

# Geodynamic Tomography: Constraining Upper Mantle Deformation Patterns from Bayesian Inversion of Surface Waves

J.K. Magali<sup>1</sup>, T. Bodin<sup>1</sup>, N. Hedjazian<sup>1</sup>, H. Samuel<sup>2</sup>, and S. Atkins<sup>3</sup>

<sup>1</sup> *Université de Lyon, UCBL, CNRS, LGL-TPE, 69622 Villeurbanne, France. Email:*

*john-keith.magali@univ-lyon1.fr*

<sup>2</sup> *Université de Paris, Institut de Physique du Globe de Paris, CNRS, F-75005 Paris, France.*

<sup>3</sup> *Laboratoire de Géologie, Ecole Normale Supérieure, PSL Res. Univ, Paris, France.*

## SUMMARY

In the Earth's upper mantle, seismic anisotropy mainly originates from the crystallographic preferred orientation (CPO) of olivine due to mantle deformation. Large-scale observation of anisotropy in surface wave tomography models provides unique constraints on present-day mantle flow. However, surface waves are not sensitive to the 21 coefficients of the elastic tensor, and therefore the complete anisotropic tensor cannot be resolved independently at every location. This large number of parameters may be reduced by imposing spatial smoothness and symmetry constraints to the elastic tensor. In this work, we propose to regularize the tomographic problem by using constraints from geodynamic modeling to reduce the number of model parameters. Instead of inverting for seismic velocities, we parameterize our inverse problem directly in terms of physical quantities governing mantle flow: a temperature field, and a temperature-dependent viscosity. The forward problem consists of three steps: (1) calculation of mantle flow induced by thermal anomalies, (2) calculation of the induced CPO and elastic properties using a micro-mechanical

model, and (3) computation of azimuthally-varying surface wave dispersion curves. We demonstrate how a fully non-linear Bayesian inversion of surface wave dispersion curves can retrieve the temperature and viscosity fields, without having to explicitly parameterize the elastic tensor. Here, we consider simple flow models generated by spherical temperature anomalies. The results show that incorporating geodynamic constraints in surface wave inversion help to retrieve patterns of mantle deformation. The solution to our inversion problem is an ensemble of models (*i.e.*, thermal structures) representing a posterior probability, therefore providing uncertainties for each model parameter.

**Key words:** Seismic tomography; Seismic anisotropy, Surface wave and free oscillations; Inverse theory; Probability distributions.

## 1 INTRODUCTION

Seismic anisotropy reveals key insights into the Earth’s interior structure and dynamics. In the upper mantle, the propagation of seismic waves appears to be anisotropic, which has generally been associated with the preferred alignment of mantle minerals (Nicolas & Christensen 1987; Montagner 1994). This so-called intrinsic anisotropy relates to the strain history induced by regional-scale convection and is observable with various seismological tools, including surface waves.

### 1.1 Surface wave tomography studies

Surface wave tomography offers a powerful technique to constrain seismic anisotropy and to image the structure of the upper mantle at both regional and global scales. With growing amounts of seismic data, tomographers have produced detailed models of azimuthal anisotropy (e.g., Debayle et al. 2005; Deschamps et al. 2008; Adam & Lebedev 2012; Yuan & Beghein 2013, 2014), and radial anisotropy (e.g., Plomerová et al. 2002; Lebedev et al. 2006; Nettles & Dziewoński 2008; Chang et al. 2014, 2015). Numerous studies have inverted dispersion curves by minimizing the difference between observed and synthetic phase and/or group velocities, proving that they can effectively constrain the depth dependence of anisotropy (e.g., Montagner & Tanimoto 1990; Ritzwoller et al. 2002).

Seismic anisotropy can be described with 21 independent components of the elastic tensor.

In practice however, the full tensor cannot be resolved by the seismic data independently at every location, and generally only a restricted number of parameters are inverted for. This is done by assuming specific symmetry classes, or by using petrological constraints to impose relations between some of the parameters. Surface waves in particular are only sensitive to 13 parameters that are just a linear combination of the elastic constants (Montagner & Nataf 1986). General practices in surface wave tomography thus investigate: (1) radial anisotropy (assuming vertical transverse isotropy, VTI, where the axis of hexagonal symmetry is vertical), constrained by comparing the speed of Rayleigh waves with that of Love waves, also known as the Rayleigh-Love discrepancy (Babuska & Cara 1991); or (2) azimuthal anisotropy, which deals with first-order variations of velocities as function of the azimuth of propagation. For example, azimuthal anisotropy can be inferred from the azimuthal terms of the Rayleigh wave phase velocities (Smith & Dahlen 1973).

Simultaneous interpretations of radial and azimuthal anisotropy have been the subject of extensive research (e.g. Beghein et al. 2014; Burgos et al. 2014). Joint efforts involving the use of *a priori* information have already been conducted to reduce the high dimensionality of anisotropic inversion. Montagner & Anderson (1989) showed that correlations exist between the elastic constants derived from petrological models, thereby reducing the total number of free parameters to be inverted for. This motivated the development of "vectorial tomography" where it involves inverting for 7 parameters instead of 13: two angles defining the strike and dip of the symmetry axis, three coefficients defining the strength of anisotropy, and finally two isotropic coefficients (Montagner & Nataf 1988; Montagner & Jobert 1988). Such a medium is also known as tilted transverse isotropy (TTI) and describes the 3D distributions of anisotropy. This further led to studies revealing that deformation-induced anisotropy can be described by a TTI medium where correlations appear to exist between P and S wave anisotropy (Becker et al. 2006). Such correlations can then be exploited to further simplify anisotropic inversion. Panning & Nolet (2008) then laid the groundwork to derive finite-frequency kernels of surface waves that are explicitly based on a TTI medium. In practice however, constraining the tilt may still be difficult due to sparse azimuthal sampling, alongside other competing factors such as non-uniqueness of the solution and poor data quality. Even so, simultaneous inversions for radial and azimuthal anisotropy using TTI models have already been applied at the regional scale using probabilistic approaches to combat these shortcomings (Xie et al. 2015, 2017).

Surface wave tomography is an ill-posed inverse problem. This arises from the uneven distribution of sources and receivers causing limited ray path coverage, and from noise in the observed seismograms. The type of spatial parameterization may also lead to ambiguity when

interpreting tomographic results. A conventional technique is to separate the problem into two steps. The first step is to construct velocity maps for each considered period, which is an almost linear inverse problem. It is followed by an inversion of each local dispersion curve to build a model of elastic parameters. The inversion is in general performed using a linearized technique, which favors a stable and unique solution through regularization, for example by adding a spatial smoothness constraint on the model parameters.

More recently, the development of probabilistic approaches using direct sampling of the model space makes it possible to handle the non-uniqueness of the solution and estimate uncertainties on the inferred parameters. These methods require the evaluation of the forward model a large number of times, and hence have a high computational cost. Nevertheless, numerous works have been successful in applying such inversion schemes to seismic data and in particular to the inversion of surface waves dispersion curves (Shapiro & Ritzwoller 2002; Shen et al. 2012; Bodin et al. 2016; Ravenna & Lebedev 2017; Xu & Beghein 2019).

In this study, we propose a complementary approach to estimate the full elastic tensor. This involves the incorporation of geodynamic and mineral physics modeling constraints: the textural evolution of peridotite aggregates during their deformation in the convective mantle. We propose a method to invert directly for the temperature field that produces convective flow and texture evolution. Modeling intrinsic anisotropy in this way removes the issue of low sensitivity from seismic waves since the elastic tensor is not explicitly inverted for, but instead computed directly from texture evolution models. Additionally, the inversion is performed using a Bayesian sampling algorithm, hence provide uncertainties on the obtained temperature field.

## 1.2 Deformation-induced seismic anisotropy

In the upper mantle, the existence of large-scale anisotropy appears to be ubiquitous in regions associated with strong deformation (McKenzie 1979). Its interpretation is based on the development of crystallographic preferred orientation (CPO) in olivine aggregates during their plastic deformation (Nicolas & Christensen 1987). Due to the physical process at its origin, seismic anisotropy can be interpreted in terms of the strain history associated with upper mantle circulation.

Different proxies have then been utilized to interpret seismic anisotropy directly in terms of mantle flow. First-order seismic observations suggest that the fast axis of azimuthal anisotropy tends to align with horizontal mantle flow (Ribe 1989; Becker et al. 2003, 2014). However, this behavior may not always be exhibited due to complex local deformation mechanisms associ-

ated with CPO evolution. Moreover, it is also important to emphasize that the development of anisotropy relates to the history of velocity gradients along a flow line, and not to the velocity field itself. Laboratory experiments of simple shear suggest that, at low strains, the orientation of the olivine fast axis tends to be aligned with the long axis of the finite strain ellipsoid (FSE) (Zhang & Karato 1995; Ribe 1992). The amplitude of anisotropy, on the other hand, can be approximated as a monotonic function relating to the ratio between the long axis and the short axis of the FSE (Ribe 1992; Hedjazian & Kaminski 2014). At sufficiently large strains however, CPO evolution deviates from the FSE due to the apparition of dynamic recrystallization. It tends to align nearly parallel to the direction of shear instead (Zhang & Karato 1995; Bystricky et al. 2000), although its transient behavior remains complex (Hansen et al. 2014a). Following this observation, a possible proxy is to interpret the orientation of the anisotropy fast axis as the infinite strain axis (ISA), that is, the axis of the FSE in the limit of infinite strains (Kaminski & Ribe 2002). In practice however, this proxy have had limited success at the global scale (Becker et al. 2014).

For that reason, an adequate interpretation of seismic anisotropy is usually based on numerical models of texture evolution. They require geodynamic flow models as inputs to provide the complete strain history. However, in some problems, the complete flow trajectory is unknown or too costly to compute, and only present-day flow is available. In this case, we propose to use a steady-state assumption to reconstruct the deformation history. This approximation is acceptable provided that the time scale of texture evolution is much smaller than that of the flow fluctuations (Kaminski & Ribe 2002).

### 1.3 Interpreting tomographic images with geodynamic modeling

In order to explain surface wave anisotropy, particularly in intra-oceanic and young continental regions, first-order interpretations involve finite strains computed from global circulation models (Becker et al. 2003). In their work, the density field derived from isotropic tomography (Becker & Boschi 2002) is used to compute instantaneous flow solutions in the upper mantle. Finite strain models derived from the flow are subsequently compared with azimuthal anisotropy in surface waves. However, as discussed above, finite strain-derived models may fall short at larger strains due to dynamic recrystallization (Zhang & Karato 1995). This urges the use of computational strategies that incorporate texture evolution models to estimate the level of CPO anisotropy.

Texture evolution can be modeled using micro-mechanical models of visco-plastic deformation of upper mantle minerals (Tommasi et al. 2000). One of which in particular uses

a kinematic formalism to model texture evolution of olivine aggregates by plastic deformation and dynamic recrystallization (Kaminski et al. 2004). It has been extensively applied to predict CPO-induced anisotropy from geodynamic flow models in a forward modeling approach at the regional (Hall et al. 2000; Lassak et al. 2006; Miller & Becker 2012; Faccenda & Capitanio 2013) and at the global scale (Becker et al. 2006, 2008). Forward models such as this assist further in the interpretation of seismic tomography models in terms of mantle circulation patterns. To cite an example, CPO-induced anisotropy resulting from 3D numerical simulations of subducting slabs shows consistency with radial anisotropy patterns inferred from global tomographic images (Ferreira et al. 2019; Sturgeon et al. 2019). However, most studies rely on visual comparisons between CPO obtained from numerical simulations and tomographic images. To the best of our knowledge, no study yet exists where mantle deformation has been inferred directly from seismic observations using an inverse approach.

#### 1.4 Geodynamic tomography

This motivated us to implement geodynamic tomography, an approach where no symmetry is imposed to the elastic tensor at the outset, and where seismic observations are inverted with constraints from geodynamic modeling, in a fully Bayesian parameter search approach. To constrain the patterns of mantle deformation, we invert Love and azimuthally-varying Rayleigh phase velocity dispersion curves to retrieve the present-day thermal structure of the upper mantle. The thermal structure relates to density anomalies through a linear equation of state. The complete forward problem proceeds as follows (see Figure 1): (1) Given a temperature field, we first numerically solve an instantaneous 3D convection problem with temperature-dependent viscosity (Samuel 2012). (2) Using the obtained velocity field and velocity gradient obtained, we track CPO evolution of olivine crystals where the steady-state assumption of the flow is implied. The result is a complete elastic tensor  $S_{ij}$  at each point in space (Kaminski et al. 2004). (3) The last step involves computing synthetic surface wave dispersion curves using normal mode summation in a spherical earth (Smith & Dahlen 1973) and their azimuthal variations from the full  $S_{ij}$  (Montagner & Nataf 1986).

The inversion explores the model space using a Markov chain Monte Carlo (MCMC) algorithm, and evaluates through Bayesian inference the posterior probability of model parameters. In opposition to conventional tomography where elastic parameters are to be inverted for, our method directly inverts for a single scalar field (*e.g.*, temperature anomalies) and extra information is driven by the physics of mantle convection. The complete solution to our problem is a probability distribution of the 3D present-day thermal structure of the upper mantle. Since

the complete elastic tensor is computed for each sampled model, we can also obtain a posterior distribution of the full elastic tensor. In fact, any variable that is implicitly computed in the forward model can be expressed as a posterior distribution in their respective model space (temperature, flow, deformation, and anisotropy). Thus, geodynamic tomography may be viewed as a technique to reduce model dimension (*i.e.*, the number of inverted parameters) in the inverse problem. Our goal in this study is to lay its proof of concept by applying it to simple synthetic temperature fields. In Section 2, we explain how geodynamic tomography is implemented, starting with the model parameterization, followed by the forward problem, the data, and finally the Bayesian inversion scheme. This is followed by Section 3, where we apply the method to synthetic data obtained from prescribed temperature fields. The last section discusses current limitations of geodynamic tomography, and its potential applications to real-Earth problems.

## 2 METHODOLOGY

Geodynamic tomography involves two main procedures: (1) evaluate the forward model completely, and (2) implement a fully Bayesian non-linear inversion scheme with an MCMC sampling technique. The solution of our inversion scheme is a posterior distribution of thermal structures and their corresponding uncertainty bounds. Fig. 1 illustrates the complete inversion scheme.

