# SItomo – A toolbox for splitting intensity tomography and application in the Eastern Alps

Frederik Link [*], Maureen D. Long

*Department of Earth and Planetary Sciences, Yale University, New Haven, CT, United States*

A B S T R A C T

The tomographic inversion of shear wave splitting data for upper mantle anisotropy has been a longstanding challenge. This is due to the ray-based approximation of classical approaches and the near-vertical incidence of the core-mantle converted phases such as SKS that are often used. Recent developments include the calculation of finite-frequency sensitivity kernels for SKS splitting intensity observations, which allows us to accurately take into account the sensitivity to anisotropic structure with depth. A requirement of this tomographic technique is a dense station spacing, which results in overlapping sensitivity kernels at depth and allows for the localization of anisotropic structure. This is satisfied by a growing number of temporary seismic deployments, which motivates the desire to image anisotropic complexities with depth. Here, we introduce and make available a toolbox for the MATLAB environment that facilitates the application of finite-frequency splitting intensity tomography to dense seismic arrays. Our implementation includes several key features, including: 1) A forward calculation of splitting intensities and sensitivity kernels for a complex anisotropic model space. 2) Consideration of the dominant period of the wave, allowing for multiple-frequency analysis, as well as the incoming wave's non-vertical incidence. 3) The inversion can be based on a classical gradient descent, on a form of the conjugate gradient method known as the BFGS algorithm, or on a gradient-informed stochastic reversible jump algorithm, allowing for a data-driven parametrization of the model space. 4) Importing splitting intensity measurements from waveforms processed in SplitRacer allows for fast pre-processing of large data sets due to its fully automatic design. To illustrate our method, we present both synthetic tests and an application to real data. We apply our inversion procedure to data from the Swath-D network, which densely covers the transition of the Central to the Eastern Alps. Previous studies showed evidence for an abrupt lateral change of layered seismic anisotropy that had been attributed to an opening for channeled asthenospheric flow. Using an SKS splitting intensity tomography approach, we can confirm previous inferences while providing additional constraints on the distribution of anisotropy laterally and with depth.

## 1. Introduction

The operation of plate tectonics at the Earth's surface and the corresponding deformation at plate boundaries are closely related to the deformation and flow in the lithospheric and asthenospheric upper mantle (e.g., Long and Becker, 2010). Due to the causative link between deformation and seismic anisotropy, observations of anisotropy provide unique insights into such deformation processes at depth (Savage, 1999). In the upper mantle, intrinsically anisotropic minerals like olivine align due to shear, resulting in a lattice preferred orientation (LPO); the resulting anisotropy can be directly linked to flow in the mantle (e.g., Karato et al. 2008; Skemer and Hansen, 2016). At more

shallow levels, structures such as aligned cracks, isotropic layers, or fluid filled pockets can result in anisotropy (Christensen, 1966; Nur and Simmons, 1969; Nur, 1971; Crampin, 1987; Yousef and Angus, 2016; Backus, 1962; Savage, 1999). Such apparent anisotropy is often referred to as shape preferred orientation (SPO). Anisotropy has several effects on the seismic wavefield. It is characterized by a directional dependence of the seismic P- and S-wave velocities (Silver, 1996; Savage, 1999). In particular, S-waves propagate with different velocities depending on their polarization direction. This results in shear wave splitting, in which two quasi-S-phases polarized roughly perpendicular to each other travel with different wave speeds. The polarization of the fast propagating phase corresponds to the fast splitting direction, and the slow phase

characterizes the corresponding slow direction (Silver and Chan, 1991). The delay time between the fast and slow phase is related to the strength of the anisotropic medium and the distance the phases travelled through it.

Shear wave splitting can be measured with the rotation correlation (Bowman and Ando, 1987) or the energy minimization (Silver and Chan, 1991) techniques, which describe the splitting using the fast direction, $\phi$, and the delay time, $\delta t$. These splitting parameters allow for a general interpretation of the anisotropy that causes the splitting, often assuming a single homogeneous anisotropic layer. The measured fast polarization typically corresponds to the fast axis orientation of the anisotropy and the delay time scales with the strength and thickness of the anisotropic layer. The trade-off between strength and thickness is complete and cannot be resolved with classical techniques. If splitting is measured over a range of backazimuths, the backazimuthal variations in apparent splitting parameters (Silver and Savage, 1994) can be used to resolve multiple anisotropic layers (typically two). The depth and thickness of anisotropic layers can only be estimated qualitatively from Fresnel-zone calculations (Alsina and Snieder, 1995; Rümpker and Ryberg, 2000) or through a spatial coherency approach (Gao and Liu, 2012).

