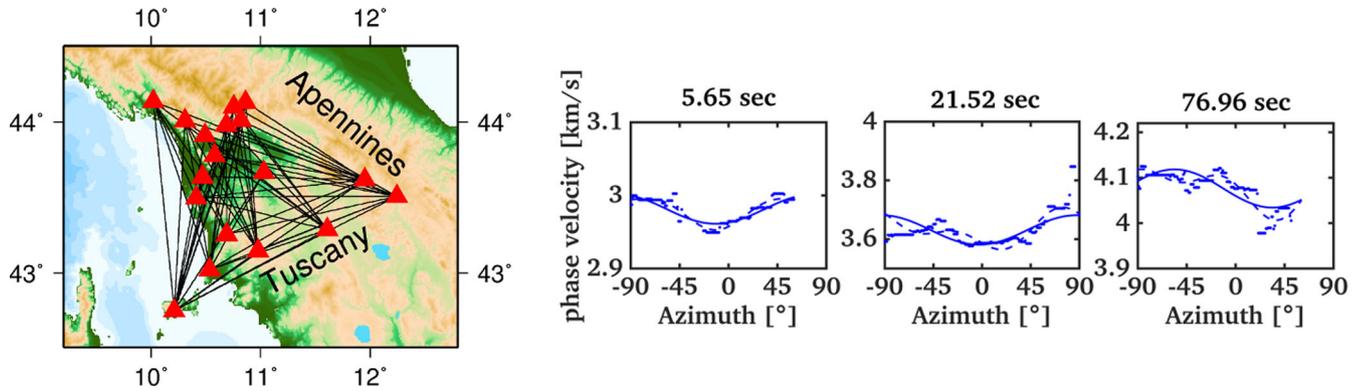
The profiles are similar, in their main features, to those obtained in previous inversions of the same data (Lebedev *et al.* 2006; Fulla *et al.* 2012). Unlike the previous inversions, however, they yield probability distributions for the inversion parameters and convey the model uncertainty. The maximum amount of radial anisotropy (7 per cent on average) is observed in the mantle-lithosphere depth range (50–80 km). The low-velocity zone below  $\sim 80$  km indicates the asthenosphere. The positive anisotropy ( $V_{sh} > V_{sv}$ ) at depths up to  $\sim 100$  km is consistent with ‘frozen-in’ fabric associated with pervasive lithospheric deformation in the past (Lebedev *et al.* 2006).

The observation of a small amount of negative anisotropy ( $V_{sv} > V_{sh}$ ) below 100 km (1 per cent on average) suggests vertical flow in the asthenosphere and is particularly interesting. Hints of this negative anisotropy have been seen earlier (Lebedev *et al.* 2006; Fulla *et al.* 2012), but the non-uniqueness of the models precluded interpreting this small signal with confidence. Our probabilistic models here show that, although relatively small in amplitude, this anisotropy is robust. We interpret it as evidence for the ascent of the asthenosphere as it flows from beneath the thick lithosphere of the Siberian Craton southeastward, to beneath the thinner lithosphere of Mongolia. The resulting decompression melting of veins of enriched material within it (Morgan & Morgan 1999) is likely to give rise to the scattered, sporadic intraplate basaltic volcanism in the Baikal-Mongolia region. The origin of Baikal-region volcanism due to subhorizontal flow of the asthenosphere beneath the lithosphere of strongly varying thickness has been discussed in detail by Lebedev *et al.* (2006), but our probabilistic inversions here are the first to provide convincing evidence for the vertical component of the flow, essential to this mechanism.

The observed posterior distribution of the sampled Moho depths (Fig. 14, bottom right) is centred on an average value of about 45.5 km, somewhat shallower than the reference value from CRUST 2.0 (47.3 km). The depth range spanned by the distribution is in agreement with the results of the previous studies (Lebedev *et al.* 2006; Fulla *et al.* 2012).

## 6 AZIMUTHAL ANISOTROPY: WEST-CENTRAL ITALY

Seismograms recorded by the RETREAT experiment in west-central Italy (Margheriti *et al.* 2006; Plomerová *et al.* 2006) were used



**Figure 15.** Left: the broad-band seismic stations of the RETREAT experiment (Margheriti *et al.* 2006; Plomerová *et al.* 2006) selected for measurements across Tuscany (Keogh *et al.* 2009), used in this study. Right: region-average, Rayleigh-wave phase velocities as a function of azimuth at three different periods, derived from the phase-velocity measurements for the various station pairs. Dots are the data, smoothed with a  $20^\circ$  sliding window, solid lines are the best-fitting curves with isotropic and  $2\varphi$  anisotropy terms only, and dashed lines are the best-fitting curves with isotropic,  $2\varphi$  and  $4\varphi$  anisotropy terms.

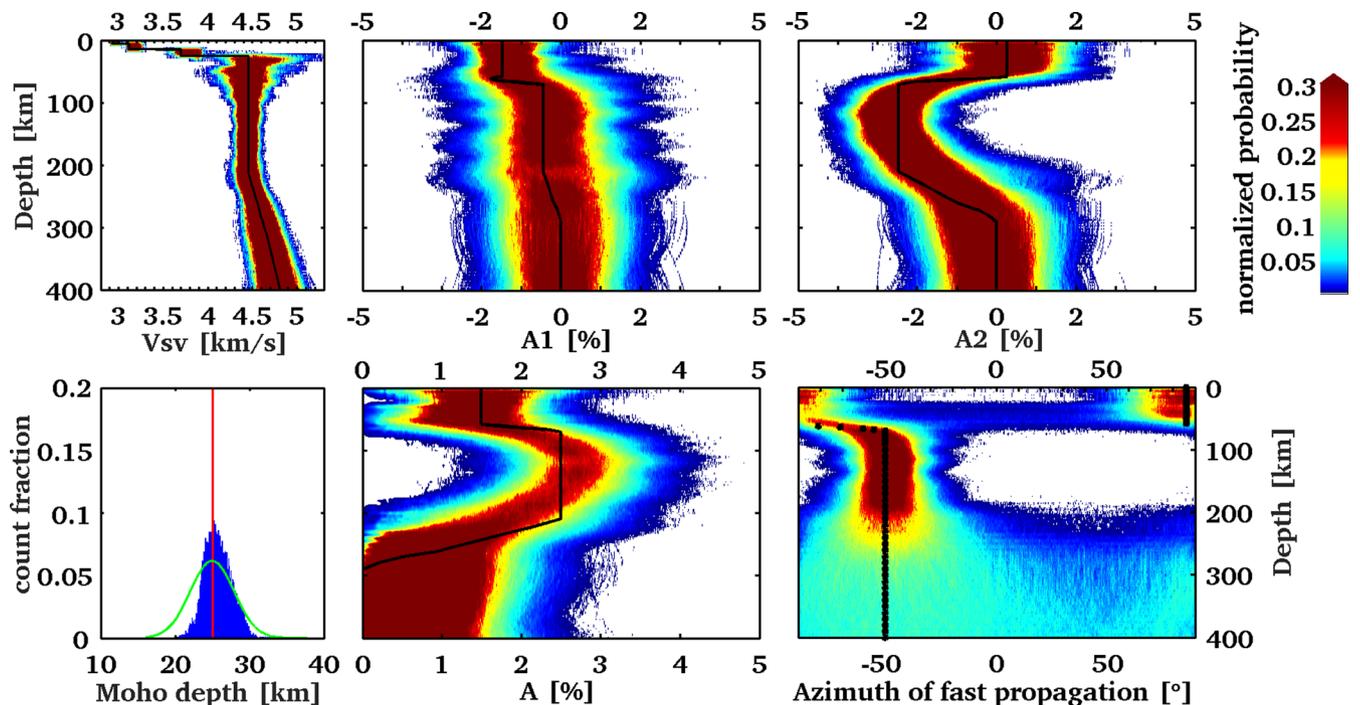
recently to measure surface-wave dispersion and derive region-average, Love and Rayleigh phase-velocity curves and Rayleigh-wave azimuthal anisotropy for Tuscany (Keogh *et al.* 2009; Lebedev *et al.* 2010, Fig. 15, left). In this section, we test our azimuthal anisotropy methods and invert the surface-wave data for azimuthal anisotropy beneath the region.