### 2.1 Model parameterization

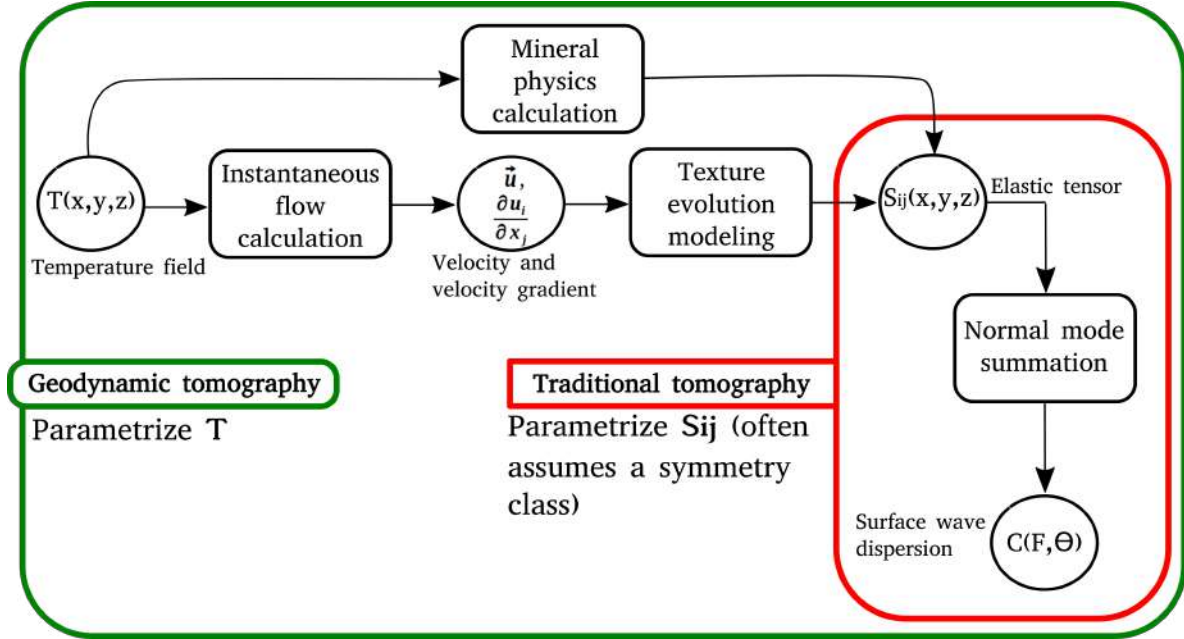
To parameterize the 3D thermal structure in a Cartesian domain  $(x, y, z)$ , we build a basis containing spherical temperature anomalies, on top of an adiabatic temperature gradient. Mathematically, this translates to:

$$T(\mathbf{r}) = T_{\text{background}}(\mathbf{r}) + \sum_{i=1}^M T_{\text{anomaly}}^i(\mathbf{r}) , \quad (1)$$

where the background temperature is assumed to be linear and only a function of depth  $z$ :

$$T_{\text{background}}(\mathbf{r}) = T_0 + \left( \frac{z}{L_s} - 1 \right) (T_0 - 1200 \text{ K}) , \quad (2)$$

and  $M$  is the number of spherical anomalies,  $\mathbf{r} = (\mathbf{x}, \mathbf{y}, \mathbf{z})$  defines any point in the 3D volume,  $T_0$  is the temperature at the bottom (*i.e.*, also the reference value), and  $L_s$  is the characteristic length scale. Each anomaly has a distinct size, temperature, and position. We define the basis



**Figure 1.** Geodynamic tomography (green) in comparison with traditional tomographic techniques (red). In geodynamic tomography, the unknown model to be inverted for is the temperature field denoted by  $T$ , whereas in traditional tomography, the model is a fourth-order elastic tensor  $S_{ij}$  with 21 independent coefficients. Often, tomographers assume a hexagonally symmetric medium onto  $S_{ij}$  to reduce model complexity. The complete forward model (in green) is cast in a Bayesian MCMC framework. One of the advantages of geodynamic tomography is the reduction of unknown model parameters due to constraints from geodynamics.

function for one given spherical anomaly as:

$$T_{\text{anomaly}}(\mathbf{r}) = -\frac{T_c}{2} \left[ 1 - \tanh \left( \frac{\beta}{L_s} \left( \|\mathbf{r} - \mathbf{r}_0\| - \frac{R}{2} \right) \right) \right], \quad (3)$$

where  $T_c$  is maximum temperature anomaly reached at the center of the sphere  $\mathbf{r}_0 = (x_0, y_0, z_0)$ , and  $R$  controls its size. These five variables are unknown model parameters to be inverted for in our problem. The non-dimensional constant  $\beta = 20$  controls the sharpness of the temperature gradient. Additional details can be found in Appendix A.

We model the medium rheology by assuming a temperature-dependent viscosity, following the Frank-Kamenetskii approximation to Arrhenius-type viscosity. Here, we only invert for a dimensionless scalar constant  $E$ , which plays a similar role to the conventional activation energy (*i.e.*, the sensitivity of viscosity to temperature). The viscosity field is described by:

$$\eta(\mathbf{r}) = \eta_0 \exp \left[ -E \frac{(T(\mathbf{r}) - T_0)}{T_0} \right], \quad (4)$$

where  $\eta_0$  is a reference value for viscosity. The total number of parameters defining the model



is therefore  $5M + 1$ , and the corresponding model vector  $\mathbf{m}$  is defined as:

$$\mathbf{m} = [E, x_0^i, y_0^i, z_0^i, R^i, T_c^i, \dots, x_0^M, y_0^M, z_0^M, R^M, T_c^M] . \quad (5)$$

## 2.2 The forward problem

The forward problem involves three main steps: (1) regional flow modeling in 3D Cartesian coordinates, (2) modeling texture evolution and computation of the full elastic tensor, and (3) computation of seismic surface wave dispersion curves. We enhance the computational efficiency in Step 2 by using a surrogate model based on an artificial neural network (ANN) to compute the deformation-induced anisotropy.

### 2.2.1 Flow model

For our instantaneous flow models, we consider the buoyancy-driven convection of a highly-viscous, Newtonian, and incompressible fluid in a 3D Cartesian coordinate system. The flow is subjected to free-slip boundary conditions. The system of equations describing the flow is given by:

$$\nabla \cdot \mathbf{u} = 0 , \quad (6)$$

and

$$-\nabla P + \nabla \cdot [\eta(\nabla \mathbf{u} + \nabla \mathbf{u}^T)] + \rho g \hat{e}_g = 0 , \quad (7)$$

where  $\mathbf{u}$  is the flow velocity,  $P$  is the dynamic pressure, and  $\hat{e}_g$  is a unit vector pointing towards the direction of gravity. We assume density  $\rho$  to be a function of temperature  $T$  using a linear equation of state controlled by a thermal expansion coefficient  $\alpha$ , where  $\rho(T) = \rho_0 - \rho_0 \alpha(T - T_0)$ . The Rayleigh number, a dimensionless quantity that relates to the level of free convection, is chosen such that it is representative of the upper mantle ( $Ra = 1.05 \times 10^6$ ). The dimensional values of the governing parameters are listed in Table 1. The Stokes equations are discretized using a finite-volume approach (e.g., Patankar 1980; Albers 2000), and are solved using the coupled iterative geometric multigrid method using V-cycles (Brandt 1982; Gerya 2010), yielding linear convergence with the number of unknowns. The complete code is parallelized with OpenMP. The accuracy of the numerical solution has been benchmarked against numerical and analytical solutions (Samuel 2012, 2018).

Although the code accommodates sharp viscosity contrasts, the latter tend to reduce the speed of convergence. Sharp viscosity contrasts are avoided in this study since smooth

**Table 1.** Dimensional parameters that define the Rayleigh number.

Symbol	Parameter	Value
$\eta_0$	Viscosity	$10^{21}$ Pa s
$\alpha$	Thermal expansion	$2 \cdot 10^{-5}$ K <sup>-1</sup>
$g$	Gravity	9.81 m/s <sup>2</sup>
$L_s$	Layer thickness	400 km
$T_0$	Temperature scale	1900 K
$k$	Thermal diffusivity	$10^{-6}$ m <sup>2</sup> /s
$\rho_0$	Density	3800 kg/m <sup>-3</sup>
$Ra$	Rayleigh number	$1.05 \cdot 10^6$

thermal structures are considered in our prior distribution. The velocity gradients are obtained by second-order finite differences of the computed velocity field.

### 2.2.2 Modeling intrinsic anisotropy

Upper mantle minerals develop CPO due to progressive shearing along a flow path. We initially model CPO evolution by employing D-Rex, a kinematic model of strain-induced crystal lattice preferred orientation of olivine and enstatite aggregates developed by (Kaminski et al. 2004). The crystal aggregates respond to an imposed macroscopic deformation by two mechanisms: (1) dislocation creep which induces reorientation of each crystallographic axis, and (2) dynamic recrystallization, which allows for the evolution of crystallographic volume fractions by grain nucleation and grain boundary migration. In this study, we only consider pure olivine of type-A fabric corresponding to dry upper mantle conditions. The raw output of D-Rex is a set of crystallographic orientations and volume fractions for a given aggregate. Finally, its effective elastic properties can be estimated with an averaging scheme such as the Voigt average (Mainprice 1990). In Voigt notation, the elastic tensor can be represented as a  $6 \times 6$  matrix with 21 independent elastic coefficients.

D-Rex does not account for pressure and temperature dependence of the single crystal elastic parameters. We model the temperature and pressure dependence of the isotropic seismic wave speeds ( $V_p$  and  $V_s$ ) using *Perple\_X*, a numerical tool that solves the Gibbs free energy minimization problem (Connolly 2005, 2009). We use the thermodynamic model from Stixrude & Lithgow-Bertelloni (2011). We assume olivine mantle composition for isotropic seismic wave speed calculations. Meanwhile, the elastic tensor given by D-Rex is at a reference temperature and pressure. It can be decomposed into an isotropic and anisotropic part  $S_{\text{iso}}$ , and  $\delta S(T_0, P_0)$ ,

respectively:

$$S(T_0, P_0) = S_{\text{iso}}(T_0, P_0) + \delta S(T_0, P_0) . \quad (8)$$

We replace the isotropic part of the tensor with the one computed from `Perple_X`. To account for the pressure and temperature dependence of the anisotropic part, it is scaled by the ratio between the shear modulus  $\mu(T, P)$  at the given pressure and temperature, and the shear modulus at the reference temperature-pressure  $\mu(T_0, P_0)$  (Gallego et al. 2013). Other methods are available, such as the use of first-order corrections around the elastic tensor at ambient  $T$  and  $P$  conditions (Estey & Douglas 1986; Becker et al. 2006). Thus, the full elastic tensor, whose isotropic part depends on pressure and temperature is:

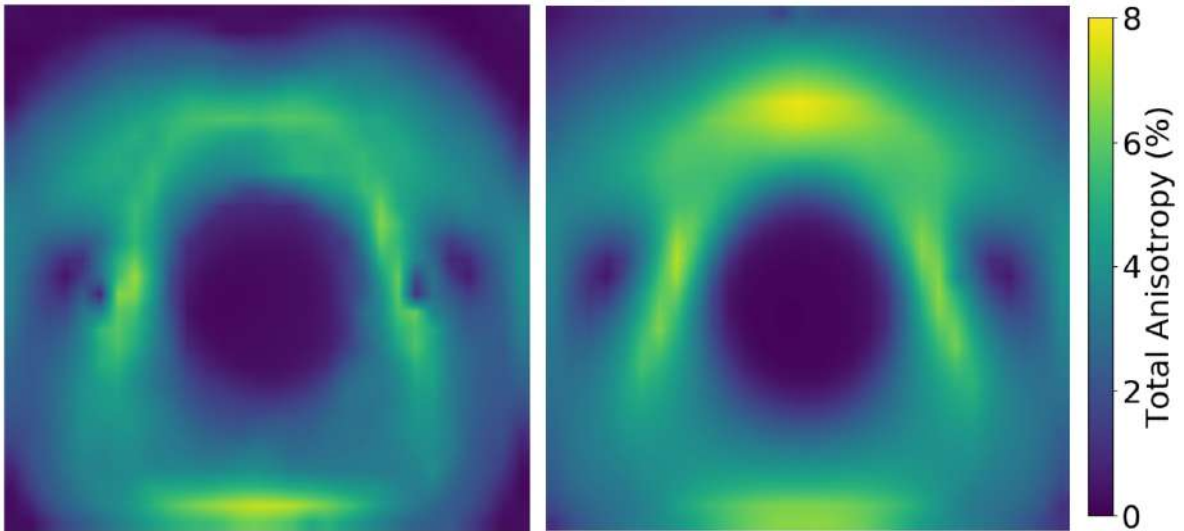
$$S(T, P) = S_{\text{iso}}(T, P) + \frac{\mu(T, P)}{\mu(T_0, P_0)} \delta S(T_0, P_0) . \quad (9)$$

### 2.2.3 Fast forward approximation for texture evolution calculations

Sampling-based techniques such as Markov chain Monte Carlo schemes can be applied to most geophysical inverse problems provided that the parameter space can be sampled efficiently. In some cases however, the forward model is computationally expensive, and sampling-based techniques may not be efficient at approximating a multi-dimensional probability distribution. Fast approximations of the forward model, such as artificial neural networks (ANN) are sometimes therefore used. Such approximations, however, lead to a theoretical error (also called modeling error). The form of these errors can be estimated and modeled as a Gaussian probability distribution with its resulting variance being accounted for in the likelihood function during the inversion process (Hansen et al. 2014b; Köpke et al. 2018). In our case, the computational bottleneck is clearly the texture evolution modeling, which we addressed by using an ANN-based surrogate model to approximate seismic anisotropy.

In the field of geophysics, these methods have already been used to approximate the inverse function in a variety of applications in seismology (e.g., Meier et al. 2007; Käuffl et al. 2014; Hansen & Cordua 2017; Hulbert et al. 2019), and in geodynamics (e.g., Shahnas et al. 2018). Among these studies, some have already applied surrogate models for fast forward approximations in sampling-based techniques (Hansen & Cordua 2017; Köpke et al. 2018; Conway et al. 2019; Moghadas et al. 2020).

These networks are composed of highly non-linear functions that can be trained to approximate a non-linear mapping between an input and an output (Bishop et al. 1995). To approximate such a function, one needs to train this network given a collection of training data consisting of a set of input and output pairs. In this work, we replicate the operator for



**Figure 2.** Vertical cross section of the percentages of total anisotropy obtained from: neural networks (left), and D-Rex (right). The total anisotropy is derived from the norm of the elastic tensor. The slices are oriented along the  $yz$ - plane, and taken at the center of the  $x$ - axis (*i.e.*,  $x = 200$  km)

texture evolution, which we now denote as  $g_{\text{CPO}}$ . Flow streamlines with assigned local velocity gradients are fed into the network as training inputs. The training output contains the anisotropic part of the elastic tensor  $\delta S(T_0, P_0)$  computed from D-Rex. The package `scikit-learn` in Python is used to train the network (Pedregosa et al. 2011)(See Appendix B for full details of the method).

Once the network is trained, which we denote as the operator  $g_{\text{nn}}$ , we perform a simple numerical test of 3D deformation due to a cold spherical temperature anomaly, and applied both operators to output seismic anisotropy. Fig. 2 shows the percentage of total anisotropy found by the two methods. We observe comparable levels of anisotropy. Moreover, the approximation also appears to capture some important features such as the absence of anisotropy at the center, which is ascribed to the larger viscosity of the anomaly in this region. However, the surrogate model tends to underestimate the total anisotropy, which may be attributed to the simplicity of the network architecture, and the number of available training data used.

#### 2.2.4 Predicting surface wave data

For any geographical location at the surface, we can extract the 1D velocity profile (*e.g.*,  $S_{ij}$  as a function of depth) and compute dispersion curves for Love and Rayleigh waves. The azimuthal dependence of surface wave phase velocity can be treated as the sum of a small anisotropic perturbation around an isotropic phase velocity model (Smith & Dahlen 1973)

giving:

$$\mathbf{c}(T, \theta) = \mathbf{c}_0(T) + \mathbf{c}_1(T) \cos(2\theta) + \mathbf{c}_2(T) \sin(2\theta) + \mathbf{c}_3(T) \cos(4\theta) + \mathbf{c}_4(T) \sin(4\theta), \quad (10)$$

where  $T$  is the period and  $\theta$  is the azimuthal angle.

In this work, we only invert  $\mathbf{c}_0(T)$ ,  $\mathbf{c}_1(T)$ , and  $\mathbf{c}_2(T)$  for Rayleigh waves and only  $\mathbf{c}_0(T)$  for Love waves. It is not common to invert other terms, due to low sensitivity or to high levels of noise. For convenience, we denote isotropic Rayleigh wave phase velocity as  $\mathbf{c}_R(T)$  and Love wave phase velocity as  $\mathbf{c}_L(T)$ .

The different terms in eq. (10) can be computed from  $S_{ij}$  in a fully non-linear fashion by normal mode summation with a Runge–Kutta matrix integration (Takeuchi & Saito 1972). We refer the reader to Montagner & Nataf (1986) and Bodin et al. (2016) for details. The seismic forward model is computed using a 1D earth assumption beneath each geographical location. We acknowledge that surface waves velocities depend on 3D heterogeneities, and particularly the fact that surface wave computations exhibit non-linearities due to mode-coupling and finite frequency effects (e.g., Sieminski et al. 2007; Ekström 2011). However, these approximations can be treated as theoretical errors and can be accounted for in the Bayesian inversion procedure.

### 2.3 Bayesian sampling scheme

We formulate the problem in a fully non-linear Bayesian framework (Box & Tiao 2011; Smith 1991; Mosegaard & Tarantola 1995), where the predicted surface wave dispersion curves estimated for a large ensemble of models (3D temperature fields) are compared to observed data. The solution of the inverse problem is the posterior distribution  $p(\mathbf{m}|\mathbf{d})$ , the probability model of parameters  $\mathbf{m}$  given the data  $\mathbf{d}$ . According to Bayes' theorem we have:

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{m}) p(\mathbf{d}|\mathbf{m}). \quad (11)$$

The prior distribution  $p(\mathbf{m})$  describes our predetermined knowledge on  $\mathbf{m}$  (*i.e.*, the position and the amplitude of thermal anomalies, as well as the activation energy). The likelihood function  $p(\mathbf{d}|\mathbf{m})$  describes the probability of observing the data given our current knowledge of the model parameters.