Other strategies allow for the inversion for a tomographic image of the anisotropic structure of the subsurface. The most advanced technique developed in the last decade is full-waveform inversion, which uses the adjoint character of the wave equation to calculate sensitivities of the waveform shape to model parameters in a 2D or 3D model volume (e.g., Tarantola, 1987; Tromp et al., 2005; Fichtner and Trampert, 2011; Zhu et al., 2021). Sensitivity kernels calculated from the adjoint methods can be used iteratively to invert for isotropic as well as anisotropic parameters. However, this technique requires large computational resources, which can limit the number of events and the frequency range used in the inversion.

Another approach involves finite-frequency modelling of splitting intensity measurements, allowing for an analytic solution of the sensitivity kernels (e.g., Favier and Chevrot, 2003; Chevrot, 2006; Long et al. 2008; Mondal and Long, 2019). The splitting intensity (SI) is a single value that expresses the strength of the splitting effect on an individual waveform (Chevrot, 2000). While a single splitting intensity measurement does not allow a direct inference of the anisotropic geometry, multiple measurements can be used to characterize anisotropy in a manner similar to the more classical rotation correlation or energy minimization methods. The calculation of finite-frequency sensitivity kernels for SI observations allows us to take into account the laterally broadened sensitivity for the anisotropic structure with depth (Favier and Chevrot, 2003; Long et al., 2008; Sieminski et al., 2009). Due to the near-vertical incidence of core-mantle converted phases, one requirement of their use in tomographic techniques is a dense station spacing, which results in overlapping sensitivity kernels at depth. This is satisfied by a growing number of temporary deployments of seismic broadband sensors (e.g., Monteiller and Chevrot, 2011; Lin et al., 2014; Mondal and Long, 2020; Huang and Chevrot, 2021).

Here, we introduce a toolbox for the MATLAB environment which facilitates the application of finite-frequency splitting intensity tomography to dense seismic arrays, building on the framework developed by Mondal and Long (2019). We carry out the forward calculation of splitting intensities and sensitivity kernels for a complex anisotropic model space. This is based on the Born approximation for an isotropic background model disturbed by arbitrarily oriented anisotropy with hexagonal symmetry. The code takes the dominant period into consideration, allowing for multiple-frequency analysis, and allows for non-vertical incidence of the incoming wave. Depending on the complexity of the problem, the user can choose to apply the inversion based on a classical gradient descent approach, a gradient descent taking into account an approximation of the Hessian matrix, or a gradient-informed stochastic inversion which includes a reversible jump algorithm, allowing for a data-driven parametrization of the model space. The code also includes the option to import splitting intensities

from waveforms processed in SplitRacer, allowing for the fast download and pre-processing of large data sets due to its fully automatic design (Link et al. 2022).

We illustrate the strengths and weaknesses of the different techniques on synthetic data sets, considering models that include the variation of anisotropic parameters laterally and with depth. To showcase the applicability of the technique, we apply the gradient descent method including the Hessian approximation to the Swath-D dense seismic network (Heit et al. 2017), a temporary broadband seismic installment in the transition from the Central to the Eastern Alps. Previous studies have shown that this area is affected by complex lateral and vertical variation of anisotropy (Link and Rümpker 2021; Qorbani et al. 2015), providing an excellent test data set for our tomographic approach.

## 2. Method

### 2.1. Forward problem and sensitivity kernels

Here, we describe the theoretical and practical underpinnings of our approach, which follows closely upon previous work (Chevrot, 2006; Long et al. 2008), particularly the framework proposed by Mondal and Long (2019). In particular, we extend the framework of Mondal and Long (2019) by including the near- and mid-field terms of the Green's functions, considering more realistic depth dependent contributions of the anisotropic structure to the calculated splitting intensities (Favier et al. 2004). While considering the previous formulation for the upper mantle anisotropy using an arbitrarily oriented effective hexagonal representation of upper mantle olivine LPO, we additionally consider the formulation defining the hexagonal anisotropic tensor based on the anisotropic perturbation of isotropic velocities (Farra et al. 1991). Due to the lack of frequency-dependent splitting intensity measurements, a variation of the dominant frequency of the incident wavelet has not been considered previously. Here, we explicitly point out the significance of the change in shape of the sensitivity kernels with frequency and introduce it as important event parameter in our formulation.