Following Adam & Lebedev (2012) and Adam (2013), Rayleigh phase-velocity measurements were sorted according to the station–station azimuth and smoothed with a  $20^\circ$  width sliding windows in order to average out noise. They were then inverted with a least-squares inversion algorithm (Adam & Lebedev 2012; Adam 2013) for best-fitting isotropic-average phase velocities and the  $2\varphi$  and  $4\varphi$  components of azimuthal anisotropy (6) (Keogh *et al.* 2009; Lebedev *et al.* 2010). Examples of best-fitting anisotropy curves for three selected periods are shown in Fig. 15 (right). Here, we apply the

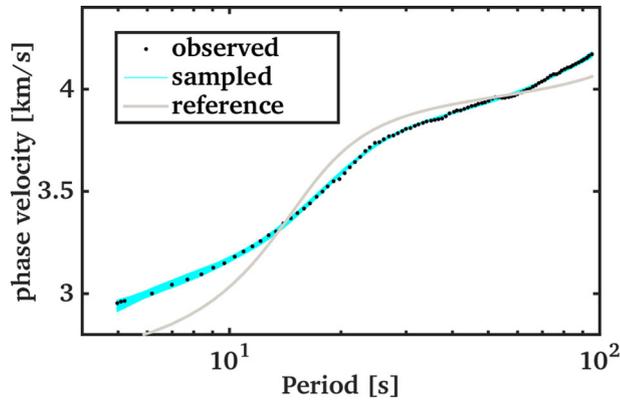
McMCM algorithm to invert the  $2\varphi$  azimuthal anisotropy of Rayleigh waves for a profile of depth-dependent azimuthal anisotropy in  $V_s$  (Section 6.2).

### 6.1 Synthetic tests

We now test the azimuthal anisotropy algorithm (Section 4.5) with synthetic data computed for an azimuthally anisotropic model (black lines in Fig. 16) that is similar in its basic features, in particular anisotropy features, to models that fit the Tuscany data. The assumed model has a Moho depth of 25 km with crustal discontinuities at 6 and 14 km depth, crustal shear velocity structure taken from Molinari *et al.* (2015), and a mantle  $V_{sv}$  structure similar to AK135. The assumed profile of the amplitude of



**Figure 16.** Marginal posterior probability distributions for  $V_{sv}^0$ ,  $A_1$ ,  $A_2$ , Moho depth, amplitude of anisotropy and the fast-propagation azimuth, resulting from the inversion of synthetic, noise-free data for  $2\varphi$  azimuthal anisotropy. The true synthetic model is shown in black. The true value of the Moho depth is shown in red, and the marginal prior probability on the Moho depth is shown as a green line superimposed on the marginal posterior distribution of the sampled Moho depths (bottom left).



**Figure 17.** Rayleigh-wave, phase-velocity curves observed and sampled in real data inversions for azimuthal anisotropy in the Tuscany region. The isotropic-average, measured phase velocities are plotted as black dots. The blue lines show the dispersion curves corresponding to the sampled models of the posterior distribution. The synthetic dispersion curve for the reference model used in the inversions (black line in Fig. 8, right) is plotted with a thick grey line.

azimuthal anisotropy ( $A(d) = \sqrt{A_1(d)^2 + A_2(d)^2}$ ) has three layers, with amplitude varying from 1.5 per cent in the crust and mantle lithosphere to 2.5 per cent at 100–200 km depth ranges, and to 0 per cent below 200 km. The fast-propagation azimuth ( $\Psi(d) = 0.5 \times \arctan(A_2(d)/A_1(d))$ ) changes sharply at 90 km depth, from  $80^\circ$  to  $-50^\circ$ .

The synthetic, fundamental-mode Rayleigh-wave dispersion curves are calculated with the MINEOS code (Section 4.5) at azimuths at a  $20^\circ$  interval between  $-90^\circ$  and  $70^\circ$  ( $-90^\circ, -70^\circ, -50^\circ, -30^\circ, -10^\circ, 10^\circ, 30^\circ, 50^\circ, 70^\circ$ ), in the same period range covered by the measured Rayleigh-wave phase velocities from Tuscany (5–95 s).

No noise is added to the synthetic dispersion curves, although the data noise standard deviations used for the computation of the covariance matrix  $C_e$  are set equal to the ones estimated for real data inversions (Section 6.2).

The parameters of the azimuthally anisotropic structure include three crustal layers and 10 control points in the mantle, both for the azimuthally isotropic shear speeds ( $\mathbf{V}_{sv}^0$ ) and the azimuthal perturbations ( $A_1, A_2$ ). The Moho depth is also a parameter of the inversion. The prior distributions chosen for  $\mathbf{V}_{sv}^0, A_1$  and  $A_2$  are shown in Fig. 8. The reference model for  $A_1$  and  $A_2$  is equal to 0 per cent in the entire depth range.

The resulting profiles (Fig. 16) were obtained by merging the posterior probability distributions of four different Markov chains, each consisting of  $6 \times 10^5$  iterations. The results demonstrate the potential of the inversion algorithm to resolve the  $2\varphi$  azimuthal anisotropy of the crust and upper mantle. Despite the smooth parametrization of the mantle structure, the sharp changes of the anisotropic parameters are well resolved. The Moho interface is also resolved, as the posterior distribution of the sampled Moho depths is peaked around the true value (red line).

## 6.2 Real-data inversions

We now use the reference model, parametrization and priors tested above to invert the observed data. Fig. 17 shows the observed isotropic-average component of Rayleigh-wave phase-velocity curves (black dots), together with the synthetic phase-velocity curves of the posterior distribution (blue lines). Resampling

with logarithmic sample spacing has been applied so as to give more equal weight to structural sensitivity at different frequencies.

Because of the high number of hyperparameters that a Hierarchical Bayesian approach would require (generally, at least one hyperparameter for each azimuth), we do not treat the data noise as an unknown in this case. The estimated errors of the Rayleigh-wave phase velocities were derived from a moderately damped gradient-search inversion (Section 3.2). The resulting estimates of data noise vary with azimuth and with period, from about 0.8 per cent to about 2 per cent of the phase-velocity values.