Since our forward problem is highly non-linear, the posterior distribution is sampled using a Markov chain Monte Carlo algorithm. It involves direct sampling of the parameter space by random iterative search, where the distribution of the sampled models asymptotically converges towards the posterior distribution.

### 2.3.1 The likelihood function

The likelihood function  $p(\mathbf{d}|\mathbf{m})$  quantifies how well the model parameters explain the observed data (*i.e.* the ensemble of local dispersion curves located at the surface). Supposing that each data-type (*i.e.*  $\mathbf{c}_R$  and  $\mathbf{c}_L$  for isotropic Rayleigh and Love wave dispersion curves, respectively;  $\mathbf{c}_1$  and  $\mathbf{c}_2$  for Rayleigh wave anisotropy) is measured independently, the likelihood function gives:

$$p(\mathbf{d}|\mathbf{m}) = p(\mathbf{c}_R|\mathbf{m}) p(\mathbf{c}_L|\mathbf{m}) p(\mathbf{c}_1|\mathbf{m}) p(\mathbf{c}_2|\mathbf{m}) . \quad (12)$$

For all dispersion curves, we assume that the errors are uncorrelated and follow Gaussian distributions with zero mean, and variances  $\sigma_{c_R}^2$ ,  $\sigma_{c_L}^2$ ,  $\sigma_{c_1}^2$ , and  $\sigma_{c_2}^2$ . For isotropic Rayleigh waves and isotropic Love waves  $\mathbf{c}_R$  and  $\mathbf{c}_L$ , respectively, we can express the likelihood function as a Gaussian distribution:

$$p(\mathbf{c}_{R,L}|\mathbf{m}) = \frac{1}{(2\pi\sigma_{c_{R,L}}^2)^{\frac{N}{2}}} \exp \left[ \frac{-\|\mathbf{c}_{R,L}^{\text{obs}} - \mathbf{c}_{R,L}(\mathbf{m})\|^2}{2\sigma_{c_{R,L}}^2} \right] . \quad (13)$$

Here, the likelihood function corresponds to a single dispersion measurement where  $N$  is the number of discrete periods. The likelihood functions of the  $2\theta$  terms,  $\mathbf{c}_1$  and  $\mathbf{c}_2$ , can be written in the same manner as eq. (13).

### 2.3.2 A maximum likelihood estimate of data errors

In general, it is difficult to estimate  $\sigma_{c_{R,L}}$  due to the lack of knowledge on the error distribution. In particular, approximating an elastic tensor with a neural network may introduce errors that are difficult to quantify.

In this work, we use a maximum likelihood estimate (MLE) of the noise parameters  $\sigma_{c_{R,L}}$  and  $\sigma_{c_{1,2}}$  following the work of Dettmer et al. (2007). This is performed by maximizing the likelihood function over the data standard deviation. The strength of this technique is that it is not necessary to estimate each contribution to the noise parameters individually. Maximizing eq. (13) over  $\sigma_{c_{R,L}}$  yields:

$$\sigma_{c_{R,L}} = \left[ \frac{1}{N} \sum_{i=1}^N (\mathbf{c}_{R,L}^{\text{obs}} - \mathbf{c}_{R,L}(\mathbf{m}))^2 \right]^{1/2} . \quad (14)$$

Substituting eq. (14) onto eq. (13), and taking the log likelihood we obtain:

$$\ln[p(\mathbf{c}_{R,L}|\mathbf{m})] = -\frac{N}{2} \ln \left[ \sum_{i=1}^N (\mathbf{c}_{R,L}^{\text{obs}} - \mathbf{c}_{R,L}(\mathbf{m}))^2 \right] . \quad (15)$$

The log likelihood functions of  $c_1$  and  $c_2$  can be defined using the same procedure. This method has two advantages: (1) the absolute value of errors need not be defined, and (2) in

the case of joint inversion, we do not have to define the relative weights between each data type. Finally, the full log likelihood function gives:

$$\ln[p(\mathbf{d}|\mathbf{m})] = \ln[p(\mathbf{c}_R|\mathbf{m})] + \ln[p(\mathbf{c}_L|\mathbf{m})] + \ln[p(\mathbf{c}_1|\mathbf{m})] + \ln[p(\mathbf{c}_2|\mathbf{m})] . \quad (16)$$

### 2.3.3 The prior distribution

In Bayesian inference, one expresses the *a priori* information in terms of a probability distribution  $p(\mathbf{m})$ . In geophysical inverse problems, model parameters are typically given a uniform prior distribution with given upper and lower bounds inferred from prior knowledge (Mosegaard & Sambridge 2002). Adopting the same formulation, the prior can be written as:

$$p(m_i) = \begin{cases} 0 & m_i > m_{max}, m_i < m_{min} \\ \frac{1}{\Delta m} & m_{min} \leq m_i \leq m_{max} , \end{cases} \quad (17)$$

where  $m_{max}$ , and  $m_{min}$  are the prior bounds for the model. Assuming that the model parameters in our inversion are prior independent, we can express the prior fully as:

$$p(\mathbf{m}) = p(E) \prod_{i=1}^M \left[ p(x_0^i) p(y_0^i) p(z_0^i) p(R^i) p(T_c^i) \right] , \quad (18)$$

where  $p(E)$  is the prior distribution for the activation energy, and  $M$  is the total number of spherical temperature anomalies. For an  $i$ th temperature anomaly,  $p(x_0^i)$ ,  $p(y_0^i)$  and  $p(z_0^i)$  are the prior distributions for position;  $p(R^i)$  and  $p(T_c^i)$  are the prior distributions for the size and temperature, respectively. We choose wide uniform prior distributions. For the prior bounds, we select: (1) the length of the spatial domain (0 km to 400 km) for the positions  $x_0$ ,  $y_0$ , and  $z_0$ , (2) 40 km to 240 km for  $R$ , (3) 500 K to 1200 K for  $T_c$ , and (3) 6 to 12 for  $E$ . Choosing wide bounds ensures that the model parameters are loosely constrained from the prior, and more emphasis is given to the information provided by the data.

### 2.3.4 A random walk to sample the posterior distribution

We use a Markov chain Monte Carlo algorithm to sample the posterior distribution. It begins by randomly selecting an initial temperature model followed by the evaluation of the initial log likelihood. At each iteration, the current model is perturbed to propose a new model. The proposal proceeds sequentially as follows:

(i) **Assign local perturbation:** One sphere is randomly picked out of  $M$  number of spheres. Once a sphere is picked, we randomly select one of four possible ways to perturb the sphere as follows:

- Perturb horizontal position; *i.e.*  $x_0$  and  $y_0$  together;
- Perturb vertical position  $z_0$ ;
- Perturb the size of the sphere  $R$ ;
- Perturb the temperature of the sphere  $T_c$ .

Each perturbation is drawn from a univariate normal distribution centered at the current value of the model parameter.

(ii) **Perturb the activation energy:** We then apply eq. (1) to define the 3D temperature field. Alongside, we perturb the activation energy  $E$  by using a normal distribution centered at the current value of  $E$ , and apply eq. (4) to define the 3D viscosity field. These two scalar fields are used as inputs in the flow calculation.

If the proposed model lies within the prior bounds following eq. (18), we evaluate the forward problem completely. The computed dispersion curves from the latter are compared with the observed data using eq. (15). The resulting likelihood is then compared to the likelihood of the current model, and the proposed model is either accepted or rejected according to an acceptance probability (Metropolis et al. 1953; Hastings 1970). If the proposed model is accepted, it becomes the current model for the next iteration. After a sufficient number of iterations, the ensemble of accepted models converges towards the posterior distribution.

### 3 APPLICATION WITH 3D SYNTHETIC TEMPERATURE FIELDS

#### 3.1 Inversion for One Spherical Anomaly

We demonstrate our proof of concept by setting up a simple temperature field consisting of one spherical negative temperature anomaly (*i.e.* negatively buoyant) placed at the middle of a  $400 \text{ km} \times 400 \text{ km} \times 400 \text{ km}$  box. The setup is a very simple toy example inspired by the work of Baumann et al. (2014) where they applied Bayesian inversion to constrain rheology from gravity anomalies and surface velocities.

Table 2 shows the complete list of true model parameters, and Fig. 3 displays a cross-sectional view of the temperature field, and its associated instantaneous velocity field.

We simulate the full forward model given the true model parameters to generate synthetic dispersion curves at periods between 10s and 200s. Fig. 4 shows a map of the computed phase velocity and azimuthal anisotropy for Rayleigh waves at 100 seconds. In Fig. 4a, the phase velocity is maximum at the middle of the region, due to the presence of the cold anomaly underneath.

Fig. 4b shows a map of azimuthal anisotropy in Rayleigh waves. Here, anisotropy is at

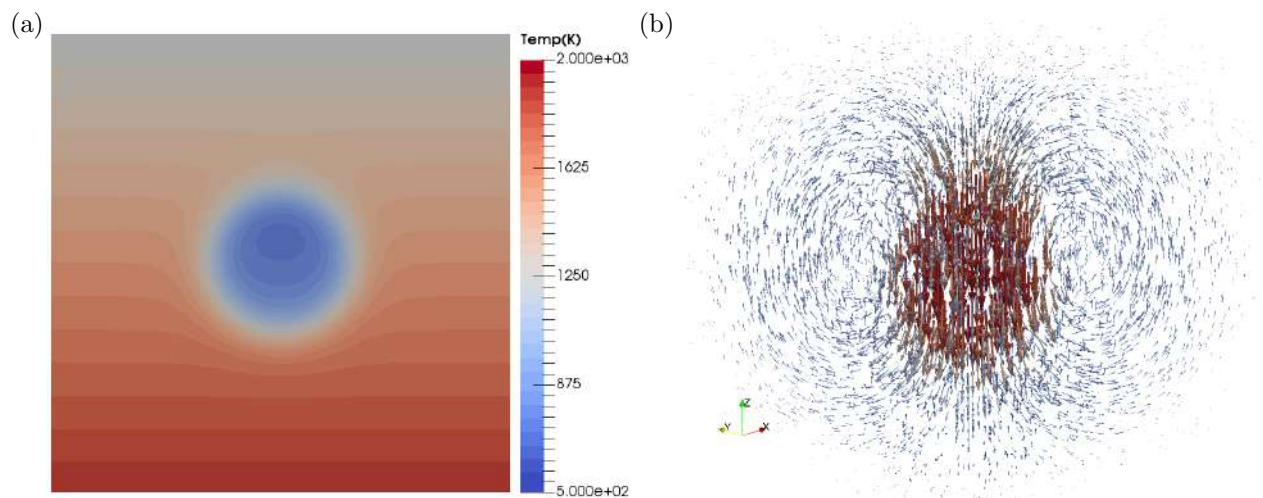


**Table 2.** True model parameters defining the synthetic temperature field.

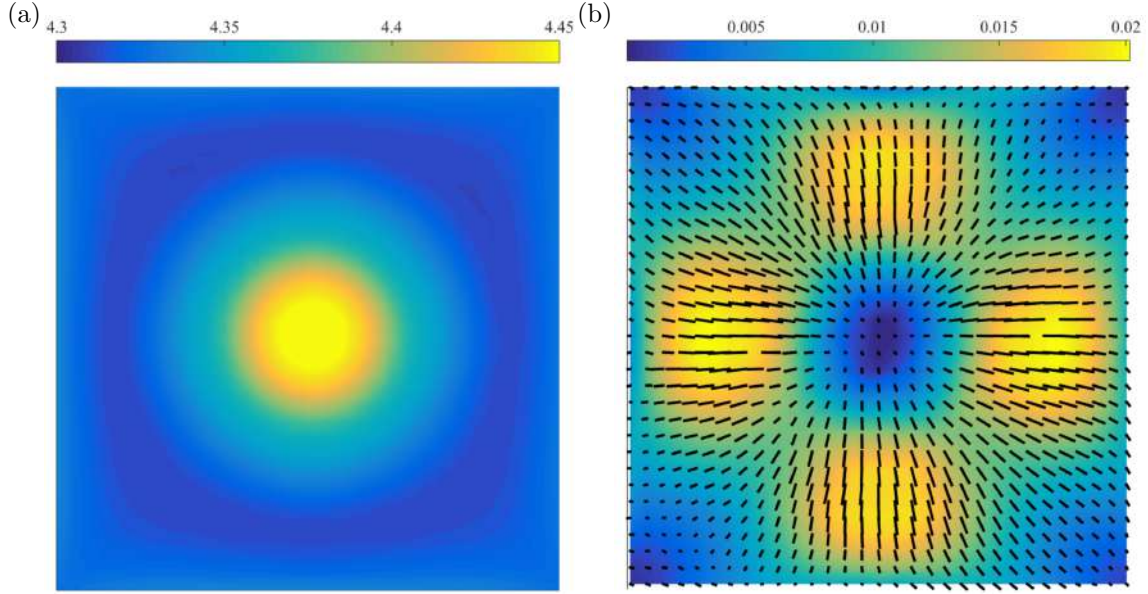
Model parameter	Assigned value
$x_0$	200 km
$y_0$	200 km
$z_0$	200 km
$R$	120 km
$T_c$	800 K
$E$	11.0

its minimum at the center, above where the cold more viscous anomaly is located. As a result of this higher rigidity, local velocity gradients are lower, resulting in smaller amounts of deformation and hence lower anisotropy. Another feature is the presence of strong anisotropy at certain locations. These regions are points where shear deformation is at its maximum due to the convergence of flow lines. On top of the level of azimuthal anisotropy is the orientation of its fast axis. Since we expect the flow direction to converge towards the center when observed from the top, the fast axis may be interpreted as the horizontal projection of the flow.

The complete data constitute a regular array of  $8 \times 8$  locations containing  $\mathbf{c}_R$ ,  $\mathbf{c}_L$ ,  $\mathbf{c}_1$ , and  $\mathbf{c}_2$  spanning the entire surface. We emphasize that the data generated comes from an elastic tensor computed with D-Rex whereas during inversion, the estimated data is obtained from an elastic tensor approximated by neural networks.



**Figure 3.** (a) Cross-sectional view in the  $yz$ - plane of the 3D temperature field. The slice is taken at the center of the  $x$ - axis. (b) 3D flow velocity due to the sinking anomaly. Largest flow magnitudes correspond to the cold anomaly.

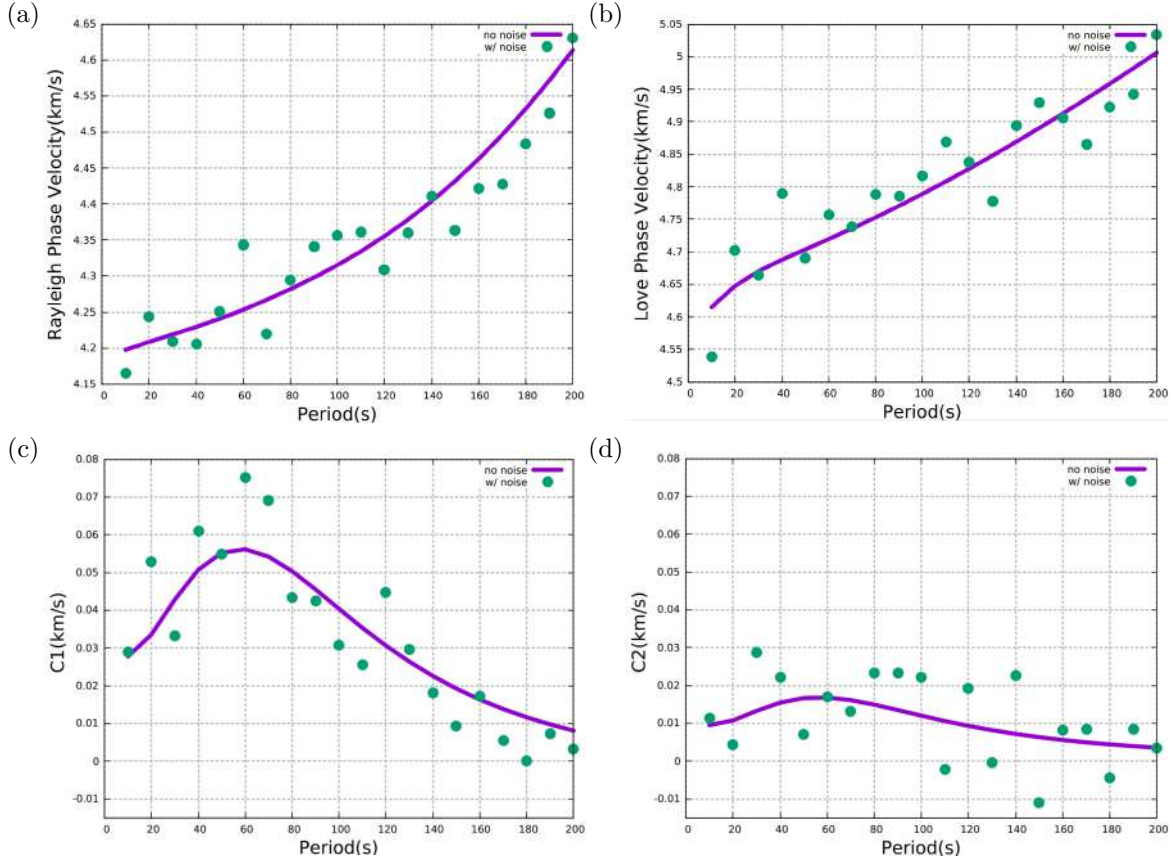


**Figure 4.** Phase velocity maps resulting from one sinking anomaly at 100s period. (a) Rayleigh wave phase velocity (km/s). (b) Azimuthal anisotropy in Rayleigh waves (km/s). The solid black lines correspond to the direction of the fast propagation axis. Surface wave maps always lie along the  $xy$ –lateral plane.