For any tomographic inversion, a forward problem must be formulated first, describing the relationship of the spatially distributed model parameters and the data. Here, we focus on shear wave splitting resulting from an interaction of an incident shear wave with an anisotropic medium. The radial and transverse components of an initially radially polarized split shear wave can be described as

$$u_r(t) = w\left(t + \frac{\delta t}{2}\right)\cos^2\phi + w\left(t - \frac{\delta t}{2}\right)\sin^2\phi,$$

$$u_t(t) = -\frac{1}{2}\left[w\left(t + \frac{\delta t}{2}\right) - w\left(t - \frac{\delta t}{2}\right)\right]\sin(2\phi), \quad (1)$$

where $w(t)$ is the initial waveform of the incoming shear wave before entering the anisotropic medium, $\delta t$ is the delay time between the fast and slow phases after the propagation through an anisotropic medium and $\phi$ is the fast axis direction (Silver and Chan, 1991). For small delay times compared to the dominant period ($T/\Delta t \geq 5$, see Rümpker and Silver, 1998), the radial waveform resembles the initial waveform, while the transverse component is proportional to the time derivative of the initial waveform (and with that to the time derivative of the radial component),

$$u_r(t) \cong w(t),$$

$$u_t(t) \cong -\frac{1}{2}\dot{w}(t)\sin(2\phi), \quad (2)$$

where $\dot{w}(t)$ describes the time derivative of the initial waveform, which can be approximated as the time derivative of the radial component (see also Chevrot, 2000). As introduced by Chevrot (2006), the splitting intensity can be constructed by projecting the transverse component on

the derivative of the radial component, which leads to

$$S = 2 \frac{\text{Re} \int i\omega \, u_t(\omega) \, w(\omega) \, d\omega}{\int \omega^2 \, |w(\omega)|^2 \, d\omega}. \tag{3}$$

where we introduce $u_r(\omega)$ and $u_t(\omega)$ in frequency domain using the Fourier transform with $\omega$ as frequency. For the forward calculation of the splitting intensity it is therefore necessary to formulate the transverse component waveform as a result from an interaction of the incident shear wave with an anisotropic structure. The contribution of an anisotropic perturbation of the elastic tensor to the particle motion can be described as (Chevrot, 2006; Mondal and Long, 2019)

$$\delta u_i = -\int_\Omega \left( \delta c_{ijkl} \, \partial_k w_l \right) \partial_i G_{ij} \, dV \,, \tag{4}$$

where $\delta c_{ijkl}$ are the anisotropic perturbations of the elastic tensor and $G_{ij}$ is the Green's function solving the inhomogeneous partial differential equations (Aki and Richards, 2002),

$$\rho \partial_t^2 G_{ij} = \partial_k \left( c_{ijkl} \, \partial_l G_{kl} \right) \delta_{ij} \, \delta(r_i - r_{S,i}) \, \delta(t - t_S). \tag{5}$$

For a homogeneous isotropic medium the perturbed wave field in Eq. (4) can be written as (Aki and Richards, 2002; Chevrot, 2006)

$$\delta u_n = \int_\Omega \left[ \frac{1}{\omega^2} \frac{15 p_n p_p p_q - 3p_n \, \delta_{pq} - 3p_p \delta_{nq} - 3p_q \delta_{np}}{4\pi \rho r^4} M_{pq} e^{-\frac{i\omega r}{\beta}} \right.$$
$$- \frac{1}{i\omega} \frac{15 p_n p_p p_q - 3p_n \, \delta_{pq} - 3p_p \delta_{nq} - 3p_q \delta_{np}}{4\pi \rho \beta r^3} M_{pq} e^{-\frac{i\omega r}{\beta}}$$
$$- \frac{6 p_n p_p p_q - p_n \, \delta_{pq} - p_p \delta_{nq} - 2p_q \delta_{np}}{4\pi \rho \beta^2 r^2} M_{pq} e^{-\frac{i\omega r}{\beta}}$$
$$\left. - i\omega \frac{p_n p_p - \delta_{np}}{4\pi \rho \beta^3 r} M_{pq} e^{-\frac{i\omega r}{\beta}} \right] dV, \tag{6}$$