The posterior distributions of the model parameters (Fig. 18) were obtained by merging the posterior probability distributions of four different Markov chains, each consisting of  $6 \times 10^5$  iterations. The superimposed best-fitting profiles (pink lines) obtained with the azimuthally anisotropic gradient-search inversion algorithm (Section 3.2) highlight the consistency between results derived from two different methods of inversion.

The  $V_{sv}^0$  profile shows a sharp positive gradient just below the Moho, underlain by a low-velocity zone at asthenospheric depths. The posterior distribution of the sampled Moho depths (Fig. 18, bottom left) is centred on an average value that is very close to the reference value (25 km).

The azimuthal anisotropy profile shows a sharp transition of the fast-propagation azimuth ( $\Psi(d) = 0.5 \times \arctan(A_2(d)/A_1(d))$ ) from  $80^\circ$  to  $-50^\circ$ , corresponding to an increase of amplitude of anisotropy ( $A(d) = \sqrt{A_1(d)^2 + A_2(d)^2}$ ) from about 1 per cent to about 3 per cent. This change in the fast-propagation azimuth occurs between 60 and 100 km depths and is likely to reflect the LAB, where the fast axis direction of anisotropy moves from E–W at crustal and lithospheric-mantle depths to NW–SE in the asthenosphere. These results are consistent with SKS splitting observations—likely to reflect asthenospheric anisotropy in regions with relatively thin lithosphere, as it is here—that show NW–SE fast directions beneath Tuscany (Salimbeni *et al.* 2007). The results are also consistent with the E–W stretching of the lithosphere (Jolivet *et al.* 2009) giving rise to the strong anisotropic fabric, with fast-propagation directions parallel to the direction of extension, as observed elsewhere (Endrun *et al.* 2008).

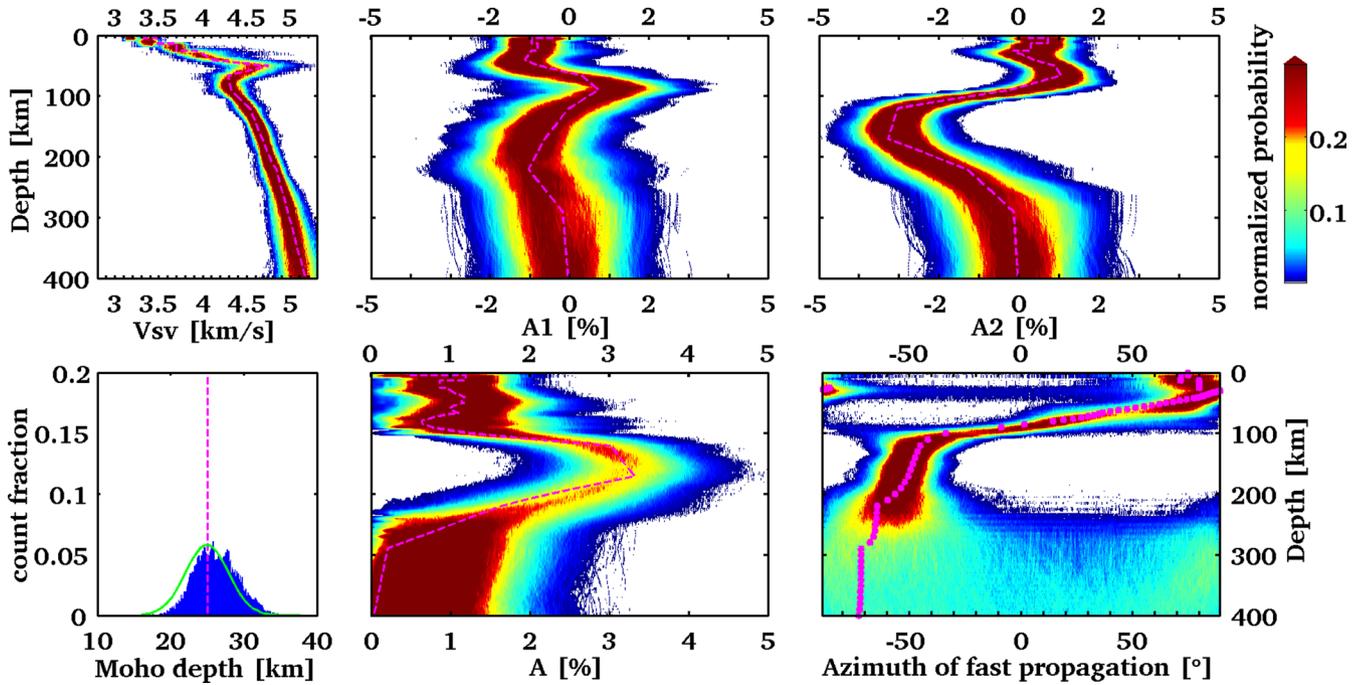
Fig. 19 (right) shows the variations of shear velocities with azimuth eq. (7) at four depths, in the lower crust, mantle lithosphere, shallow asthenosphere and below. The  $2\varphi$  anisotropy patterns are clearly different in the lithosphere and asthenosphere. At 320 km depth, no anisotropy is detected, which may well be due to the lack of sensitivity of our limited-frequency-range data in the deep upper mantle.

In Fig. 20, we show the distribution of  $2\varphi$  anisotropy with period. The red circles show the data; the solid and dotted lines show the phase velocities of the best-fitting models obtained with the MCMC inversion.

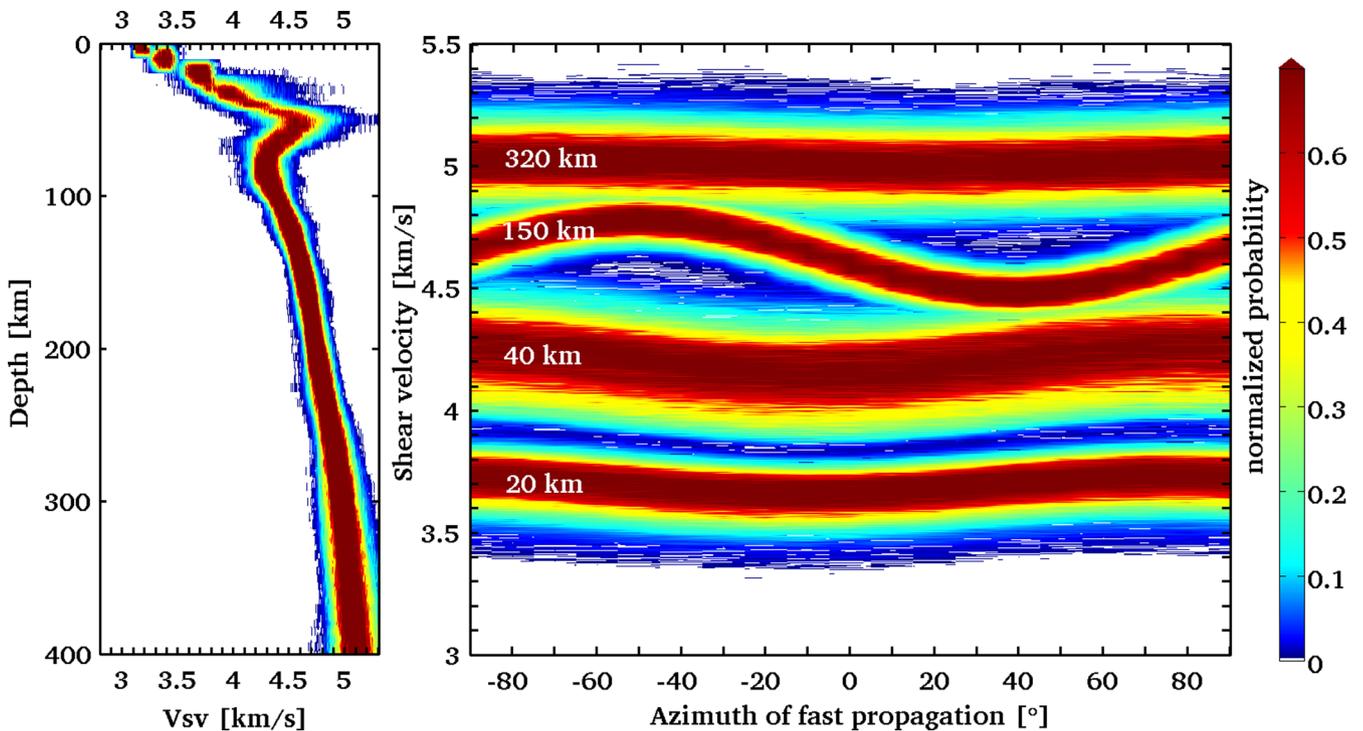
We can see that the fast azimuth is constrained better than the amplitude of anisotropy. The large amplitude of anisotropy differences observed at adjacent periods (red circles) cannot be reconciled with realistic Earth structures; the more gradual changes with period observed for the best-fitting curves (solid lines) reflect the inherent smoothness of phase-velocity curves.