Finally, we added random uncorrelated noise onto  $\mathbf{c}_R$ ,  $\mathbf{c}_L$ ,  $\mathbf{c}_1$ , and  $\mathbf{c}_2$ . Standard deviations for Love and Rayleigh are set at  $\sigma_R = 0.05$  km/s and  $\sigma_L = 0.05$  km/s, whereas  $\sigma_{c_1} = 0.01$  km/s and  $\sigma_{c_2} = 0.01$  km/s. Fig. 5 illustrates the resulting dispersion curves at one given location with and without noise.

The inversion consists of 20 independent Markov chains each containing 40,000 samples initiated at a random temperature structure. We demonstrate two cases. First is an isotropic inversion, where no anisotropy is involved in the forward model. In this case, it is not necessary to compute instantaneous flow and anisotropy, as isotropic seismic velocities  $V_p$  and  $V_s$  can be directly scaled with temperature. The inverted data are the isotropic phase velocities  $\mathbf{c}_R$  and  $\mathbf{c}_L$ . Secondly, we present an anisotropic inversion (geodynamic tomography). Both isotropic and anisotropic inversions are given the same wide uniform priors allowing for more mobility when searching the parameter space. We initiate geodynamic tomography by first employing an isotropic inversion. Once the chains have converged in this phase, we then start the actual anisotropic inversion procedure.

The diagonal panels of Figs 6 and 7 illustrate the ensemble of models recovered from isotropic inversion and anisotropic inversion. The off-diagonal panels depict 2D marginal distributions as 2D histograms to explore possible trade-offs. The black circles indicate the values of the true model parameters. Compared to isotropic inversion, the width of the posterior

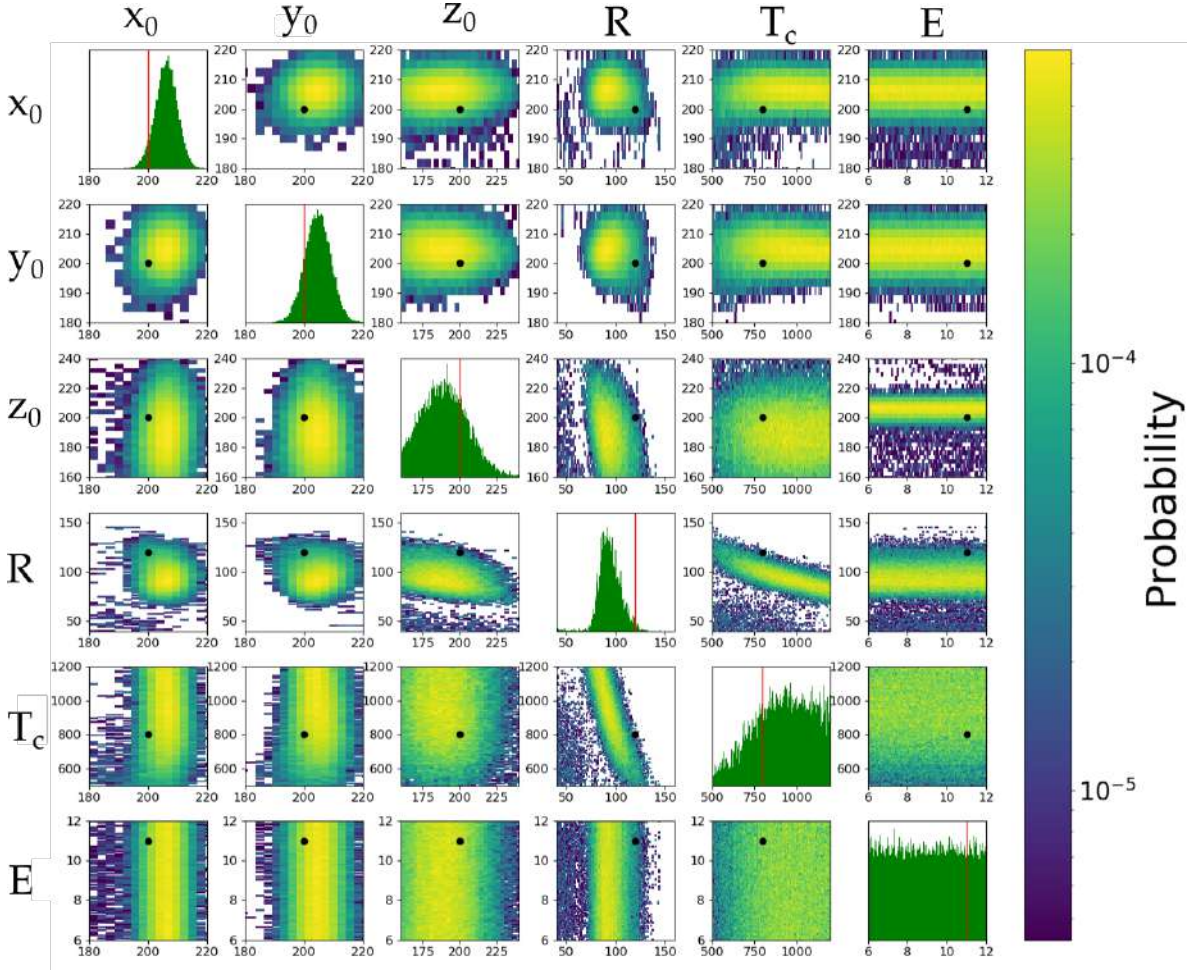


**Figure 5.** Synthetic surface wave dispersion curves from 10s to 200s at a given location: (a) Rayleigh wave phase velocity, (b) Love wave phase velocity, (c) Rayleigh anisotropy  $c_1$ , (d) Rayleigh anisotropy  $c_2$ . Scatter plot: observed dispersion curve with added noise. Line plot: observed dispersion curve without noise.

distribution inferred from geodynamic tomography has been reduced considerably. More information is thus added by introducing geodynamic constraints in the tomographic problem.

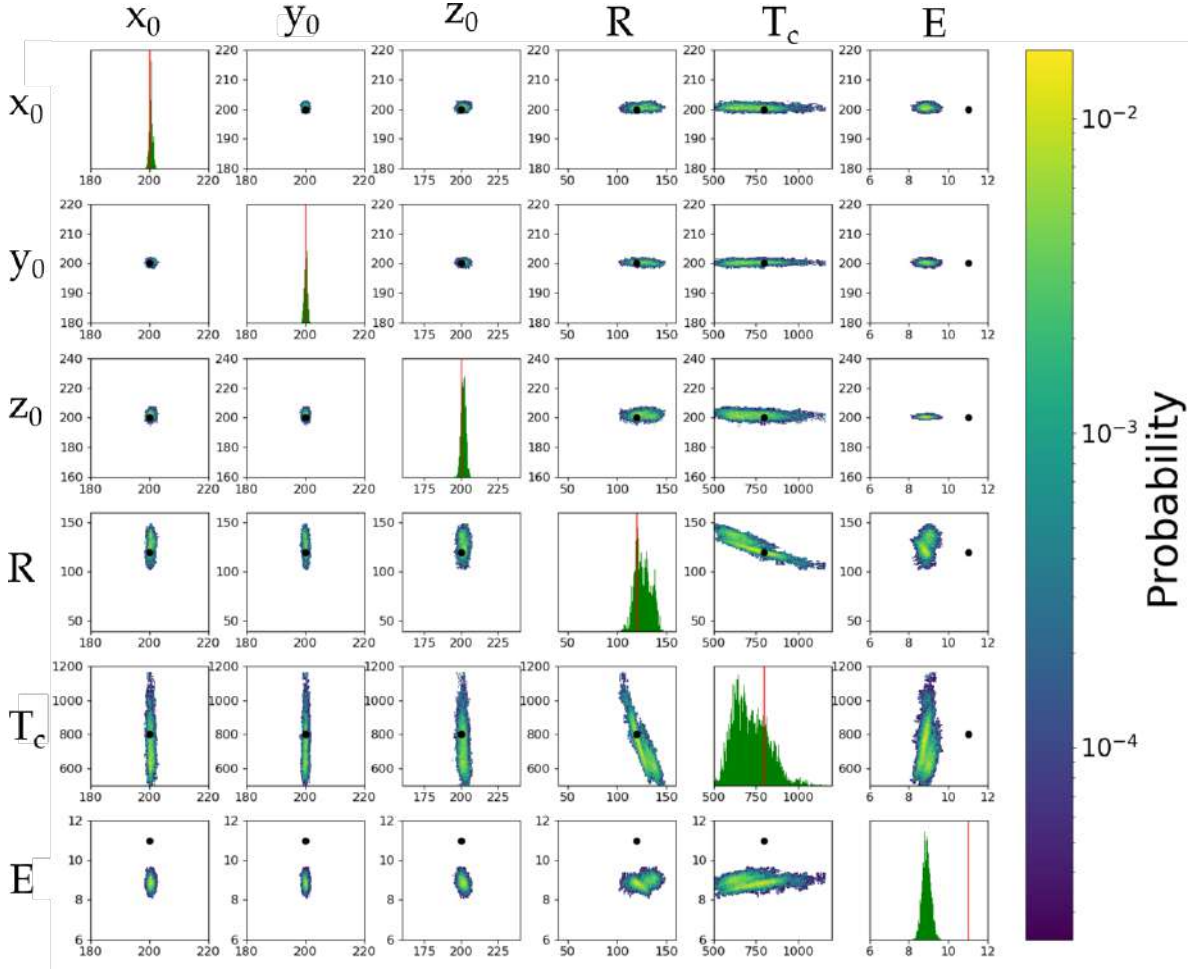
As expected, the posterior distribution on the activation energy  $E$  in the isotropic case is flat, as isotropic velocities are only sensitive to temperature and not to viscosity. Anisotropic inversion, on the other hand, constrains  $E$  as shown in Fig. 7. The distribution, however, appears to be distant from the correct value of  $E$ . Such a behavior is also evident in its 2D marginal posterior where the true value is outside the inferred distribution. This clearly exhibits a bias which is deduced from the imperfections of the neural network when computing anisotropy. This effect is eliminated when one uses the correct forward operator for modeling anisotropy. Another distinct feature in these figures is the negative trade-off between  $T_c$  and  $R$ , which may be attributed to the symmetry of the problem considered. An increase in temperature of the anomaly compensates for an increase in its radius. Such trade-offs may be reduced in the case where the true model exhibits less symmetry.





**Figure 6.** Posterior probability distribution in the 6-dimensional parameter space inferred from the isotropic inversion  $p(\mathbf{m}|\mathbf{c}_R, \mathbf{c}_L)$ . Diagonal panels show 1D marginal distributions for each model parameter. Off-diagonal panels show 2D marginal distributions and depict possible trade-offs between pairs of model parameters. The red vertical lines and the black markers indicate the true model values for the diagonal and the off-diagonal panels, respectively. The intensity pertains to the level of posterior probability (*i.e.*, high intensity means high probability, and thus low misfit).

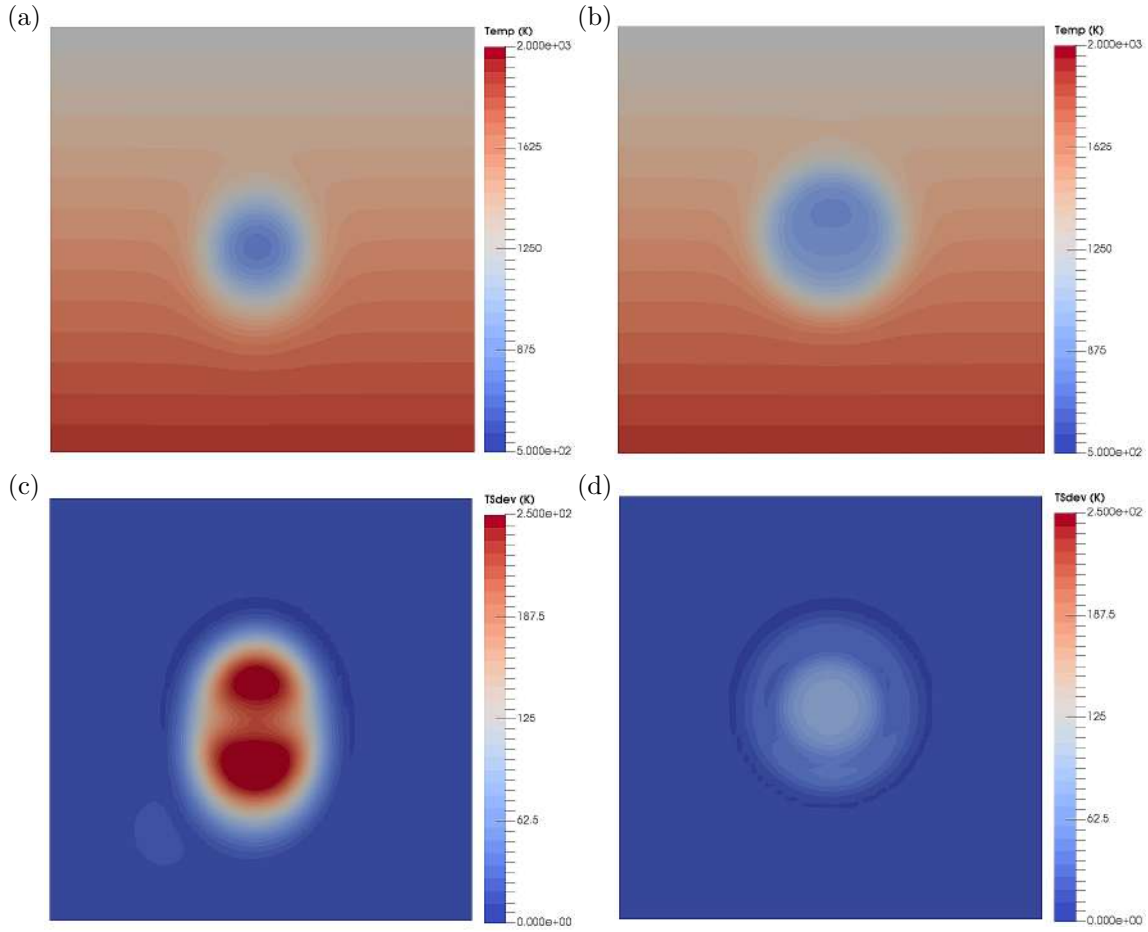
We also plot the mean temperature models from both inversions (see Fig. 8). The figures are obtained by averaging the temperature values at each point. By visual inspection, anisotropic inversion better resolves the 3D thermal structure. This is further supported by the standard deviation computed around the mean temperature at a given pixel as shown in Figs 8c and 8d. In both cases, the standard deviations is higher at the center of the box, where the spherical anomaly is located. This is due to the variations in the location and amplitude of the sphere in the ensemble of sampled models. In the anisotropic case, the vertical position of the sphere is less constrained than its horizontal position, as can be seen in the 2D histograms. The ensemble of sampled spheres therefore share the same horizontal position but have a vari-



**Figure 7.** Posterior probability distribution in the 6-dimensional parameter space inferred from the anisotropic inversion  $p(\mathbf{m}|\mathbf{c}_R, \mathbf{c}_L, \mathbf{c}_1, \mathbf{c}_2)$ .

able vertical position, which explains the shape of the standard deviation map in Fig. 8c. The posterior uncertainties are also relatively small compared to the recovered temperature field, implying that sufficient information can be retrieved from the noisy dispersion curves.