where $p_n$ is the unit slowness vector of the outgoing wave, $\delta_{pq}$ is the Kronecker delta, $\beta$ is the isotropic reference shear velocity and $M_{pq}$ is the moment tensor. This can be constructed from the perturbed elasticity and strain tensor (Chevrot, 2006)

$$M_{pq} = \frac{i\omega}{\beta} \widetilde{M}_{pq} w_0 e^{i\omega T}, \tag{7}$$

with

$$\widetilde{M}_{pq} = 2\rho \alpha^2 \epsilon \left( p_k' s_k \right) (g_l s_l) (s_p s_q - \delta_{pq})$$
$$+ \rho \alpha^2 \delta \left( p_k' s_k \right) (g_l s_l) \left( \delta_{pq} - 2 s_p s_q \right)$$
$$+ 2\rho \beta^2 \gamma \left[ 2 \left( p_k' s_k \right) (g_l s_l) \delta_{pq} - (g_k s_k) \left( p_p' s_q + s_p p_q' \right) - \left( p_k' s_k \right) (g_p s_q + s_p g_q) \right], \tag{8}$$

where $s_p$ is the orientation of the symmetry axis of the anisotropic tensor and $p_p'$, $g_p$ are the slowness unit vector and the polarization vector of the incoming wave, respectively. This assumes a transverse isotropic medium formulated using the dimensionless Thomsen parameters (Mensch and Rasolofosaon, 1997)

$$\epsilon = \frac{(C_{11} - C_{33})}{2\rho\alpha^2} \,,$$

$$\gamma = \frac{C_{66} - C_{44}}{2\rho\beta^2} \,,$$

$$\delta = \frac{C_{13} - C_{33} + 2C_{44}}{\rho\alpha^2}, \tag{9}$$

where we use the Voigt notation for the elastic tensor. This allows the definition of the elastic constants independent of the orientation relative to the coordinate system. This formulation can be used for any trans-

verse isotropic material. However, minerals considered to produce anisotropy in the mantle, such as olivine, have mostly orthorhombic symmetry, and require a form of angular averaging (Mondal and Long, 2019) to be approximated with hexagonal symmetry. The Thomsen parameters can then be replaced by the averaged single crystal or aggregate parameters, while the strength of anisotropy can be expressed as the fraction parameter $\chi$ of crystal alignment (Mondal and Long, 2019)

$$C_{ij} = \chi C_{ij,single} + (1 - \chi) C_{ij,Isotropic}, \tag{10}$$

which uses the separation of the elastic tensor into isotropic and hexagonal parts (Browaeys and Chevrot, 2004). There is also the possibility to describe the elastic constants in terms of anisotropic velocity perturbations to the isotropic velocities (Farra et al. 1991)

$$A = \rho v_P \left( 1 - \frac{a}{2} \right)^2, C = \rho v_P \left( 1 + \frac{a}{2} \right)^2,$$

$$L = \rho v_S \left( 1 + \frac{a}{2} \right)^2, N = \rho v_S \left( 1 - \frac{a}{2} \right)^2,$$

$$F = \mu(A - 2L), \mu = 1.03, \tag{11}$$

where $\rho$ denotes the density, $v_P$ is the P-wave velocity, and $v_S$ is the S-wave velocity in the medium. For simplification, the anisotropy for the P-wave and the S-wave propagation is assumed to be equal and described by the variable $a$, which is the ratio of the difference between the maximum and minimum anisotropic velocities divided by its average,

$$a = \frac{v_1 - v_2}{\bar{v}}. \tag{12}$$

We make this simplification because shear wave splitting is not sensitive to differences in P- and S-wave anisotropic fractions. The constants $(A, C, L, N, F)$ fully describe the elastic tensor for hexagonal symmetry as described by Love (1920). To avoid the simplifications involved in this approximation, more realistic elastic tensors based on laboratory experiments and/or field observations may be used, as mentioned previously. Instead of inverting for the Thomsen parameters in addition to the orientation of the elastic tensor, which would require the inversion for 5 parameters, we define a single crystal estimate for the elastic tensor and invert for the fraction paramter $\chi$ to describe the strength of the anisotropic medium, following Mondal and Long (2019). Together with the angles $(\phi, \theta)$ describing the azimuth and dip, respectively, of the symmetry axis, this reduces the inversion to only 3 parameters.