## 7 DISCUSSION AND CONCLUSIONS

The increasing abundance of Rayleigh- and Love-wave phase-velocity measurements presents opportunities for the imaging of seismic anisotropy more accurately and in more locations than ever



**Figure 18.** Marginal posterior probability distributions for  $V_{sv}^0$ ,  $A_1$ ,  $A_2$ , Moho depth, amplitude of azimuthal anisotropy and the fast-propagation azimuth, resulting from the inversions of the azimuthally anisotropic data. Pink lines show the best-fitting profiles and Moho depth value obtained from a moderately damped azimuthally anisotropic non-linear gradient-search inversion. The marginal prior probability on the Moho depth is shown as a green line superimposed on the marginal posterior distribution of the sampled Moho depths (bottom left).

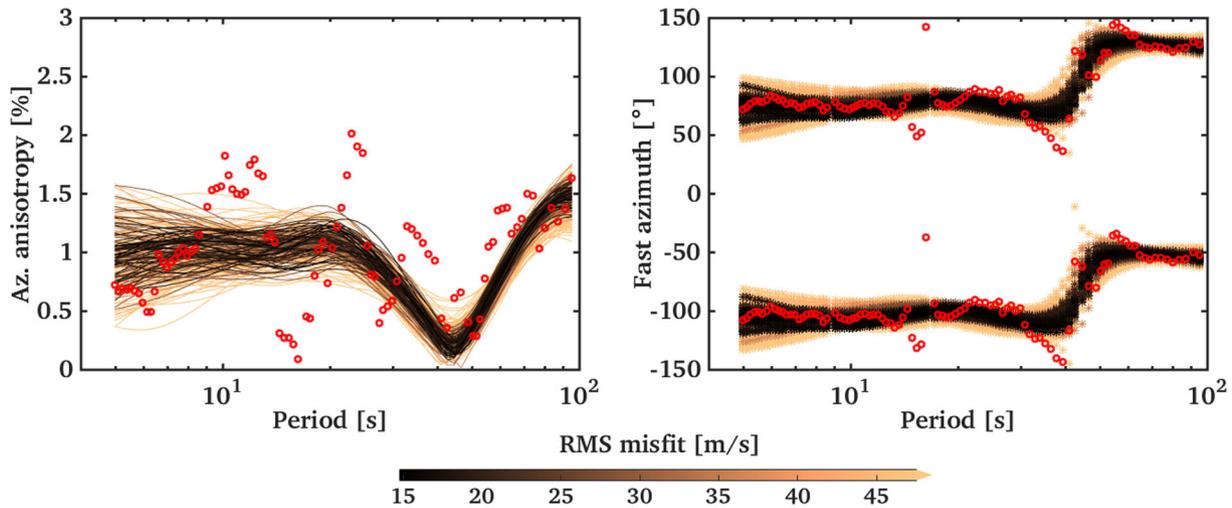


**Figure 19.** Left: marginal posterior probability distribution for  $V_{sv}^0$  (isotropic-average component of shear velocity) resulting from the inversions of the azimuthally anisotropic data. Right: the dependence of shear velocity on azimuth at different depths, retrieved by the McMC inversion.

before. These data come from measurements using station pairs, from array analysis, and from phase-velocity maps, regional or global. The 1-D, Bayesian Monte Carlo approach developed in this study is well suited for the inversion of these data. The methods can extract structural information from the data effectively and constrain

the depth-dependent radial and azimuthal shear-wave anisotropies in the crust and upper mantle.

Importantly, the probabilistic basis of the method enables us to quantify the non-uniqueness of solutions by drawing samples from the posterior distributions of Earth model parameters. Treating the



**Figure 20.** Distribution of Rayleigh-wave,  $2\phi$  anisotropy with period in the west-central Italy region of study. Red circles show the amplitude (half peak to peak) (left) and fast-propagation azimuth (right) of phase-velocity anisotropy in the region (Keogh *et al.* 2009; Lebedev *et al.* 2010). Solid and dotted lines show the amplitude (left) and fast-propagation direction (right) of anisotropy of the phase velocities corresponding to the best-fitting models yielded by our McMC inversions.

data noise as an unknown in the inversions for radial anisotropy, we can let the data dictate the closeness of the fit that is required, which shows very promising results. Accurate estimation of errors in phase-velocity data is often difficult, especially if the data themselves are results of inversions (phase-velocity maps, for example). It is thus useful to have the means to estimate the errors, for which we have implemented and tested two approaches. In one, we use gradient-search inversions to identify the component of the signal that cannot be fit by any 1-D Earth model, and in the other, we make the amplitude of noise an unknown in McMC inversions and solve for it.

Our Bayesian inversion for depth-dependent azimuthal anisotropy of shear-waves differs from previous methods in that it is fully non-linear. The relation between the model parameters and surface-wave phase velocities is not linearized around a model that is averaged azimuthally, with the phase velocities at different azimuths computed numerically for every model realization. This ensures that any linearization biases are avoided.

We also developed a faster, non-linear gradient-search inversion of azimuthally anisotropic phase-velocity curves for depth-dependent  $V_s$  anisotropy. The gradient-search and the McMC algorithms were tested with synthetic and measured data and yielded mutually consistent models; both are ready for applications with either small or large data sets. In particular, the methods can be applied for a point-by-point inversion of Rayleigh- and Love-wave phase-velocity maps for 3-D anisotropic models.