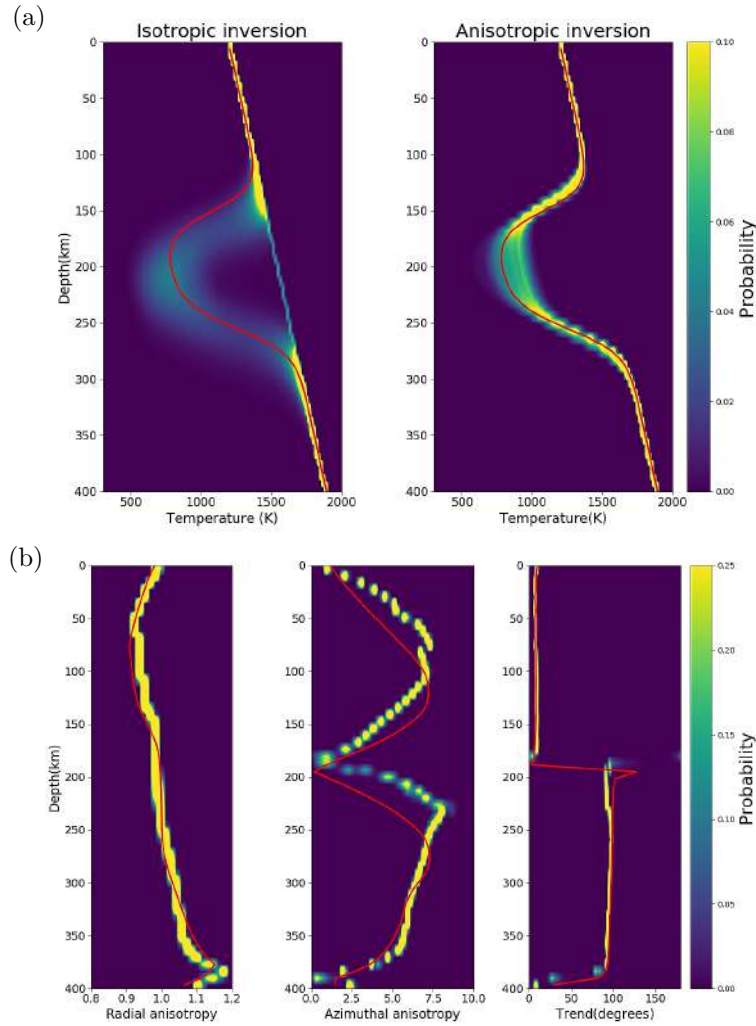
Fig. 9a shows the 1D depth marginal posterior probability profiles (see captions for further details) for temperature, and Fig. 9b for radial anisotropy  $\xi$ , peak-to-peak azimuthal anisotropy, and the trend of the fast-axis of azimuthal anisotropy at a given location. Both methods capture the 1D structures for temperature. However, by adding geodynamic constraints (*i.e.*, anisotropic inversion), we observe that the temperature is much better resolved. Additionally, we successfully recover radial anisotropy and azimuthal anisotropy without having to explicitly invert for the elastic tensor (see Fig. 9b). Here, due to the positioning of the chosen depth profile for temperature (passing nearly through the center of the anomaly), the azimuthal anisotropy appears to be nonexistent at this location. For that reason, we consider



**Figure 8.** Upper panel: Cross-sectional view in the  $yz$ - plane of the mean temperature field recovered from (a) isotropic inversion, and (b) anisotropic inversion. Lower panel: Standard deviations around the mean temperature fields from (c) isotropic inversion, and (d) anisotropic inversion. These cross-sections are taken at the center of the  $x$ - axis.

another depth profile ( $x = 325$  km,  $y = 225$  km) where azimuthal anisotropy is noticeable (Fig. 9b–middle).

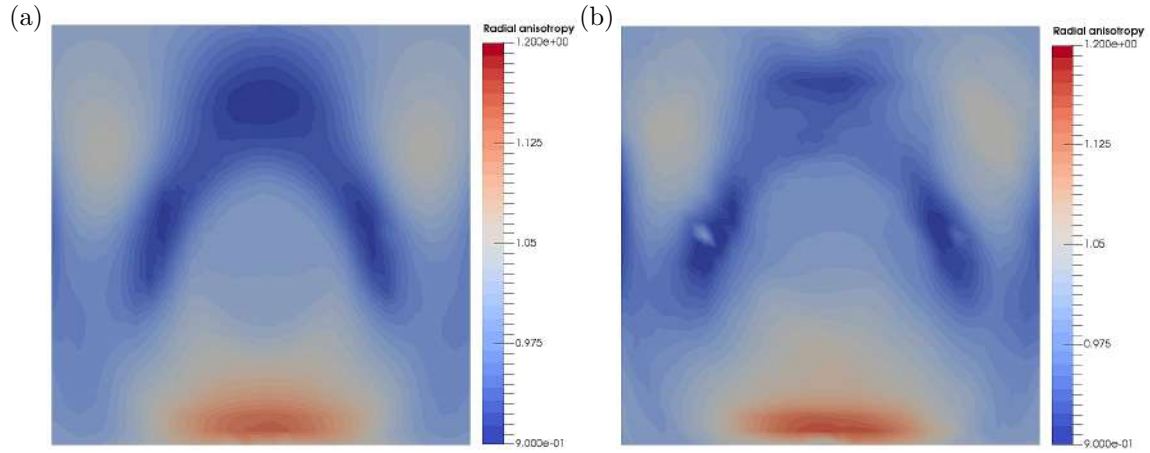
This method also allows us to resolve 3D structures of seismic properties. In fact, any implicitly computed variable can be restructured in 3D. Figs 10 and 11 show the resulting structures computed from the mean temperature model placed side by side with that of the true model. It appears that the value of anisotropy computed with the neural network is underestimated compared to that of D-Rex when using the same input model. This explains why the activation energy  $E$  resulting from the inversion is lower compared to the true value: to produce larger anisotropy and replicate the same output as obtained from D-Rex, one has to reduce the value of  $E$ . Indeed, reducing the viscosity of the material allows for a stronger deformation. The resulting percentage of total anisotropy from both figures are nearly



**Figure 9.** Upper panel: Probability density plots of temperature with depth. Lower panel: Probability density plots of radial anisotropy, peak-to-peak azimuthal anisotropy, and its fast axis with depth. The depth profiles of temperature and radial anisotropy are taken nearly through the center of the sphere. To show that azimuthal anisotropy is also well-constrained, we took a depth profile at  $(x = 325 \text{ km}, y = 225 \text{ km})$ , where azimuthal anisotropy is large. Geodynamic tomography offers the capability to constrain seismic anisotropy. The solid red lines indicate the true structures.

identical. Fig. 10 shows the presence of positive radial anisotropy at the bottom, indicating horizontal flow. Due to the imposition of free-slip boundary conditions combined with zero normal velocities imposed on all surfaces, the flow at the bottom of the box is oriented nearly horizontally. The negative radial anisotropy we observe implies vertical flow (see Fig. 10 caption for details). This is a result of convection cells forming at the sides of the anomaly as it sinks. At the top of the anomaly, negative radial anisotropy also indicates vertical flow due to downwelling. Finally and as we expect, radial anisotropy at the middle is nearly unity due to the presence of the more viscous anomaly. The difference in the structures may be

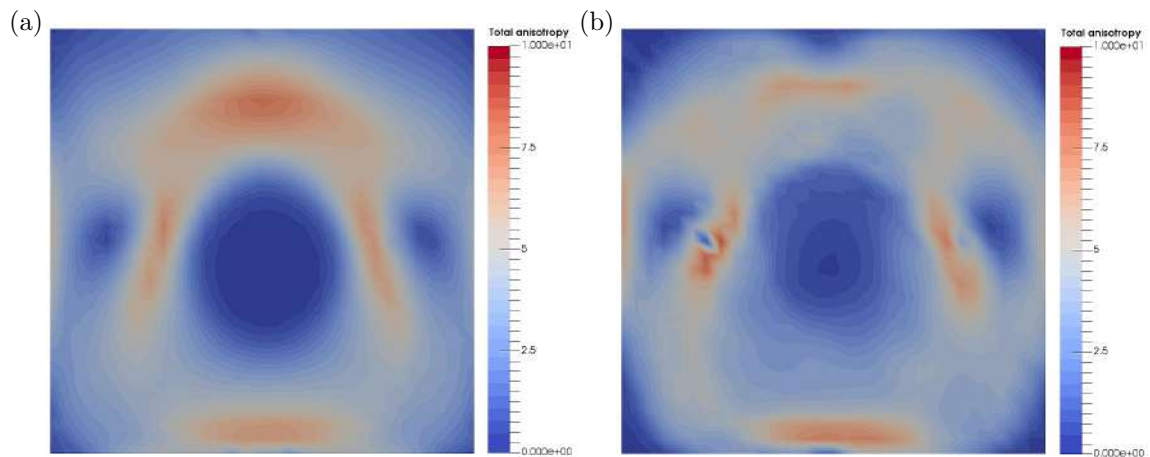




**Figure 10.** Cross-sectional view in the  $yz$ - plane of the radial anisotropy  $\xi$  inferred from (a) true model, (b) mean model. Radial anisotropy is often used as a proxy to infer flow orientation. A  $\xi > 1$  (positive radial anisotropy) is often interpreted as horizontal flow. A  $\xi < 1$  (negative radial anisotropy) on the other hand, pertains to vertical flow. A  $\xi = 1$  indicates the absence of radial anisotropy. The cross sections are taken at the center of the  $x$ - axis.

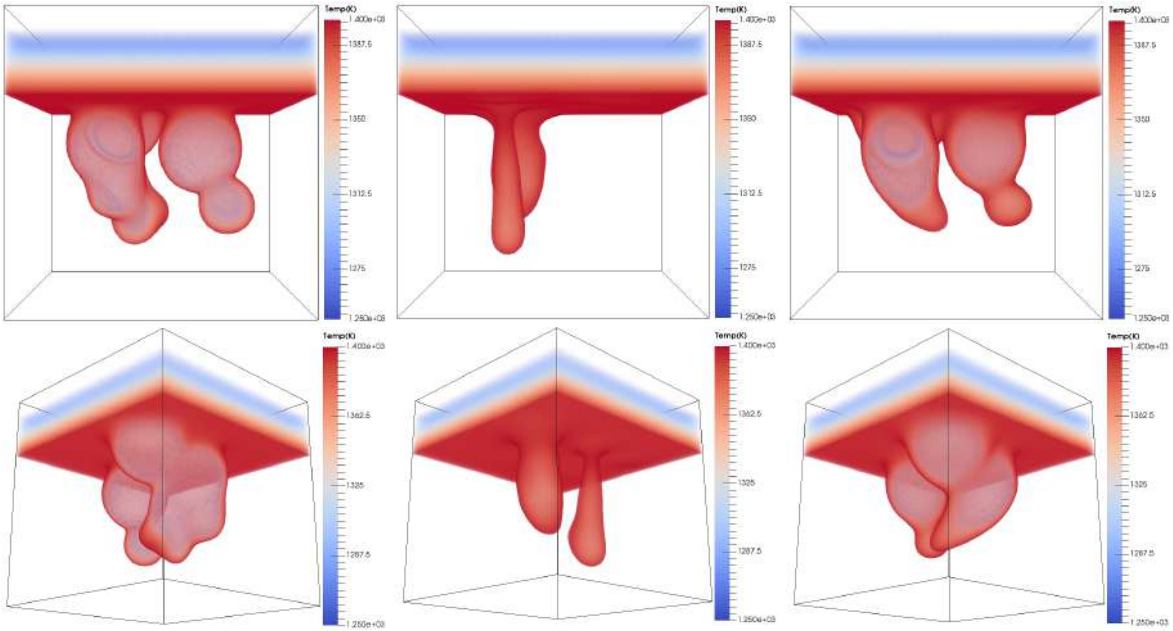
attributed to the following: (1) imperfections of the forward model used in the inversion; (2) information loss related to data sensitivity, and data noise.

We tested the convergence of the Markov chain by plotting the estimates for data errors with MC steps. For further details, refer to Appendix C.



**Figure 11.** Cross-sectional view in the  $yz$ - plane of the percentage of total anisotropy (*i.e.*, norm of  $S_{ij}$ ) inferred from (a) true model, (b) mean model. The absence of anisotropy at the center corresponds to a region of minimal deformation for the cold and highly-viscous anomaly. The cross sections are taken at the center of the  $x$ - axis.



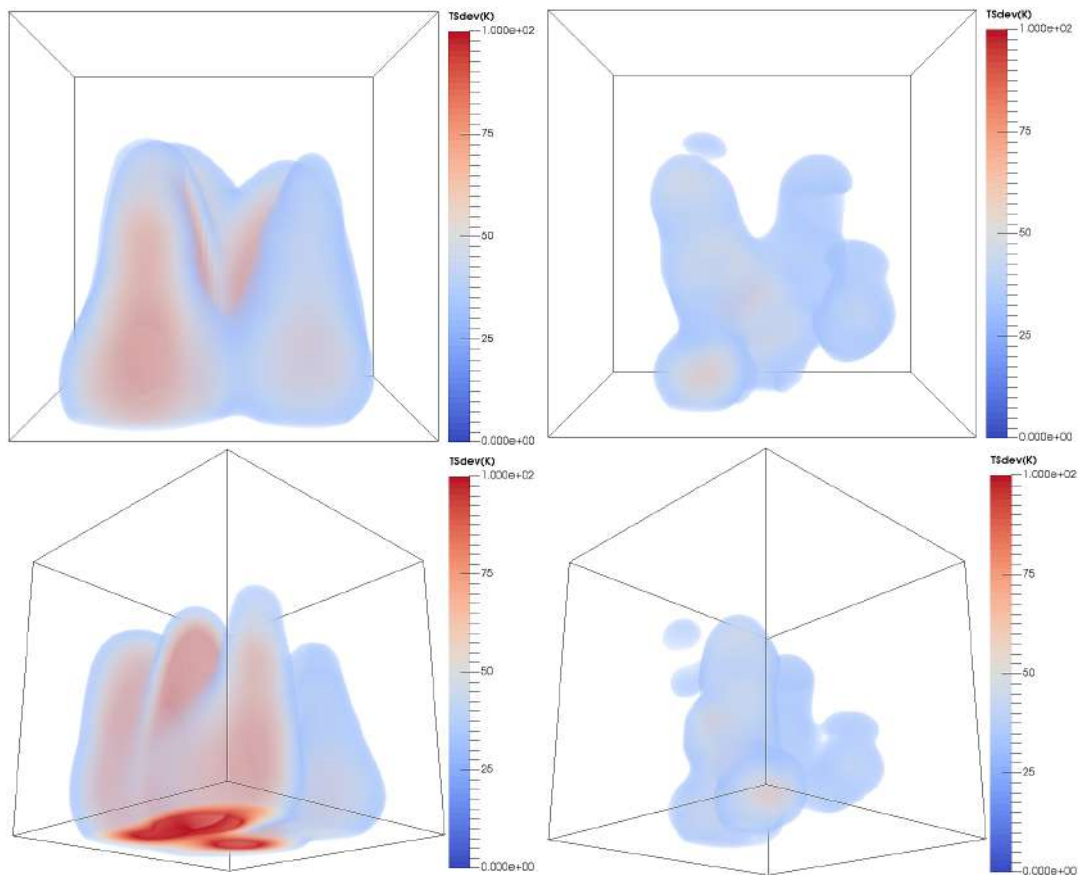


**Figure 12.** Isovolumetric view of the temperature fields. Left: True temperature field. Middle: Mean temperature field from isotropic inversion. Right: Mean temperature field from anisotropic inversion.