Using the Eqs. (3), (6)–(9) the splitting intensity can be formulated as

$$S = \int_\Omega K_\chi(\theta, \phi) \, \chi \, d^3 r, \tag{13}$$

where $K_\chi(\theta, \phi)$ is equivalent to the sensitivity kernel with respect to the fraction parameter $\chi$. The kernel can be simply expressed by the derivative of the observable for the model parameter of interest

$$\frac{\partial S}{\partial m_i} = K_{m_i}, \tag{14}$$

with the parameter vector $m = (\chi, \phi, \theta)$. Fig. 1 shows example sensitivity kernels in a homogeneous anisotropic model calculated with this formulation. We use the first derivative of a Gaussian as an incident wave with periods of 8 and 16 s. The broadening of the kernels with an increase of the dominant period of the incident wave is evident, as well as asymmetry due to non-vertical incidence, which can be considered in the forward calculation. The kernels of strength of anisotropy and fast axis orientation are nearly symmetric, while the dip of the symmetry axis shows strong asymmetric behavior, which is a result of the non-
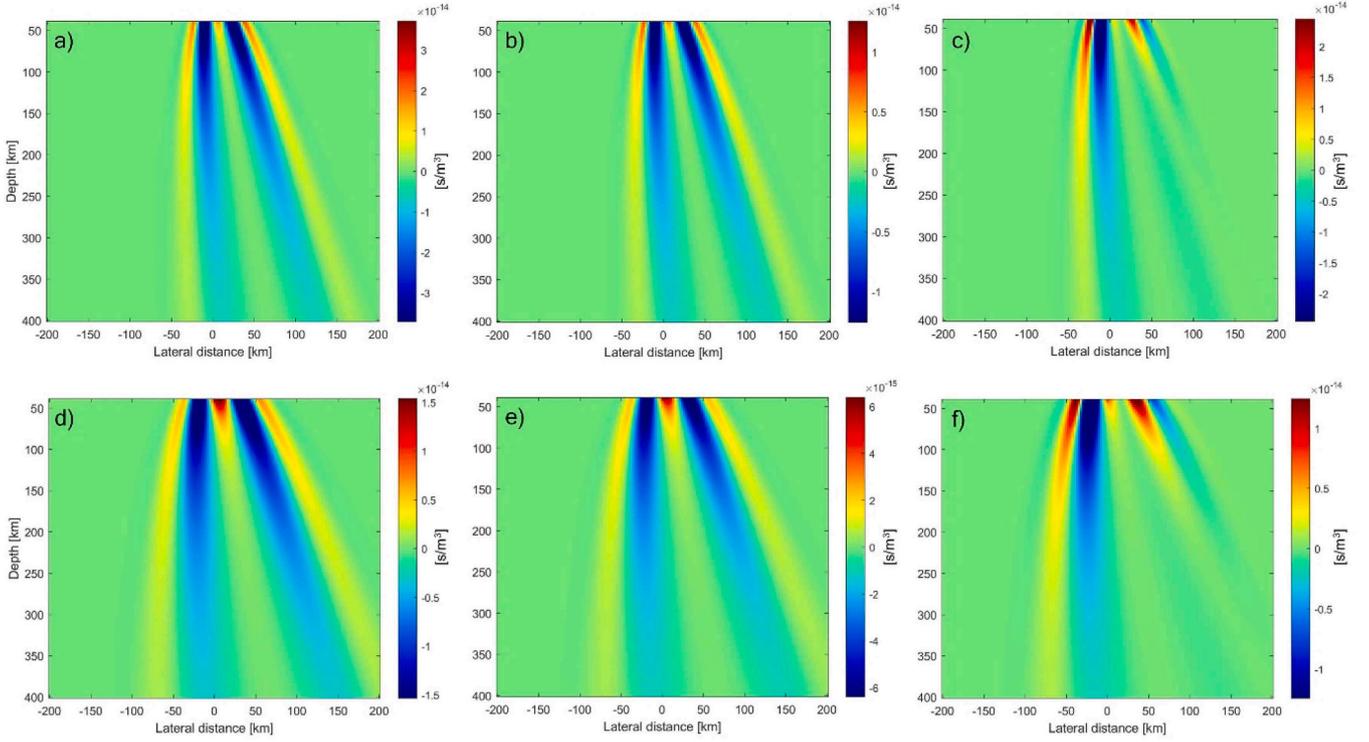
**Fig. 1.** 2D cut through the y plane at y = 0 km of sensitivity kernels for an incident plane wave of period 8 s, incidence angle 15° and backazimuth of 90° for the (a) strength parameter χ, (b) fast axis orientation with respect to north ϕ, and (c) plunge of the symmetry axis from horizontal θ. The model is chosen to be homogenously anisotropic with χ = 0.1, 45° plunge and 40° fast axis. (d-f) Same as (a-c) but for an incident plane wave of period 16 s.

vertical incidence in a medium with a dipping symmetry axis.