There remain aspects of the methods where refinements and improvements would be beneficial, at least in some applications. First, the parametrization that we have defined is not adaptive, therefore the number of model parameters is not treated as one of the unknowns during the inversions but is chosen beforehand. Although the locations of the control points in the mantle can be changed in each Markov chain in order to allow flexibility of the model space sampling, the number of layers and the depths of discontinuities in the crust (except for the Moho depth) are fixed in each inversion and set equal to some reference values. A possible improvement of the algorithms would be achieved by using an adaptive parametrization of the crustal structure, or by treating the crustal discontinuity depths as additional parameters of the inversions (which can be incorporated easily).

Secondly, although we try to estimate the dependency of the data noise with period and we treat the amplitude of the noise as an unknown in the radial anisotropy inversions, we assume uncorrelated measurement errors and diagonal covariance matrices of the noise. This may not be true for surface-wave dispersion data. One approach to take into account the correlation between neighbouring data points would be to introduce off-diagonal terms in the covariance matrix of data noise and, possibly, treat both the variance and correlation parameters of data noise as unknowns in the inversions (as done in Bodin *et al.* 2012 for receiver-function inversions).

Thirdly, our algorithms have been applied only to the fundamental-mode data. A future implementation could include a joint inversion of the fundamental and higher mode phase-velocity curves, provided that phase-velocity measurements for higher modes are available. This development would yield resolution of Earth structure down to the transition-zone and lower mantle depths (e.g. Khan *et al.* 2011).

The applications of the methods to measured data presented in this paper were not only used to test the methods, but also provided useful new information on Earth structure and dynamics. Our radially anisotropic profile for the Baikal-central Mongolia region yielded resolution from the crust down to the transition-zone depths. It is consistent with those from previous studies in its main features (Lebedev *et al.* 2006; Fullea *et al.* 2012). However, it also provides more convincing evidence than previously for a strong vertical component of flow in the asthenosphere. This is consistent with an upward asthenospheric flow from below the thick lithosphere of the Siberian Craton to below the thinner lithosphere of central Mongolia, which is likely to cause decompression melting of veins of enriched material within the asthenosphere and give rise to the scattered, sporadic volcanism observed in the Baikal Rift area.

The McMC inversion of the phase-velocity data from west-central Italy (Keogh *et al.* 2009; Lebedev *et al.* 2010) for azimuthal anisotropy revealed the layering of anisotropic fabric in the lithosphere–asthenosphere depth range. A clear change in the fast-propagation direction was detected in the 70–100 km depth range, probably at or just below the LAB. The orientation of the fabric in the lithosphere is roughly E–W, parallel to the direction of stretching over the last 10 m.y. (Jolivet *et al.* 2009). The orientation of the fabric we detect in the asthenosphere is NW–SE, matching the

fast directions inferred from shear-wave splitting (Salimbeni *et al.* 2007), which thus occurs in the asthenosphere.

## ACKNOWLEDGEMENTS

We are grateful to Sung-Joon Chang, Malcolm Sambridge and the editor, Ana Ferreira, for constructive comments and suggestions that helped us improve the manuscript. We thank Nicola Piana-Agostinetti and Pierre Arroucau for insightful discussions on Bayesian Inference and Andrew Schaeffer, Andrea Licciardi and Eric Mandolesi for useful feedback. This work was funded by the Dublin Institute for Advanced Studies and by Science Foundation Ireland (grant 13/CDA/2192). The authors wish to acknowledge the DJEI/DES/SFI/HEA Irish Centre for High-End-Computing for the computational facilities and support. Fig. 10 and the map in Fig. 15 were prepared using the Generic Mapping Tool of Wessel & Smith (1998).

## REFERENCES

- Adam, J.M.-C., 2013. Seismic structure and anisotropy in southern Africa's lithosphere: constraints from broad-band surface-wave dispersion, *PhD thesis*, Dublin Institute for Advanced Studies, Dublin.
- Adam, J.M.-C. & Lebedev, S., 2012. Azimuthal anisotropy beneath southern Africa from very broad-band surface-wave dispersion measurements, *Geophys. J. Int.*, **191**(1), 155–174.
- Agius, M.R. & Lebedev, S., 2013. Tibetan and Indian lithospheres in the upper mantle beneath Tibet: evidence from broadband surface-wave dispersion, *Geochem. Geophys. Geosyst.*, **14**(10), 4260–4281.
- Agius, M.R. & Lebedev, S., 2014. Shear-velocity structure, radial anisotropy and dynamics of the Tibetan crust, *Geophys. J. Int.*, **199**(3), 1395–1415.
- Agius, M.R. & Lebedev, S., 2017. Complex, multilayered azimuthal anisotropy beneath Tibet: evidence for co-existing channel flow and pure-shear crustal thickening, *Geophys. J. Int.*, **210**(3), 1823–1844.
- Anderson, D.L., 1961. Elastic wave propagation in layered anisotropic media, *J. geophys. Res.*, **66**(9), 2953–2963.