### 3.2 Inversion for Multiple Spherical Anomalies

This section covers the inversion for ten spherical temperature anomalies with different properties (*i.e.*, temperature  $T_c$  and radius  $R$ ), positioned randomly in 3D space. Such parameterization scheme may be essential to represent anomalies with complex shapes (*e.g.*, subducting slab) using a collection of several spheres with different characteristics. The synthetic data is generated from a true temperature model consisting of ten spherical anomalies as well. We compare the true temperature model with the mean temperature models obtained from isotropic and anisotropic inversions (Fig. 12). Even with this much more complex structure, we are able to recover the main features of the temperature field. Also, as in the test of Section 3, anisotropic inversion better recovers the structure than isotropic inversion. Posterior uncertainties are represented in Fig. 13 and support this observation. However, some differences with the exact true structure remain, even using anisotropic inversion. Surface waves are long period observations and hence, small and sharp thermal anomalies may not be resolved. Other contributing factors involve the very nature of the tomographic problem itself as enumerated earlier (*e.g.*, data and modeling errors).

In Fig. 14, we choose one depth profile to show the 1D marginal posterior probability densities for temperature, radial anisotropy, and azimuthal anisotropy. The dashed black lines represent the true model. Based on the recovered profiles, anisotropic inversion resolves temperature better than the isotropic case again due to the complementing information brought



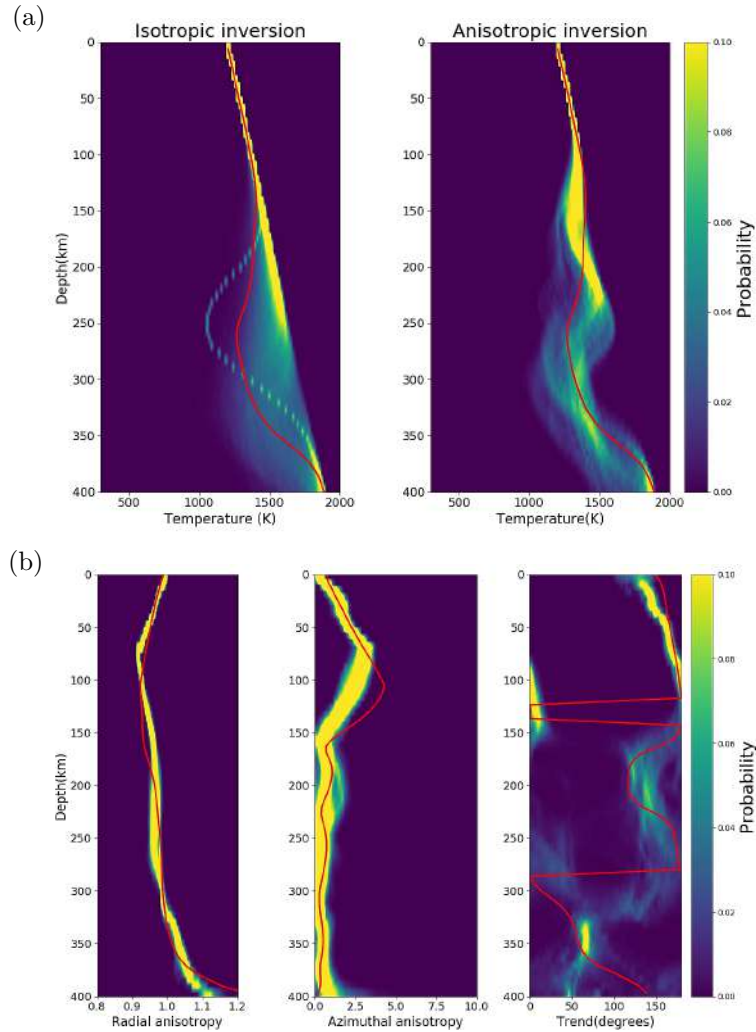
**Figure 13.** Isovolumetric view of the standard deviations around the mean temperature models. Left: Standard deviation for the isotropic inversion. Right: Standard deviation for the anisotropic inversion.

by geodynamic constraints. Radial and azimuthal anisotropy still appears to be tightly constrained; however with some noticeable deviations from the true model.

## 4 DISCUSSION

### 4.1 Additional Comments on the Method

**Model parameterization.** The goal of this study was to test the method in the most simple cases, and we acknowledge that our parameterization of the temperature field in terms of a sum of spherical anomalies is simplistic. However, such parameterization can be applied to invert for more complex geometries such as a detached slab, a homogeneous plume, or upper mantle structures beneath cratons. A step further will be to test more realistic approaches. One possible alternative parameterization is the use of initial temperature models inferred from isotropic tomography, and an iterative update of the structure based on the anisotropy signature at the surface (*i.e.*, anisotropic surface wave dispersion curves). This, however, may



**Figure 14.** Upper panel: Comparison between isotropic and anisotropic inversion. Probability density plots of temperature with depth. The profiles are taken nearly through the center of the sphere. Lower panel: Anisotropic inversion: probability density plots of radial anisotropy, peak-to-peak azimuthal anisotropy, and its fast axis with depth. All profiles correspond to the temperature profile above. The solid red lines indicate the true structures.

only be feasible at the global scale due to boundary effects. It should still be possible to apply this technique at the regional scale, but the structure of interest should be far from the borders of the region considered in order to avoid these boundary effects. Another simple yet effective parameterization would be to invert for constant parameters (*e.g.*, density, viscosity) within geometrical blocks defined from *a priori* information regarding the tectonics of the region (Baumann et al. 2014). In general, the quality of the results will depend on the choice of the model parameters, and the prior information available for the region of interest.

**Neural network-based approach to texture evolution.** The computational demands

of direct sampling techniques such as MCMC is high, as it requires evaluating the forward model a large number of times. Among all routines involved in the forward model, calculating CPO anisotropy proved to be the most costly. We therefore devised a surrogate model that computes texture evolution via a neural network, thus reducing the computation time by three orders of magnitude compared to D-Rex (see Appendix B for absolute computation times of both methods).

However, the surrogate model introduces theoretical errors, which can be reduced by using a network architecture or a training procedure more adapted to the problem at hand. More accurate predictions could be obtained by using a larger training data set, but this has a higher initial computational cost. We observed that the surrogate model does not generalize well. It has been trained for a specific type of flow (convective flows due to spherical temperature anomalies), and thus provides correct predictions only for flow models of the same nature. However, only these specific flow types are tested in the MCMC scheme, and it is therefore not necessary here to have a general neural network that applies to any type of flows.

The success of our synthetic tests is in some ways a proof of the quality of the neural network. The inverted anisotropic seismic data sets were calculated using the exact D-Rex model. Therefore, any errors introduced by the network would manifest themselves by producing a poor fit to the observed data. These theoretical errors have been quantified and accounted for in the Bayesian inversion (see Section 2.3.2). If we want to treat another problem, such as a sinking slab with complex geometry, one needs to re-train the surrogate model for the specific parameterization and prior distribution used. A possible future avenue of geodynamic tomography that is independent of this specific step would be to directly parameterize mantle flow, and build a family of expected convection patterns (together with their predicted anisotropy) to investigate flow patterns underneath mid-ocean ridges and subduction zones. Such parameterization can be easily extended to the global scale by treating these patterns in terms of source and sink models derived from prescribed plate velocities (Bercovici 1995).

The Bayesian formulation is a practical tool to quantify and account for the theoretical errors introduced by the parameterization choice and the surrogate model. Statistics of these errors can be studied by comparing responses obtained with the true forward model and the surrogate model. If the distribution of residuals is approximated as a normal distribution, theoretical errors can be accounted for in the likelihood function (Hansen et al. 2014b). However, the size of the residual vector may not be large enough to properly represent the statistics of errors. Here instead, we used a maximum likelihood estimation (MLE) to implicitly account for these theoretical errors (Dettmer et al. 2007).

**The data.** In this work, we assume that the measurement errors in the data are uncorrelated. In reality however, surface wave dispersion measurements are inherently smooth, and correlated both in space and in frequency. A simple improvement when modeling noise can be made by introducing a function that varies with period while still maintaining the assumption of uncorrelated errors, as in the work of Ravenna & Lebedev (2017). One may proceed a step further by constructing a covariance matrix of data noise, more importantly when working on highly spatially-correlated data sets.

It is also worth mentioning that the method is not limited to the use of a single data-type (*i.e.*, surface wave measurements) to effectively constrain the patterns of upper mantle deformation. This calls for the inclusion of other data-types such as gravity anomalies, surface topography, and/or surface velocities in a joint or separate approach. Such strategies have already been successfully implemented to invert for the 3D density structure of the mantle. (Ricard & Wuming 1991).

## 4.2 Physical Assumptions

The trade-off between physical complexity and computational cost is evident in every geophysical problem considered. In this work, we chose to decrease the computational cost to massively explore the parameter space (using an inverse problem formulation) but at the price of using simplified physical assumptions.

**Nature of the flow model.** We assumed that the flow is in steady-state in order to trace the flow streamlines, which is a prerequisite to compute CPO anisotropy. However, this may not be the case in regions where flow appears to be time-dependent such as migrating trenches and mid ocean ridges (Heuret & Lallemand 2005; Masalu 2007). A time-dependent flow could be implemented by accounting for the evolution of the surface tectonics (Ricard et al. 1993) and the retrodiction of internal heterogeneities (Bunge et al. 2003; Steinberger et al. 2004). Nevertheless, steady-state assumption is still valid in some places such as intra-oceanic regions where flow has been observed to be in steady-state over the last 40 Myr (Becker et al. 2003, 2006).

Another limiting factor is the imposition of arbitrary boundary conditions on the sides of the model domain which strongly impact the nature of the flow. Note that the boundary conditions could be treated as an unknown parameter to be inverted for. An obvious way to address this issue is also to work at the global scale. In this case, a fast and reliable method to compute geodynamic flow in a spherical Earth is indispensable. To cite an example, semi-

analytical circulation models such as that of Hager & O’Connell (1981) can be computed from simple density distributions assuming no lateral variations in viscosity. However, the latter may not render a reasonable assumption within the context of geodynamic tomography since lateral viscosity variations affect the flow significantly, and thus may also strongly influence the resulting anisotropy.

In the context of inverse modeling, the inclusion of lateral viscosity variations is indeed computationally more challenging. However, it remains attainable by performing these calculations in a coarser grid to obtain the general pattern of the flow. This step can be followed by interpolating the coarse grid solution on a finer grid prior to the computation of CPO. Using iterative approaches to flow calculations, another practical approach is to degrade the accuracy of the solution should convergence be an impediment. When cast in a Bayesian formulation, the modeling error due to the approximation of the flow can be accounted for in the inversion process, similar to how the errors due to the ANN were dealt with (see Section 2.3.2). Consequently, texture evolution modeling at the global scale could reasonably be achieved from flows of this nature. The availability of global surface wave maps on the other end should thus make geodynamic tomography feasible at the global scale.

**Composition of the mantle.** Here, we assumed that the composition of the mantle to be olivine, with an A-type crystal fabric, corresponding to dry upper mantle conditions. In the real Earth, seismic wave velocities not only depend on temperature and pressure variations, but also on the compositional structure of the minerals. Recently, self-consistent thermodynamic models have already been incorporated in seismic inversion schemes to interpret tomographic images in terms of mantle composition (Ricard et al. 2005; Cammarano et al. 2009). While the bulk properties (*i.e.*, seismic wave speeds) obtained from Gibbs minimization are isotropic, to our knowledge, deformation-induced anisotropy has not yet been formulated cohesively with thermodynamic models, let alone casting it in an inverse problem.

In general, intrinsic anisotropy in the upper mantle results from complex deformation processes, which depend on a plethora of physical parameters that may be linked to one another. Unlike conventional tomographic techniques, the elastic structure recovered in our scheme directly depends on the assumptions made on these upper-mantle processes. As an example, one would expect that the inclusion of enstatite in our models would dilute the overall amplitude of anisotropy in surface waves. In addition, inversion results depend on control parameters for CPO modeling such as the choice of the slip systems of olivine. For the moment, the value of these parameters have been chosen *ad hoc*, using current available

knowledge mostly originating from laboratory experiments, and thus can be viewed as prior (regularization). Ultimately, the flexibility of Bayesian inference would allow us to treat these parameters as unknown parameters to be inverted for in geodynamic tomography.

## 5 CONCLUSION

We have laid the groundwork for geodynamic tomography, a novel approach that involves constraints from geodynamic modeling to invert seismic surface waves. Imposing these geodynamic constraints reduces the number of model parameters to a single scalar field (*i.e.*, temperature) and one scalar variable (*i.e.*, activation energy for viscosity). The inverse problem is cast using Bayesian inference where we directly sample the model space using Markov chain Monte Carlo algorithm. Here, instantaneous flow, deformation history, and finally seismic anisotropy are computed in our forward problem. The model space is reduced further by parameterizing the temperature field as a sum of spherical temperature anomalies with variable position, size, and temperature.

We tested geodynamic tomography in simple cases, where we successfully recovered synthetic 3D temperature fields, by jointly inverting fundamental mode anisotropic Rayleigh wave and isotropic Love wave phase velocities. In the process, we are also able to constrain the complete deformation pattern, to provide a quantitative interpretation of seismic anisotropy in the mantle. Given the Bayesian formulation, one may express the ensemble of temperature models, and any implicitly computed variables (such as deformation or anisotropy) as posterior probability distributions, and quantify their associated uncertainties. Geodynamic tomography is therefore a potentially powerful technique to study the structure of the upper mantle, and interpret seismic observations in terms of mantle deformation patterns.

## ACKNOWLEDGMENTS

We thank Yanick Ricard for his valuable comments on the manuscript. This work was funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 716542.

## DATA AVAILABILITY

The code underlying the inversion scheme, the instantaneous flow computation, as well as the neural networks will be shared on reasonable request to the authors.

## REFERENCES

- Adam, J. M.-C. & Lebedev, S., 2012. Azimuthal anisotropy beneath southern africa from very broad-band surface-wave dispersion measurements, *Geophysical Journal International*, **191**(1), 155–174.
- Albers, M., 2000. A local mesh refinement multigrid method for 3-d convection problems with strongly variable viscosity, *Journal of Computational Physics*, **160**(1), 126–150.
- Babuska, V. & Cara, M., 1991. *Seismic anisotropy in the Earth*, vol. 10, Springer Science & Business Media.
- Baumann, T. S., Kaus, B. J., & Popov, A. A., 2014. Constraining effective rheology through parallel joint geodynamic inversion, *Tectonophysics*, **631**, 197–211.
- Becker, T. W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochemistry, Geophysics, Geosystems*, **3**(1).
- Becker, T. W., Kellogg, J. B., Ekström, G., & O’Connell, R. J., 2003. Comparison of azimuthal seismic anisotropy from surface waves and finite strain from global mantle-circulation models, *Geophysical Journal International*, **155**(2), 696–714.
- Becker, T. W., Chevrot, S., Schulte-Pelkum, V., & Blackman, D. K., 2006. Statistical properties of seismic anisotropy predicted by upper mantle geodynamic models, *Journal of Geophysical Research: Solid Earth*, **111**(B8).
- Becker, T. W., Kustowski, B., & Ekström, G., 2008. Radial seismic anisotropy as a constraint for upper mantle rheology, *Earth and Planetary Science Letters*, **267**(1-2), 213–227.
- Becker, T. W., Conrad, C. P., Schaeffer, A. J., & Lebedev, S., 2014. Origin of azimuthal seismic anisotropy in oceanic plates and mantle, *Earth and Planetary Science Letters*, **401**, 236–250.
- 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.
- Bercovici, D., 1995. A source-sink model of the generation of plate tectonics from non-newtonian mantle flow, *Journal of Geophysical Research: Solid Earth*, **100**(B2), 2013–2030.
- Bishop, C. M. et al., 1995. *Neural networks for pattern recognition*, Oxford university press.
- Bodin, T., Leiva, J., Romanowicz, B., Maupin, V., & Yuan, H., 2016. Imaging anisotropic layering with bayesian inversion of multiple data types, *Geophysical Journal International*, **206**(1), 605–629.
- Box, G. E. & Tiao, G. C., 2011. *Bayesian inference in statistical analysis*, vol. 40, John Wiley & Sons.
- Brandt, A., 1982. Guide to multigrid development, *Lecture Notes in Mathematics*, **960**, 220–312.
- Bunge, H.-P., Hagelberg, C., & Travis, B., 2003. Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography, *Geophysical Journal International*, **152**(2), 280–301.
- Burgos, G., Montagner, J.-P., Beucler, E., Capdeville, Y., Mocquet, A., & Drilleau, M., 2014. Oceanic lithosphere-asthenosphere boundary from surface wave dispersion data, *Journal of Geophysical Research: Solid Earth*, **119**(2), 1079–1093.