### 2.2. Inversion procedures

In the inversion for the anisotropic parameters $\boldsymbol{m} = (\chi, \phi, \theta)$, we wish to minimize the difference between the forward calculated $S_m$ and the observed splitting intensities, $S_{obs}$. This can be formulated as minimizing their $L_2$-norm, resulting in the cost function (e.g., Tarantola, 1987)

$$X(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{S}_m(\boldsymbol{m}) - \boldsymbol{S}_{obs}\|_2^2. \tag{15}$$

In practice, the inverse problem is usually ill-posed and underdetermined. Usually, a smooth result is preferred in tomographic imaging, which can be achieved by a Tikhonov regularization (e.g., Fernández-Martínez et al., 2014) in the form of

$$X(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{S}_m(\boldsymbol{m}) - \boldsymbol{S}_{obs}\|_2^2 + \frac{1}{2}\|\boldsymbol{\Gamma}\boldsymbol{m}\|_2^2, \tag{16}$$

where we use the second order finite difference operator $L_2(\boldsymbol{m}) = \frac{\partial^2 \boldsymbol{m}}{\partial x_i \partial x_j}$ as a smoothness constraint (Chevrot, 2006) in combination with the regularization factor $\alpha$ for the Tikhonov-matrix, $\boldsymbol{\Gamma}$,

$$X(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{S}_m(\boldsymbol{m}) - \boldsymbol{S}_{obs}\|_2^2 + \frac{1}{2}\|\alpha L_2(\boldsymbol{m})\|_2^2. \tag{17}$$

As we use angular model parameters, it is necessary to reformulate our model for the smoothing constraint to avoid penalizing neighboring model parameters at the start and end of an angular cycle (e.g., $\phi_{i-1,j} = 359°$ and $\phi_{i,j} = 1°$) as

$$\boldsymbol{m}' = \begin{pmatrix} \chi\sin(\phi)\cos(\theta) \\ \chi\cos(\phi)\sin(\theta) \\ \chi\sin(\theta) \end{pmatrix}. \tag{18}$$

The inversion of the cost function for the model parameters can be performed using different strategies. Here, we choose to implement and

compare three algorithms:

1) The gradient descent (GD) is the simplest algorithm, following the direction of the gradient of the cost function, $\frac{\partial\chi(\boldsymbol{m})}{\partial\boldsymbol{m}}$.

2) The BFGS-algorithm is a form of the conjugate gradient method, which adds information based on approximations of the Hessian matrix in each iteration to introduce additional information on the sensitivity of the model parameters to changes in the cost function.

3) A gradient-based stochastic inversion using a reversible jump Monte Carlo Markov chain (rj-McMC) is a fast-converging stochastic algorithm that allows for a data adaptive parametrization of the model space, which becomes significantly more efficient for large models compared to standard stochastic inversion methods.

#### 2.2.1. The gradient descent method

In the gradient descent method, we aim to optimize the cost function iteratively, beginning the search from a starting model $m_0$. As direction for the improvement of the data fit, we take the negative gradient of the cost function $\Delta m' = -\frac{\partial X}{\partial m}\big|_{m_0}$ (see also Boyd and Vandenberghe, 2004). We perform a line-search for the step-size, which avoids an over-shooting of the step over the improvement of the cost function. A new model at iteration $n+1$ is therefore constructed as $m_{n+1} = m_n - k\frac{\partial X}{\partial m}\big|_{m_n}$, where $k$ is the current step-size found in the line-search.