- Anderson, D.L., 1967. Latest information from seismic observations, in *The Earth's Mantle*, pp. 355–420, ed. Gaskell, T.F., Academic Press, New York.
- Auer, L., Boschi, L., Becker, T.W., Nissen-Meyer, T. & Giardini, D., 2014. Savani: a variable resolution whole-mantle model of anisotropic shear velocity variations based on multiple datasets, *J. geophys. Res.*, **119**(4), 3006–3034.
- Babuska, V. & Cara, M., 1991. *Seismic Anisotropy in the Earth*, Kluwer Academic Publishers.
- Backus, G.E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.*, **67**(11), 4427–4440.
- Bartzsch, S., Lebedev, S. & Meier, T., 2011. Resolving the lithosphere-asthenosphere boundary with seismic Rayleigh waves, *Geophys. J. Int.*, **186**(3), 1152–1164.
- Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface-wave tomography in North America, *EOS, Trans. Am. geophys. Un.*, **81**, F897.
- Bayes, T., 1763. An essay towards solving a problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S., *Phil. Trans. R. Soc. Lond.*, **53**, 370–418. [Reprinted in *Biometrika*, **45**, 295–315, 1958.]
- Becker, T.W., Lebedev, S. & Long, M.D., 2012. On the relationship between azimuthal anisotropy from shear-wave splitting and surface-wave tomography, *J. geophys. Res.*, **117**(B1), 1–17.
- 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., Schaeffer, A.J., Lebedev, S. & Conrad, C.P., 2015. Toward a generalized plate motion reference frame, *Geophys. Res. Lett.*, **42**(9), 3188–3196.
- Beghein, C. & Trampert, J., 2004. Probability density functions for radial anisotropy: implications for the upper 1200 km of the mantle, *Earth planet. Sci. Lett.*, **217**(1), 151–162.
- Beghein, C., Trampert, J. & van Heijst, H., 2006. Radial anisotropy in seismic reference models of the mantle, *J. geophys. Res.*, **111**, B02303, doi:10.1029/2005JB003728.
- Beghein, C., Yuan, K., Schmerr, N. & Xing, Z., 2014. Changes in seismic anisotropy shed light on the nature of the Gutenberg discontinuity, *Science*, **343**(6176), 1237–1240.
- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K. & Rawlinson, N., 2012. Transdimensional inversion of receiver functions and surface-wave dispersion, *J. geophys. Res.*, **117**, B02301, doi:10.1029/2011JB008560.
- Bodin, T., Capdeville, Y., Romanowicz, B. & Montagner, J.-P., 2015. Interpreting radial anisotropy in global and regional tomographic models, in *The Earth's Heterogeneous Mantle*, pp. 105–143, eds Khan, A., Deschamps, F. & Kawai, K., Springer International Publishing.
- Bodin, T., Leiva, J., Romanowicz, B., Maupin, V. & Yuan, H., 2016. Imaging anisotropic layering with Bayesian inversion of multiple data types, *Geophys. J. Int.*, **206**(1), 605–629.
- Burgos, G., Montagner, J.-P., Beucler, E., Capdeville, Y., Mocquet, A. & Drilleau, M., 2014. Oceanic lithosphere-asthenosphere boundary from surface-wave dispersion data, *J. geophys. Res.*, **119**(2), 1079–1093.
- Calò, M., Bodin, T. & Romanowicz, B., 2016. Layered structure in the upper mantle across North America from joint inversion of long and short period seismic data, *Earth planet. Sci. Lett.*, **449**, 164–175.
- Chang, S.-J., Ferreira, A.M.G., Ritsema, J., Van Heijst, H.J. & Woodhouse, J.H., 2014. Global radially anisotropic mantle structure from multiple datasets: a review, current challenges, and outlook, *Tectonophysics*, **617**, 1–19.
- Chang, S.J. & Ferreira, A.M.G., 2017. Improving global radial anisotropy tomography: the importance of simultaneously inverting for crustal and mantle structure, *Bull. seism. Soc. Am.*, **107**(2), 624–638.
- Chang, S.-J., Ferreira, A.M.G., Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., 2015. Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations, *J. geophys. Res.*, **120**, doi:10.1002/2014JB011824.
- Crampin, S., 1984. Effective anisotropic elastic constants for wave propagation through cracked solids, *Geophys. J. R. astr. Soc.*, **76**, 135–145.
- Darbyshire, F.A., Eaton, D.W. & Bastow, I.D., 2013. Seismic imaging of the lithosphere beneath Hudson Bay: episodic growth of the Laurentian mantle keel, *Earth planet. Sci. Lett.*, **373**, 179–193.
- Debayle, E. & Ricard, Y., 2012. A global shear velocity model of the upper mantle from fundamental and higher Rayleigh mode measurements, *J. geophys. Res.*, **117**(B10), 1–24.
- Debayle, E. & Kennett, B.L.N., 2000. Anisotropy in the Australasian upper mantle from Love and Rayleigh waveform inversion, *Earth planet. Sci. Lett.*, **184**, 339–351.
- Debayle, E., Kennett, B. & Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia, *Nature*, **433**(7025), 509–512.
- Debayle, E., Dubuffet, F. & Durand, S., 2016. An automatically updated S-wave model of the upper mantle and the depth extend of azimuthal anisotropy, *Geophys. Res. Lett.*, **43**(2), 674–682.
- Denison, D.G.T., Holmes, C.C., Mallik, B.K. & Smith, A.F.M., 2002. *Bayesian Methods for Nonlinear Classification and Regression*. Wiley, Chichester.
- Deschamps, F., Lebedev, S., Meier, T. & Trampert, J., 2008a. Azimuthal anisotropy of Rayleigh-wave phase velocities in the East-central United States, *Geophys. J. Int.*, **173**, 827–843.
- Deschamps, F., Lebedev, S., Meier, T. & Trampert, J., 2008b. Stratified seismic anisotropy reveals past and present deformation beneath the East-central United States, *Earth planet. Sci. Lett.*, **274**(3), 489–498.
- Drilleau, M., Beucler, E., Mocquet, A., Verhoeven, O., Moebs, G., Burgos, G., Montagner, J.-P. & Vacher, P., 2013. A Bayesian approach to infer