- Bystricky, M., Kunze, K., Burlini, L., & Burg, J.-P., 2000. High shear strain of olivine aggregates: Rheological and seismic consequences, *Science*, **290**(5496), 1564–1567.
- Cammarano, F., Romanowicz, B., Stixrude, L., Lithgow-Bertelloni, C., & Xu, W., 2009. Inferring the thermochemical structure of the upper mantle from seismic data, *Geophysical Journal International*, **179**(2), 1169–1185.
- Chang, S.-J., Ferreira, A. M., 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., Ritsema, J., Heijst, H. J., & Woodhouse, J. H., 2015. Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations, *Journal of Geophysical Research: Solid Earth*, **120**(6), 4278–4300.
- Connolly, J., 2009. The geodynamic equation of state: what and how, *Geochemistry, Geophysics, Geosystems*, **10**(10).
- Connolly, J. A., 2005. Computation of phase equilibria by linear programming: a tool for geodynamic modeling and its application to subduction zone decarbonation, *Earth and Planetary Science Letters*, **236**(1-2), 524–541.
- Conway, D., Alexander, B., King, M., Heinson, G., & Kee, Y., 2019. Inverting magnetotelluric responses in a three-dimensional earth using fast forward approximations based on artificial neural networks, *Computers & Geosciences*, **127**, 44–52.
- Debayle, E., Kennett, B., & Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of australia, *Nature*, **433**(7025), 509.
- Deschamps, F., Lebedev, S., Meier, T., & Trampert, J., 2008. Azimuthal anisotropy of rayleigh-wave phase velocities in the east-central united states, *Geophysical Journal International*, **173**(3), 827–843.
- Dettmer, J., Dosso, S. E., & Holland, C. W., 2007. Uncertainty estimation in seismo-acoustic reflection travel time inversion, *The Journal of the Acoustical Society of America*, **122**(1), 161–176.
- Ekström, G., 2011. A global model of love and rayleigh surface wave dispersion and anisotropy, 25-250 s, *Geophysical Journal International*, **187**(3), 1668–1686.
- Estey, L. H. & Douglas, B. J., 1986. Upper mantle anisotropy: a preliminary model, *Journal of Geophysical Research: Solid Earth*, **91**(B11), 11393–11406.
- Faccenda, M. & Capitanio, F., 2013. Seismic anisotropy around subduction zones: Insights from three-dimensional modeling of upper mantle deformation and sks splitting calculations, *Geochemistry, Geophysics, Geosystems*, **14**(1), 243–262.
- Ferreira, A. M., Faccenda, M., Sturgeon, W., Chang, S.-J., & Schardong, L., 2019. Ubiquitous lower-mantle anisotropy beneath subduction zones, *Nature Geoscience*, **12**(4), 301–306.
- Gallego, A., Ito, G., & Dunn, R., 2013. Investigating seismic anisotropy beneath the reykjanes ridge using models of mantle flow, crystallographic evolution, and surface wave propagation, *Geochemistry, Geophysics, Geosystems*, **14**(1), 243–262.

- Geophysics, Geosystems*, **14**(8), 3250–3267.
- Gerya, T. V., 2010. *Introduction to numerical geodynamic modeling*, Cambridge University Press.
- Hager, B. H. & O’Connell, R. J., 1981. A simple global model of plate dynamics and mantle convection, *Journal of Geophysical Research: Solid Earth*, **86**(B6), 4843–4867.
- Hall, C. E., Fischer, K. M., Parmentier, E., & Blackman, D. K., 2000. The influence of plate motions on three-dimensional back arc mantle flow and shear wave splitting, *Journal of Geophysical Research: Solid Earth*, **105**(B12), 28009–28033.
- Hansen, L. N., Zhao, Y.-H., Zimmerman, M. E., & Kohlstedt, D. L., 2014a. Protracted fabric evolution in olivine: Implications for the relationship among strain, crystallographic fabric, and seismic anisotropy, *Earth and Planetary Science Letters*, **387**, 157–168.
- Hansen, T. M. & Cordua, K. S., 2017. Efficient monte carlo sampling of inverse problems using a neural network-based forward—applied to gpr crosshole travelttime inversion, *Geophysical Journal International*, **211**(3), 1524–1533.
- Hansen, T. M., Cordua, K. S., Jacobsen, B. H., & Mosegaard, K., 2014b. Accounting for imperfect forward modeling in geophysical inverse problems—exemplified for crosshole tomography, *Geophysics*, **79**(3), H1–H21.
- Hastings, W., 1970. Monte carlo sampling methods using markov chains and their applications, *Biometrika*, **57**(1), 97–109.
- Hedjazian, N. & Kaminski, E., 2014. Defining a proxy for the interpretation of seismic anisotropy in non-newtonian mantle flows, *Geophysical Research Letters*, **41**(20), 7065–7072.
- Heuret, A. & Lallemand, S., 2005. Plate motions, slab dynamics and back-arc deformation, *Physics of the Earth and Planetary Interiors*, **149**(1-2), 31–51.
- Hulbert, C., Rouet-Leduc, B., Johnson, P. A., Ren, C. X., Rivière, J., Bolton, D. C., & Marone, C., 2019. Similarity of fast and slow earthquakes illuminated by machine learning, *Nature Geoscience*, **12**(1), 69–74.
- Kaminski, É. & Ribe, N. M., 2002. Timescales for the evolution of seismic anisotropy in mantle flow, *Geochemistry, Geophysics, Geosystems*, **3**(8), 1–17.
- Kaminski, E., Ribe, N. M., & Browaeys, J. T., 2004. D-rer, a program for calculation of seismic anisotropy due to crystal lattice preferred orientation in the convective upper mantle, *Geophysical Journal International*, **158**(2), 744–752.
- Käüfl, P., Valentine, A. P., O’Toole, T. B., & Trampert, J., 2014. A framework for fast probabilistic centroid-moment-tensor determination—inversion of regional static displacement measurements, *Geophysical Journal International*, **196**(3), 1676–1693.
- Köpke, C., Irving, J., & Elsheikh, A. H., 2018. Accounting for model error in bayesian solutions to hydrogeophysical inverse problems using a local basis approach, *Advances in water resources*, **116**, 195–207.
- Lassak, T. M., Fouch, M. J., Hall, C. E., & Kaminski, É., 2006. Seismic characterization of mantle

- flow in subduction systems: Can we resolve a hydrated mantle wedge?, *Earth and Planetary Science Letters*, **243**(3-4), 632–649.
- Lebedev, S., Meier, T., & van der Hilst, R. D., 2006. Asthenospheric flow and origin of volcanism in the baikal rift area, *Earth and Planetary Science Letters*, **249**(3-4), 415–424.
- LeCun, Y., Bengio, Y., & Hinton, G., 2015. Deep learning, *nature*, **521**(7553), 436.
- Mainprice, D., 1990. A fortran program to calculate seismic anisotropy from the lattice preferred orientation of minerals, *Computers & Geosciences*, **16**(3), 385–393.
- Masalu, D. C., 2007. Mapping absolute migration of global mid-ocean ridges since 80 ma to present, *Earth, planets and space*, **59**(9), 1061–1066.
- McKenzie, D., 1979. Finite deformation during fluid flow, *Geophysical Journal International*, **58**(3), 689–715.
- Meier, U., Curtis, A., & Trampert, J., 2007. Global crustal thickness from neural network inversion of surface wave data, *Geophysical Journal International*, **169**(2), 706–722.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E., 1953. Equation of state calculations by fast computing machines, *The journal of chemical physics*, **21**(6), 1087–1092.
- Miller, M. S. & Becker, T. W., 2012. Mantle flow deflected by interactions between subducted slabs and cratonic keels, *Nature Geoscience*, **5**(10), 726.
- Moghadas, D., Behroozmand, A. A., & Christiansen, A. V., 2020. Soil electrical conductivity imaging using a neural network-based forward solver: Applied to large-scale bayesian electromagnetic inversion, *Journal of Applied Geophysics*, p. 104012.
- Montagner, J. & Nataf, H., 1988. Vectorial tomography. part i: Theory, *Geophysical Journal International*, **94**, 295–307.
- Montagner, J.-P., 1994. Can seismology tell us anything about convection in the mantle?, *Reviews of Geophysics*, **32**(2), 115–137.
- Montagner, J.-P. & Anderson, D. L., 1989. Petrological constraints on seismic anisotropy, *Physics of the earth and planetary interiors*, **54**(1-2), 82–105.
- Montagner, J.-P. & Jobert, N., 1988. Vectorial tomography—ii. application to the indian ocean, *Geophysical Journal International*, **94**(2), 309–344.
- Montagner, J.-P. & Nataf, H.-C., 1986. A simple method for inverting the azimuthal anisotropy of surface waves, *Journal of Geophysical Research: Solid Earth*, **91**(B1), 511–520.
- Montagner, J.-P. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *Journal of Geophysical Research: Solid Earth*, **95**(B4), 4797–4819.
- Mosegaard, K. & Sambridge, M., 2002. Monte carlo analysis of inverse problems, *Inverse problems*, **18**(3), R29.
- Mosegaard, K. & Tarantola, A., 1995. Monte carlo sampling of solutions to inverse problems, *Journal of Geophysical Research: Solid Earth*, **100**(B7), 12431–12447.
- Nettles, M. & Dziewoński, A. M., 2008. Radially anisotropic shear velocity structure of the upper

- mantle globally and beneath north america, *Journal of Geophysical Research: Solid Earth*, **113**(B2).
- Nicolas, A. & Christensen, N. I., 1987. Formation of anisotropy in upper mantle peridotites—a review, *Composition, structure and dynamics of the lithosphere-asthenosphere system*, **16**, 111–123.
- Panning, M. P. & Nolet, G., 2008. Surface wave tomography for azimuthal anisotropy in a strongly reduced parameter space, *Geophysical Journal International*, **174**(2), 629–648.
- Patankar, S. V., 1980. *Numerical Heat Transfer and Fluid Flow*, Hemisphere Publishing Corporation, New York.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E., 2011. Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research*, **12**, 2825–2830.
- Plomerová, J., Kouba, D., & Babuška, V., 2002. Mapping the lithosphere–asthenosphere boundary through changes in surface-wave anisotropy, *Tectonophysics*, **358**(1-4), 175–185.
- Ravenna, M. & Lebedev, S., 2017. Bayesian inversion of surface-wave data for radial and azimuthal shear-wave anisotropy, with applications to central mongolia and west-central italy, *Geophysical Journal International*, **213**(1), 278–300.
- Ribe, N. M., 1989. Seismic anisotropy and mantle flow, *Journal of Geophysical Research: Solid Earth*, **94**(B4), 4213–4223.
- Ribe, N. M., 1992. On the relation between seismic anisotropy and finite strain, *Journal of Geophysical Research: Solid Earth*, **97**(B6), 8737–8747.
- Ricard, Y. & Wuming, B., 1991. Inferring the viscosity and the 3-d density structure of the mantle from geoid, topography and plate velocities, *Geophysical Journal International*, **105**(3), 561–571.
- Ricard, Y., Richards, M., Lithgow-Bertelloni, C., & Le Stunff, Y., 1993. A geodynamic model of mantle density heterogeneity, *Journal of Geophysical Research: Solid Earth*, **98**(B12), 21895–21909.
- Ricard, Y., Mattern, E., & Matas, J., 2005. Mineral physics in thermo-chemical mantle models, *Hilst, R., Bass, JD, Matas, Trampert, J., (Eds.), Com-position, Structure and Evolution of the Earth Mantle, AGU monograph*, **160**, 283–300.
- Ritzwoller, M. H., Shapiro, N. M., Barmin, M. P., & Levshin, A. L., 2002. Global surface wave diffraction tomography, *Journal of Geophysical Research: Solid Earth*, **107**(B12), ESE–4.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J., 1985. Learning internal representations by error propagation, Tech. rep., California Univ San Diego La Jolla Inst for Cognitive Science.
- Samuel, H., 2012. Time-domain parallelization for computational geodynamics, *G-cubed*, p. doi:10.1029/2011GC003905.
- Samuel, H., 2018. A deformable particle-in-cell method for advective transport in geodynamic modelling, *gji*, **214**, 1744–1773, doi:10.1093/gji/ggy231.
- Shahnas, M., Yuen, D., & Pysklywec, R., 2018. Inverse problems in geodynamics using machine learning algorithms, *Journal of Geophysical Research: Solid Earth*, **123**(1), 296–310.

- Shapiro, N. & Ritzwoller, M., 2002. Monte-carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophysical Journal International*, **151**(1), 88–105.
- Shen, W., Ritzwoller, M. H., Schulte-Pelkum, V., & Lin, F.-C., 2012. Joint inversion of surface wave dispersion and receiver functions: a bayesian monte-carlo approach, *Geophysical Journal International*, **192**(2), 807–836.
- Sieminski, A., Liu, Q., Trampert, J., & Tromp, J., 2007. Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods, *Geophysical Journal International*, **168**(3), 1153–1174.
- Smith, A. F. M., 1991. Bayesian computational methods, *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, **337**(1647), 369–386.
- Smith, M. L. & Dahlen, F., 1973. The azimuthal dependence of love and rayleigh wave propagation in a slightly anisotropic medium, *Journal of Geophysical Research*, **78**(17), 3321–3333.
- Steinberger, B., Sutherland, R., & O’connell, R. J., 2004. Prediction of emperor-hawaii seamount locations from a revised model of global plate motion and mantle flow, *Nature*, **430**(6996), 167–173.
- Stixrude, L. & Lithgow-Bertelloni, C., 2011. Thermodynamics of mantle minerals-ii. phase equilibria, *Geophysical Journal International*, **184**(3), 1180–1213.
- Sturgeon, W., Ferreira, A. M., Faccenda, M., Chang, S.-J., & Schardong, L., 2019. On the origin of radial anisotropy near subducted slabs in the midmantle, *Geochemistry, Geophysics, Geosystems*, **20**(11), 5105–5125.
- Takeuchi, H. & Saito, M., 1972. Seismic surface waves, *Methods in computational physics*, **11**, 217–295.
- Tommasi, A., Mainprice, D., Canova, G., & Chastel, Y., 2000. Viscoplastic self-consistent and equilibrium-based modeling of olivine lattice preferred orientations: Implications for the upper mantle seismic anisotropy, *Journal of Geophysical Research: Solid Earth*, **105**(B4), 7893–7908.
- 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, *Geophysical Journal International*, **201**(2), 996–1021.
- Xie, J., Ritzwoller, M. H., Shen, W., & Wang, W., 2017. Crustal anisotropy across eastern tibet and surroundings modeled as a depth-dependent tilted hexagonally symmetric medium, *Geophysical Journal International*, **209**(1), 466–491.
- Xu, H. & Beghein, C., 2019. Measuring higher-mode surface wave dispersion using a transdimensional bayesian approach, *Geophysical Journal International*.
- Yuan, K. & Beghein, C., 2013. Seismic anisotropy changes across upper mantle phase transitions, *Earth and Planetary Science Letters*, **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, *Journal of Geophysical Research: Solid Earth*, **119**(4), 3232–3255.
- Zhang, S. & Karato, S.-i., 1995. Lattice preferred orientation of olivine aggregates deformed in simple

shear, *Nature*, **375**(6534), 774.