The model variables are not fully independent; therefore, we introduce two additional strategies in the inversion. We alternate the search for model improvement in fast axis direction, strength of anisotropy and dip of the symmetry axis (if the latter is considered in the inversion; in some applications, the symmetry axis is assumed to be horizontal). The inversion depends on the starting model. While we generally start with an effectively isotropic model (homogeneous anisotropy of 0.001), we choose an arbitrary fast axis direction, producing several starting models with a homogeneous orientation for the entire model area. To

investigate the influence of the starting model to the inversion and to identify consistent models required by the data, we perform the analysis for a range of different starting models, each with an arbitrary choice of the fast axis direction. The second strategy that we introduce in the inversion is to update the model only in areas with high sensitivity. This is implemented by identifying the area of the gradient with at least 50% of the maximum absolute sensitivity at each iteration and setting the remaining gradient to zero. This confines the changes of the model in space and reduces leakage of the anisotropic areas.

The distribution and quality of the data also potentially affect the inversion procedure. To ameliorate this effect, we use a subset of randomly selected events in each inversion. This procedure limits the influence of outlier traces with high noise levels, which otherwise might dominate the inversion procedure. The final model is estimated from the average of the model parameters for the inversion runs with the best data fit (we choose the best two third of the runs in this study). This allows us to also estimate the uncertainty of the model parameters based on their standard deviation after the model runs. At the same time, assuming that the inversion converges to the real model, the residuals of the observed data and the forward modelled splitting intensities based on the recovered model represent the uncertainty or error of the data.

### 2.2.2. The BFGS algorithm

A more frequently used and sophisticated approach is the BFGS algorithm (named after Broyden, 1967, Fletcher and Watson, 1980, Goldfarb, 1976, and Shanno, 1978). Here, an approximation of the Hessian matrix is introduced to allow a faster convergence while maintaining computational efficiency (see also Bonnans et al. 2006). The direction of the model step is determined considering the gradient and an approximation of the Hessian, $B_n$, resulting in $\Delta m' = -\frac{\partial X}{\partial m}|_{m_n} B_n^{-1}$. The actual step size is found in a line search similar to the gradient descent method, which results in the model update $m_{n+1} = m_n - \frac{\partial X}{\partial m}|_{m_n} B_n^{-1} k$. While for the first iteration an initial guess for the Hessian matrix is used (for example, the identity matrix), the estimate is improved in each iteration according to the formula

$$B_{n+1} = B_n - \frac{s y^T B_n + B_n y s^T}{y^T s} + \left[1 + \frac{y^T B_n y}{y^T s}\right] \frac{s s^T}{y^T s}. \tag{19}$$

with $s = s_n = m_{n+1} - m_n$ and $y = y_n = \frac{\partial X}{\partial m}|_{n+1} - \frac{\partial X}{\partial m}|_n$ (Bonnans et al. 2006). This new estimate for the Hessian matrix is then used in the next iteration for updating the step in the model towards improvement of the data fit. As for the gradient descent method, we introduce different starting models with an effective isotropic background and an arbitrarily chosen fast axis direction. As for the GD-algorithm, we choose to run each inversion with a different randomly selected subset of the data. The final model and the uncertainty of the model parameters are estimated from the mean and standard deviation of the model runs with the best fit to the data (from the best two thirds of the model runs in this study).

### 2.2.3. Formulation of the gradient-informed rj-McMC algorithm

For the reversible jump Markov chain Monte Carlo algorithm, we mostly follow the formulation by Bodin and Sambridge (2009), who describe this approach in mathematical detail. The algorithm estimates the posterior of an open set of model parameters using a random walk through the model space, where the number of cells in space is usually fixed. The model space itself is divided in Voronoi cells. During the random walk, one of these cells is selected and its model parameters are altered. The model fit of the new step is determined by calculating the likelihood and is then compared to the previous model fit (previous likelihood). Similar to the Metropolis-Hastings algorithm, an acceptance level is introduced with some random variability, which leads to the acceptance of the new model or its rejection. If the model is accepted, the model is taken as starting model for the next step; if it is rejected, the previous model is kept as starting model. In addition to the random walk, in which model parameters are altered in the current set of locally

fixed Voronoi cells, a three-step strategy is introduced in the reversible jump algorithm: an additional cell is added to the set of Voronoi cells, a cell is eliminated, or a cell is shifted in space. To allow faster convergence, we additionally use the gradient during the random walk to determine the direction in which the model parameters of the randomly chosen Voronoi cell can be improved. This greatly diminishes the number of rejected steps and improves convergence. We follow the strategy of a Metropolis-adjusted Langevin algorithm (MALA; Roberts and Tweedie, 1996), in which the samples of the Markov chain, $M_i$, are constructed based on Langevin diffusion as