- radial models of temperature and anisotropy in the transition zone from surface-wave data, *Geophys. J. Int.*, **195**, 1165–1183.
- Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, **25**, 297–356.
- Ekstrom, G. & Dziewonski, A.M., 1998. The unique anisotropy of the Pacific upper mantle, *Nature*, **394**, 168–172.
- Ekstrom, G., Tromp, J. & Larson, E.W.F., 1997. Measurements and global models of surface-wave propagation, *J. geophys. Res.*, **102**(B4), 8137–8157.
- Endrun, B., Meier, T., Lebedev, S., Bohnhoff, M., Stavrakakis, G. & Harjes, H.P., 2008. S velocity structure and radial anisotropy in the Aegean region from surface-wave dispersion, *Geophys. J. Int.*, **174**(2), 593–616.
- Erduran, M., Endrun, B. & Meier, T., 2008. Continental vs. oceanic lithosphere beneath the eastern Mediterranean Sea—implications from Rayleigh wave dispersion measurements, *Tectonophysics*, **457**, 42–52.
- Ferreira, A.M.G., Woodhouse, J.H., Visser, K. & Trampert, J., 2010. On the robustness of global radially anisotropic surface-wave tomography, *J. geophys. Res.*, **115**, B04313, doi:10.1029/2009JB006716.
- 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), 270–280.
- French, S., Lekic, V. & Romanowicz, B., 2013. Waveform tomography reveals channeled flow at the base of the oceanic asthenosphere, *Science*, **342**(6155), 227–230.
- Fry, B., Davey, F., Eberhart-Phillips, D. & Lebedev, S., 2014. Depth variable crustal anisotropy, patterns of crustal weakness, and destructive earthquakes in Canterbury, New Zealand, *Earth planet. Sci. Lett.*, **392**, 50–57.
- Fuller, J., Lebedev, S., Agius, M.R., Jones, A.G. & Afonso, J.C., 2012. Lithospheric structure in the Baikal–central Mongolia region from integrated geophysical-petrological inversion of surface-wave data and topographic elevation, *Geochem. Geophys. Geosyst.*, **13**, Q0AK09, doi:10.1029/2012GC004138.
- Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M. & Stephenson, J., 2009. Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for Earth Science problems, *Mar. Petrol. Geol.*, **26**(4), 525–535.
- Gelman, A., Carlin, J., Stern, H. & Rubin, D., 2004. *Bayesian Data Analysis*, CRC Press.
- Gilks, W.R., Richardson, S. & Spiegelhalter, D., 1996. *Markov chain Monte Carlo in Practice*, Chapman & Hall.
- Gung, Y., Panning, M. & Romanowicz, B., 2003. Global anisotropy and the thickness of continents, *Nature*, **422**(6933), 707–711.
- Harkrider, D.G. & Anderson, D.L., 1962. Computation of surface-wave dispersion for multilayered anisotropic media, *Bull. seism. Soc. Am.*, **52**, 321–332.
- Hastings, W., 1970. Monte Carlo simulation methods using Markov chains and their applications, *Biometrika*, **57**, 97–109.
- Huang, H., Yao, H. & van der Hilst, R.D., 2010. Radial anisotropy in the crust of SE Tibet and SW China from ambient noise interferometry, *Geophys. Res. Lett.*, **37**, L21310, doi:10.1029/2010GL044981.
- Jolivet, L., Faccenna, C. & Piromallo, C., 2009. From mantle to crust: stretching the Mediterranean, *Earth planet. Sci. Lett.*, **285**(1–2), 198–209.
- Kawakatsu, H., Kumar, P., Takei, Y., Shinohara, M., Kanazawa, T., Araki, E. & Suyehiro, K., 2009. Seismic evidence for sharp lithosphere-asthenosphere boundaries of oceanic plates, *Science*, **24**, 499–502.
- Kennett, B.L.N., Engdahl, E.R. & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, **122**(1), 108–124.
- Keogh, P., Adam, J. & Lebedev, S., 2009. A surface-wave study of structure and anisotropy of Tuscany, in *AGU Fall Meeting Abstracts*, abstract #T51B-1524.
- Khan, A., Boschi, L. & Connolly, J.A.D., 2011. Mapping the Earth's thermochemical and anisotropic structure using global surface-wave data, *J. geophys. Res.*, **116**, B01301, doi:10.1029/2010JB007828.
- Lebedev, S., Nolet, G., Meier, T. & van der Hilst, R.D., 2005. Automated multimode inversion of surface and S waveforms, *Geophys. J. Int.*, **162**, 951–964.
- Lebedev, S. & van der Hilst, R.D., 2008. Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms, *Geophys. J. Int.*, **173**, 505–518.
- Lebedev, S., Meier, T. & van der Hilst, R.D., 2006. Asthenospheric flow and origin of volcanism in the Baikal Rift area, *Earth Planet. Sci. Lett.*, **249**, 415–424.
- Lebedev, S., Adam, J., Keogh, P. & Tirel, C., 2010. Structure and subduction-induced deformation of Tuscany's lithosphere, from broadband surface-wave analysis, in *Abstract T/MO/O2, ESC General Assembly 2010*, Montpellier, France, 2010 September 6–10.
- Lebedev, S., Adam, J.M.-C. & Meier, T., 2013. Mapping the Moho with seismic surface waves: a review, resolution analysis, and recommended inversion strategies, *Tectonophysics*, **609**, 377–394.
- Lekic, V. & Romanowicz, B., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method, *Geophys. J. Int.*, **185**(2), 799–831.
- Lévesque, J., Debayle, E. & Maupin, V., 1998. Anisotropy in the Indian Ocean upper mantle from Rayleigh and Love waveform inversion, *Geophys. J. Int.*, **133**, 529–540.
- Levshin, A. & Ratnikova, L., 1984. Apparent anisotropy in inhomogeneous media, *Geophys. J. R. astr. Soc.*, **76**(1), 65–69.
- Lin, F.C., Ritzwoller, M.H., Yang, Y., Moschetti, M.P. & Fouch, M.J., 2011. Complex and variable crustal and uppermost mantle seismic anisotropy in the western United States, *Nature Geosci.*, **4**(1), 55–61.
- Long, M.D. & Becker, T.W., 2010. Mantle dynamics and seismic anisotropy, *Earth planet. Sci. Lett.*, **297**(3–4), 341–354.
- Love, A.E.H., 1927. *A Treatise on the Mathematical Theory of Elasticity*, Cambridge Univ. Press.
- Malinverno, A., 2002. Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, *Geophys. J. Int.*, **151**(3), 675–688.
- Malinverno, A. & Briggs, V.A., 2004. Expanded uncertainty quantification in inverse problems: hierarchical Bayes and empirical Bayes, *Geophysics*, **69**(4), 1005–1016.
- Margheriti, L. *et al.*, 2006. RETREAT seismic deployment in the Northern Apennines, *Ann. Geophys.*, **49**(4/5), 1005–1017.
- Masters, G., Barmine, M.P. & Kientz, S., 2007. Mineos user's manual, in *Computational Infrastructure for Geodynamics*, California Institute of Technology, Pasadena.
- Maupin, V., 1985. Partial derivatives of surface-wave phase velocities for anisotropic models, *Geophys. J. Int.*, **83**(2), 379–398.
- McEvilly, T.V., 1964. Central US crust-upper mantle structure from Love and Rayleigh wave phase velocity inversion, *Bull. seism. Soc. Am.*, **54**, 1997–2015.
- Meier, T., Dietrich, K., Stockhert, B. & Harjes, H.-P., 2004. One-dimensional models of shear-wave velocity for the eastern Mediterranean obtained from the inversion of Rayleigh wave phase velocities and tectonic implications, *Geophys. J. Int.*, **156**(1), 45–58.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E., 1953. Equations of state calculations by fast computational machine, *J. Chem. Phys.*, **21**(6), 1087–1091.
- Molinari, I., Verbeke, J., Boschi, L., Kissling, E. & Morelli, A., 2015. Italian and Alpine three-dimensional crustal structure imaged by ambient-noise surface-wave dispersion, *Geochem. Geophys. Geosyst.*, **16**(12), 4405–4421.
- 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. & Anderson, D.L., 1989. Petrological constraints on seismic anisotropy, *Phys. Earth planet. Inter.*, **54**, 82–105.
- Montagner, J.-P. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. geophys. Res.*, **95**, 4797–4819.
- Montagner, J.-P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. geophys. Res.*, **96**, 20 337–20 351.