## APPENDIX A: PARAMETERIZING TEMPERATURE WITH SPHERICAL ANOMALIES

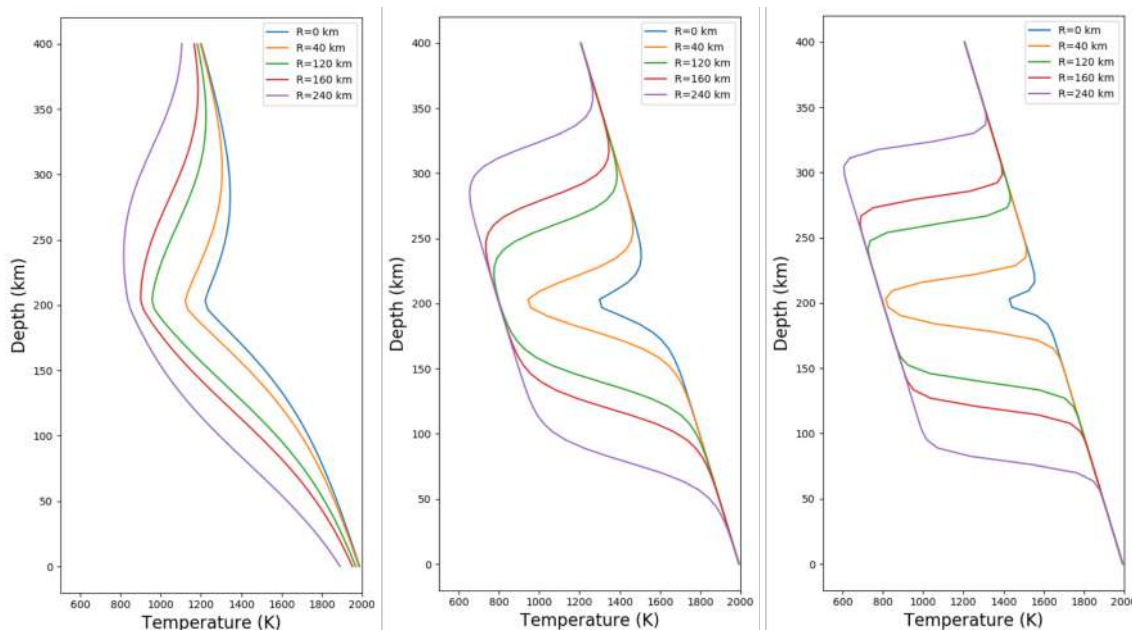
For a given anomaly, we define a basis function corresponding to that anomaly using eq. (3). The negative sign indicates that the anomaly is colder than the background temperature if  $T_c$  is positive (a negatively buoyant anomaly). Should  $T_c$  be negative, then the anomaly adds up with the background temperature resulting to a positively buoyant anomaly. The function is designed such that: (1) When  $\left\| \mathbf{r} - \frac{\mathbf{r}_0}{L_s} \right\| > \frac{R}{L_s}$  and  $\tanh$  returns a value of nearly one, then the temperature is just the background temperature. (2) When  $\left\| \mathbf{r} - \frac{\mathbf{r}_0}{L_s} \right\| = \frac{R}{L_s}$ , then the temperature at just half of the radius of the anomaly is equal to  $T_{\text{background}} - \frac{T_c}{2}$ . (3) Finally, when  $\left\| \mathbf{r} - \frac{\mathbf{r}_0}{L_s} \right\| < \frac{R}{L_s}$  and  $\tanh$  returns a value of minus one, this corresponds to the temperature at the center of the anomaly  $T_{\text{background}} - T_c$ .

Here,  $\beta$  controls the sharpness of the temperature gradient and is held at a fixed value. Choosing a very large value for  $\beta$  results in a sharp temperature gradient (see Fig. A1). In addition, opting for a smooth function such as hyperbolic tangent avoids very sharp viscosity contrasts when computing for the flow. The advantage of building a basis set is to reduce the number of model parameters. In conventional inversion schemes of scalar fields, we usually invert for a scalar at a given grid point. Hence, the number of model parameters depends on the grid size. In a cube, this would result to  $N^3$  model parameters to constrain, where  $N^3$  is the size of the 3D block. In our case, this gives us  $5M$  parameters to be inverted, where  $M$  is the number of spherical anomalies. Finally, we define the 3D scalar temperature field as the sum of the background temperature and the spherical anomalies as shown in eq. (1).

## APPENDIX B: A NEURAL NETWORK-BASED APPROXIMATION TO D-REX

In this work, we use an ANN as a surrogate model  $g_{\text{nn}}$ , to approximate the forward operator for texture evolution  $g_{\text{CPO}}$ . We consider a simple architecture of feedforward neural network called a multi-layer perceptron (MLP) with two hidden layers similar to the work of LeCun et al. (2015) defined by:

$$g_{\text{nn}}(X_l) = \hat{Y}_l = a_1 \left( \sum_{k=1}^{N_{h1}} w_{kl}^3 a_2 \left( \sum_{j=1}^{N_{h2}} w_{jk}^2 a_3 \left( \sum_{i=1}^{N_x} w_{ij}^1 X_i \right) \right) \right). \quad (\text{B.1})$$



**Figure A1.** 1D temperature profiles with depth for different values of  $R$  and  $\beta$ . Left:  $\beta = 5$ . Middle:  $\beta = 20$ . Right:  $\beta = 50$ . Here, we consider a spherical anomaly with  $T_c = 800$  K located at the center of the 3D volume. The plots refer to 1D depth profiles of temperature through the middle of the sphere at specified values of  $R$  and  $\beta$ . The  $x$  and  $y$  axes correspond to temperature and depth, respectively. Based on our parameterization, increasing the value of  $R$  at constant  $\beta$  increases the size of the temperature anomaly. At constant  $R$ , the anomalies retain their respective sizes but the temperature gradient becomes sharper at increasing  $\beta$ . Thus, choosing an appropriate  $\beta$  is important so as to avoid sharp viscosity contrasts (since  $\eta$  depends on  $T$ ) when computing flow. In our inversion, we choose to fix  $\beta = 20$ , and invert for  $R$ .

The output  $\hat{Y}_l$  of the MLP is an estimate of the 21 independent coefficients of the stiffness tensor where  $l$  is the index pertaining to one element in the tensor.  $N_{h1}$  and  $N_{h2}$  are the sizes of the two hidden layers considered, and  $N_x$  is the size of the input vector. We design the network such that the input  $X$  contains the deformation history along a flow streamline. The streamline is divided into 200 time steps. Each step contains one  $L_{ij}$  matrix and one corresponding  $dt$ . Thus, each step has 10 independent components as inputs. The number of inputs in the neural network first layer is  $N_x = 2000$  (see eq. (B.1)). The functions  $a_1, a_2$ , and  $a_3$  are known as activation functions whose purpose are to introduce non-linearity to the output of one neuron and to constrain its output to a desired range and distribution. Here, we choose them as default rectified linear unit functions to allow faster convergence (Pedregosa et al. 2011). Lastly, the  $w$ 's refer to the weights which reflect the significance of a given neuron.

To build a suitable surrogate model to D-Rex, the weights  $w^1$ ,  $w^2$ , and  $w^3$  have to be adjusted to the proper value. This is performed by minimizing a loss function which is the

difference between the training outputs  $g_{\text{CPO}}(X)$  and the output of the network itself  $g_{\text{nn}}(X)$  using a stochastic gradient descent algorithm (Rumelhart et al. 1985). Formally, the loss function is a squared  $L_2$  norm and takes the form:

$$\text{Loss}(Y, \hat{Y}, w) = \frac{1}{2} \|Y - \hat{Y}\|_2^2 + \frac{\lambda}{2} \|w\|_2^2. \quad (\text{B.2})$$

The second term constrains the weights to avoid data over-fitting, where  $\alpha$  is a regularization parameter that quantifies the degree of penalization. The weights are updated iteratively by subtracting its current value from the gradient of the loss function with respect to the weights:

$$w_{i+1} = w_i - \epsilon \nabla \text{loss}_i, \quad (\text{B.3})$$

where  $\epsilon$  is the learning rate which controls the step-size for updating the weights, and  $i$  is the iteration step. The training achieves convergence when the tolerance value *tol* for the loss function is reached. However, the algorithm may also be stopped once the maximum number of iterations is reached.

The network is trained by considering 30 flow models, each comprising  $M$  spherical anomalies to drive thermal convection. Each sphere has a random position and size, and can either be positively or negatively buoyant. This is to ensure that each flow path we define is unique enough so that the network can learn a variety of input-output combinations. Here, we acknowledge that the choice of flow models is not enough to be able to predict seismic anisotropy in the most general case. However, in this work, we only attempt to predict anisotropy for a small class of flow models (convection due to a collection of spherical temperature anomalies). Since only such classes of models are tested, we can restrict ourselves to this type of model when training the network.

One training input corresponds to one deformation history along a streamline whereas one training output corresponds to one stiffness matrix computed with D-Rex. The training set can be represented as a matrix containing the stiffness coefficients and the input parameters given by  $[Y_{l=1,21}, X_{i=1,N_x}]_{n=1, N_{\text{train}}}$  where  $N_{\text{train}}$  is the number of training sets. Thus, the training inputs are of the size  $[2000, N_{\text{train}}]$  and the training outputs are of the size  $[21, N_{\text{train}}]$ . In this problem,  $16^3$  input-output combinations for each 3D flow model are used to train the network. In total, there are  $M = 1.2288 \times 10^5$  training sets for the network to learn from.

We adopt the Python package scikit-learn to train the network (Pedregosa et al. 2011). Table A1 below summarizes the parameters used to design and build the network.

The network is tested by considering a 3D deformation due to a sinking anomaly that is not part of the training input. Table A2 shows the computation times for computing anisotropy



**Table A1.** Neural network parameters.

$N_{train}$	$N_x$	$N_y$	$N_{h1}$	$N_{h2}$	$\lambda$	$\epsilon$	$tol$	Max iterations
$1.2288 \times 10^5$	2000	21	100	50	0.1	$1.0 \times 10^{-3}$	$1.0 \times 10^{-4}$	1000

from both D-Rex and neural networks. The relative speed-up of using neural networks is over three orders of magnitude compared to performing texture evolution calculations with D-Rex. For reference, we also give the computation times for network training, flow modeling, as well as for surface wave dispersion curves calculations. Each routine in the forward problem has been executed in a serial fashion for the sake of comparison.

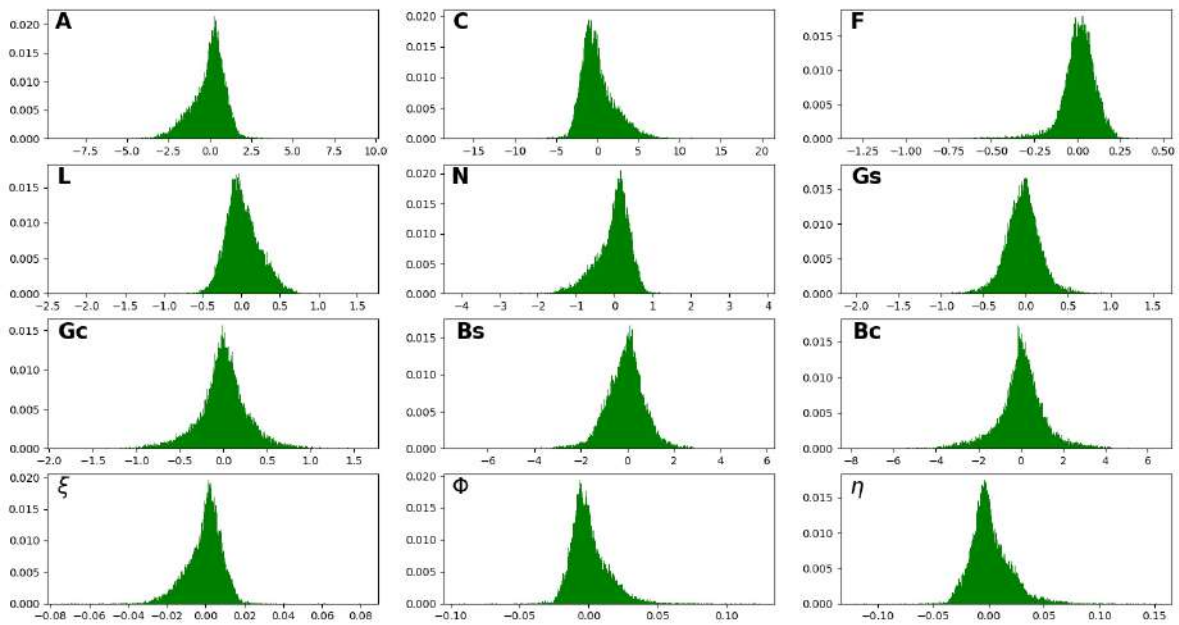
The elastic tensor computed from  $g_{nn}$  is projected into both a VTI medium, thus having elastic parameters  $A$ ,  $C$ ,  $F$ ,  $L$ , &  $N$ , and radial anisotropy strength  $\xi$ ,  $\phi$ , &  $\eta$ ; and an HTI medium, with parameters  $G_s$ ,  $G_c$ ,  $B_s$ , &  $B_c$ . Aside from plotting the percentage of total anisotropy (as in Section 2.2.3), we compare the results further with D-Rex by plotting 1-D marginal distributions of the residuals of each seismic parameter. Each parameter contains a small bias very close to zero which is attributed to the minimization of the  $L_2$  loss function.

### APPENDIX C: A SIMPLE TEST FOR CONVERGENCE

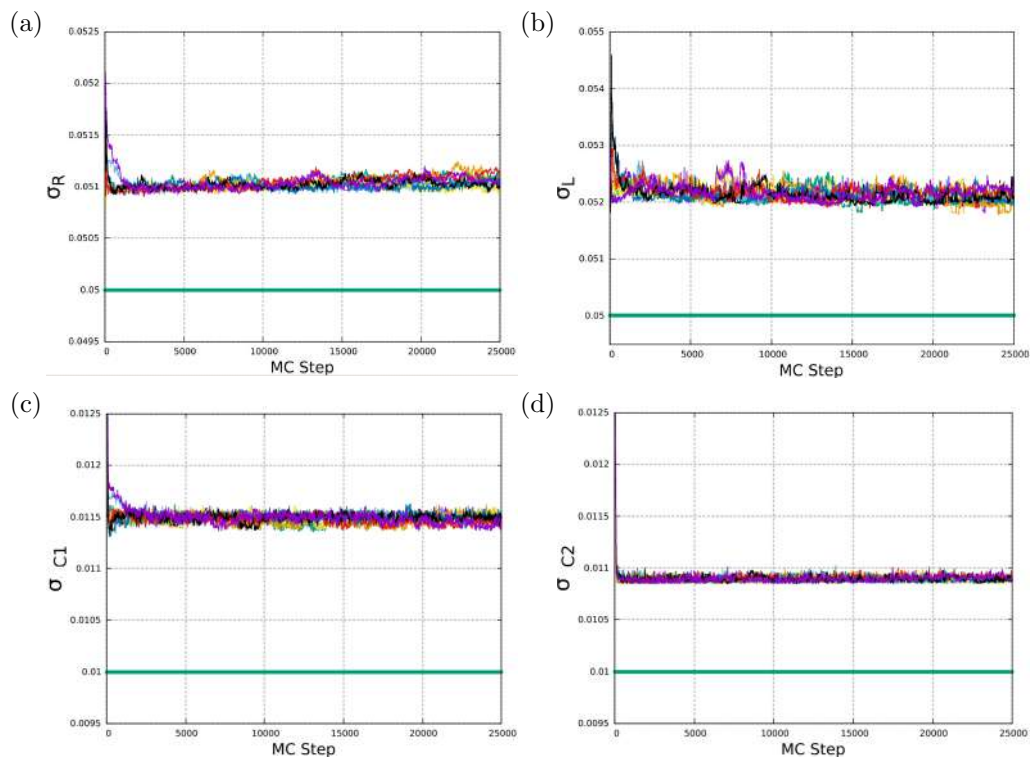
Fig. A3 shows the noise estimate plotted against MC step in the one sphere case. The standard deviation of data noise is implicitly computed with MLE (see Section 2.3.2), and is simply given by the level of data fit. The starting point for each plot is the iteration at which anisotropic tomography commences. The trends exhibit well-mixed random walk behaviors indicating that convergence has been achieved. This level of noise estimated by MLE represents the combination of observational errors (white noise added to the data), and theoretical errors (errors of the surrogate model used for texture evolution).

**Table A2.** Computation times for each subroutine in the forward model.

Routine	D-Rex	ANN	Flow	Dispersion	Training
Time (s)	73919.83	21.55	6.6	119.63	603.85



**Figure A2.** 1-D marginal distribution of the difference between  $g_{\text{CPO}}(X)$  and  $g_{\text{nn}}(X)$  in terms of the VTI and HTI-projected elastic tensor.



**Figure A3.** Noise estimate with MC step for (a) Rayleigh waves, (b) Love waves, (c)  $c_1$ , and (d)  $c_2$ . Each colored line plot is associated with one independent Markov chain. Solid green line indicates the standard deviation of random errors added to the data.