- Morgan, J.P. & Morgan, W.J., 1999. Two-stage melting and the geochemical evolution of the mantle: a recipe for mantle plum-pudding, *Earth planet. Sci. Lett.*, **170**, 215–239.
- Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.*, **100**(B7), 12–431.
- Mosegaard, K., Singh, S.C., Snyder, D. & Wagner, H., 1997. Monte Carlo analysis of seismic reflections from Moho and the W-reflector, *J. geophys. Res.*, **102**(B2), 2969–2981.
- Nettles, M. & Dziewonski, A.M., 2008. Radially anisotropic shear velocity structure of the upper mantle globally and beneath North America, *J. geophys. Res.*, **113**(B2), 1–27.
- Nolet, G., 1990. Partitioned wave-form inversion and 2D structure under the NARS array, *J. geophys. Res.*, **95**, 8513–8526.
- Nolet, G., 2008. *A Breviary of Seismic Tomography*. Cambridge University Press.
- Nur, A., 1971. Effects of stress on velocity anisotropy in rocks with cracks, *J. geophys. Res.*, **76**(8), 2022–2034.
- Panning, M.P. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.*, **167**(1), 361–379.
- Pedersen, H.A., Bruneton, M. & Maupin, V. & SVEKALAPKO Seismic Tomography Working Group, 2006. Lithospheric and sublithospheric anisotropy beneath the Baltic shield from surface-wave array analysis, *Earth planet. Sci. Lett.*, **244**(3), 590–605.
- Piana Agostinetti, N. & Malinverno, A., 2010. Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophys. J. Int.*, **181**(2), 858–872.
- Plomerová, J. *et al.*, 2006. Seismic anisotropy beneath the Northern Apennines (Italy): mantle flow or lithosphere fabric?, *Earth planet. Sci. Lett.*, **247**(1), 157–170.
- Polat, G., Lebedev, S., Readman, P.W., O'Reilly, B.M. & Hauser, F., 2012. Anisotropic Rayleigh-wave tomography of Ireland's crust: implications for crustal accretion and evolution within the Caledonian Orogen, *Geophys. Res. Lett.*, **39**(4), 1–5.
- Ribe, N.M., 1989. Seismic anisotropy and mantle flow, *J. geophys. Res.*, **94**, 4213–4223.
- Ritzwoller, M.H., Shapiro, N.M., Barmin, M.P. & Levshin, A.L., 2002. Global surface-wave diffraction tomography, *J. geophys. Res.*, **107**(B12), 2335, doi:10.1029/2002JB001777.
- Roux, E., Moorkamp, M., Jones, A.G., Bischoff, M., Endrun, B., Lebedev, S. & Meier, T., 2011. Joint inversion of long-period magnetotelluric data and surface-wave dispersion curves for anisotropic structure: application to data from Central Germany, *Geophys. Res. Lett.*, **38**, L05304, doi:10.1029/2010GL046358.
- Salimbeni, S., Pondrelli, S., Margheriti, L., Levin, V., Park, J., Plomerova, J. & Babuska, V., 2007. Abrupt change in mantle fabric across northern Apennines detected using seismic anisotropy, *Geophys. Res. Lett.*, **34**, L07308, doi:10.1029/2007GL029302.
- Sambridge, M., Gallagher, K., Jackson, A. & Rickwood, P., 2006. Trans-dimensional inverse problems, model comparison and the evidence, *Geophys. J. Int.*, **167**(2), 528–542.
- Shapiro, N.M. & Ritzwoller, M.H., 2002. Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int.*, **151**, 88–105.
- Silver, P.G. & Chan, W.W., 1988. Implications for continental structure and evolution from seismic anisotropy, *Nature*, **335**, 34–39.
- Schaeffer, A.J. & Lebedev, S., 2013. Global shear-speed structure of the upper mantle and transition zone, *Geophys. J. Int.*, **194**, 417–449.
- Schaeffer, A.J., Lebedev, S. & Becker, T.W., 2016. Azimuthal seismic anisotropy in the Earth's upper mantle and the thickness of tectonic plates, *Geophys. J. Int.*, **207**(2), 901–933.
- Simons, F.J., van Der Hilst, R.D., Montagner, J.-P. & Zielhuis, A., 2002. Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle, *Geophys. J. Int.*, **151**(3), 738–754.
- Smith, M.L. & Dahlen, F., 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, *J. geophys. Res.*, **78**(17), 3321–3333.
- Snieder & R., 1988. Large-scale waveform inversions of surface waves for lateral heterogeneity, 2. Application to surface waves in Europe and the Mediterranean, *J. geophys. Res.*, **93**, 12067–12080.
- Soudou, F., Yuan, X., Kind, R., Lebedev, S., Adam, J., Kastle, E. & Tilmann, E., 2013. Seismic evidence for stratification in composition and anisotropic fabric within the thick lithosphere of Kalahari Craton, *Geochem. Geophys. Geosyst.*, **14**, 5393–5412.
- Takeuchi, H. & Saito, M., 1972. Seismic surface waves, *Methods Comput. Phys.*, **11**, 217–295.
- Tatham, D., Lloyd, G., Butler, R. & Casey, M., 2008. Amphibole and lower crustal seismic properties, *Earth planet. Sci. Lett.*, **267**(1–2), 118–128.
- Trampert, J., Woodhouse & J.H., 1996. High resolution phase velocity maps for the whole Earth, *Geophys. Res. Lett.*, **23**, 21–24.
- Visser, K., Trampert, J., Lebedev, S. & Kennett, B.L.N., 2008. Probability of radial anisotropy in the deep mantle, *Earth planet. Sci. Lett.*, **270**, 241–250.
- Wang, N., Montagner, J.-P., Fichtner, A. & Capdeville, Y., 2013. Intrinsic versus extrinsic seismic anisotropy: the radial anisotropy in reference earth models, *Geophys. Res. Lett.*, **40**(16), 4284–4288.
- Weiss, T., Siegesmund, S., Rabbel, W., Bohlen, T. & Pohl, M., 1999. Seismic velocities and anisotropy of the lower continental crust: a review, *Pure appl. Geophys.*, **156**(1–2), 97–122.
- Wessel, P. & Smith, W.H.F., 1998. New, improved version of the generic mapping tools released, *EOS, Trans. Am. geophys. Un.*, **79**, 579.
- Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle; three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. geophys. Res.*, **89**(B7), 5953–5986.
- Xie, J., Ritzwoller, M.H., Shen, W., Yang, Y., Zheng, Y. & Zhou, L., 2013. Crustal radial anisotropy across Eastern Tibet and the Western Yangtze Craton, *J. geophys. Res.*, **118**(8), 4226–4252.
- Xie, J., Ritzwoller, M.H., Brownlee, S. & Hacker, B., 2015. Inferring the oriented elastic tensor from surface-wave observations: preliminary application across the western United States, *Geophys. J. Int.*, **201**(2), 996–1019.
- Yao, H., van der Hilst, R.D. & Montagner, J.-P., 2010. Heterogeneity and anisotropy of the lithosphere of SE Tibet from surface-wave array tomography, *J. geophys. Res.*, **115**, B12307, doi:10.1029/2009JB007142.
- Yuan, H. & Romanowicz, B., 2010. Lithospheric layering in the North American craton, *Nature*, **466**, 1063–1068.
- Yuan, K. & Beghein, C., 2013. Seismic anisotropy changes across upper mantle phase transitions, *Earth planet. Sci. Lett.*, **374**, 132–144.
- Yuan, K. & Beghein, C., 2014. Three-dimensional variations in Love and Rayleigh wave azimuthal anisotropy for the upper 800 km of the mantle, *J. geophys. Res.*, **119**(4), 3232–3255.
- Zhang, S. & Karato, S.-I., 1995. Lattice preferred orientation of olivine aggregates deformed in simple shear, *Nature*, **375**, 774–777.
- Zhang, X., Paulssen, H., Lebedev, S. & Meier, T., 2009. 3D shear velocity structure beneath the Gulf of California from Rayleigh wave dispersion, *Earth planet. Sci. Lett.*, **279**(3–4), 255–262.
- Zhu, H. & Tromp, J., 2013. Mapping tectonic deformation in the crust and upper mantle beneath Europe and the North Atlantic Ocean, *Science*, **341**(6148), 871–875.
- Zhu, H., Bozdag, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nature Geosci.*, **5**, 493–498.

## SUPPORTING INFORMATION

Supplementary data are available at [GJI](#) online.

**SUPPLEMENT A:** THE PRIOR

**SUPPLEMENT B:** PROPOSAL DISTRIBUTIONS

**SUPPLEMENT C:** THE ACCEPTANCE PROBABILITY

**SUPPLEMENT D:** EFFICIENCY AND CONVERGENCE OF THE ALGORITHMS

**Figure D1.** Left panel: negative log-likelihood as a function of chain step, for the four chains run in the synthetic inversions described in section 5.1. Right panels: hyperparameter values, as functions of chain step, sampled during the inversions. The red lines indicate the true values of the two hyperparameters.

**Figure D2.** Marginal posterior probability distributions for isotropic average shear velocity ( $V_s$ ) and radial anisotropy at different depths, resulting from the synthetic inversions described in section 5.1. The Gaussian prior probability distributions (normal fits of the sampled prior distributions) are shown in green, and are centered, at each depth, at the reference model (dashed green lines). The values of the true synthetic model at the given depths are indicated by the red lines.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.