$$M_{n+1} = M_n - \frac{h}{2}\nabla\log(\pi(M_n)) + \sqrt{h}\epsilon_n, \tag{20}$$

with a step size $h > 0$ a random draw from a multivariate normal distribution with zero mean, $\epsilon_n$, and a probability density function, $\pi(M_i)$, that is equivalent, here, to the maximum likelihood using Eq. (17):

$$\log(\pi(M_i)) \propto \frac{\left(\frac{1}{N}\sum_j \left(S_{m,j}(M_i) - S_{obs,j}\right)^2 + \frac{\alpha}{K}\sum_i \left(L_{2,i}(M_n)\right)^2\right)}{2\sigma^2}, \tag{21}$$

where $N$ is the number of measurements, $K$ is the number of model parameters, and $\sigma$ is an estimate for the average data uncertainty. This formulation allows us to include a regularization in the rj-McMC procedure (Vidal et al. 2020). While this regularization is optional for the user of the toolbox, we emphasize that the use of regularization ensures the smoothness of the likelihood function, which is a fundamental requirement for the convergence of this algorithm. Each iteration in MALA is checked against a Metropolis-Hastings acceptance criterion, which is equivalent to the acceptance criterion of the rj-McMC step used by Bodin and Sambridge (2009).

We perform multiple Markov chain calculations with different starting models by randomly selecting different geometries of the initial Voronoi cells. The priors in the MALA formulation are characterized by the gradient. However, to ensure that the model parameters remain within realistic range, we alter the prior distribution by defining boundaries for the model parameters. This is usually introduced by a uniform distribution with a value of 1 within and 0 outside of the given boundaries (0 to 1 for the anisotropic fraction, 0 to 180° for the fast axis, and −60° to 60° for the dip of the symmetry axis). As with the previously introduced algorithms, we start with an effectively isotropic medium with randomly selected initial fast axis orientation (with homogeneous orientation over the model space) for the different Markov chains. As for the other algorithms, we choose to run each inversion with a different randomly selected subset of the data. Finally, we construct a common posterior distribution by randomly selecting subsets (the number of Voronoi cells is chosen by the user of the toolbox) of the individual posterior distributions of the individual chains and interpolating them using a distance weighted algorithm (Shepard, 1968). Here we draw only from one third of the samples with the highest likelihood (corresponding to the best data fit). That leaves a variable amount of burn-in steps of the Markov chains, which are usually low in likelihood, and discards non-convergent chains. The interpolation produces a relatively smooth model, which is less affected by the appearance and sharp discontinuities of the Voronoi cells. However, the distance-weighted algorithm allows us to resolve sharp boundaries where it is required by the data, expressed by areas of high density of Voronoi cells. The final model is estimated from the mean of the different subsets and the uncertainty is estimated from the standard deviation.

### 2.3. Frequency dependent splitting intensity measurements

The lateral and vertical sensitivity of the observables depends strongly on the frequency of the incident wave (see Fig. 1 and Rümpker and Ryberg, 2000 for shear wave splitting more generally). However, the splitting intensity is mostly estimated as a characteristic value for a

measured waveform over a broad frequency band, as

$$S = 2 \frac{\mathrm{Re} \int_\Omega i\omega \, v_t(\omega) \, v_r(\omega) \, d\omega}{\int_\Omega \omega^2 \, |v_r(\omega)|^2 \, d\omega}, \tag{22}$$

where $v_r$ and $v_t$ denote the measured radial and transverse component of an observed shear-phase and both the numerator and the denominator are integrated over the entire frequency band. However, in practice the waveforms $v_r$ and $v_t$ are initially filtered to suppress noise and (ideally) increase the clarity of the signal. Here, we suggest reformulating the definition of the splitting intensity to explicitly incorporate this idea, while allowing a determination of the dominant frequency and a stable estimate of splitting intensity at the same time, with

$$S(\omega') = 2 \frac{\mathrm{Re} \int_{\omega'-d\omega}^{\omega'+d\omega} i\omega \, v_t(\omega) \, v_r(\omega) \, d\omega}{\int_{\omega'-d\omega}^{\omega'+d\omega} \omega^2 \, |v_r(\omega)|^2 \, d\omega}, \tag{23}$$

where $\omega'$ is the center frequency, and $d\omega$ defines the width of the frequency band. In our test, we use a frequency dependent band defined by a constant number of Fourier coefficients. We select the number of coefficients from a reference band corresponding to the periods of 6 s and 10 s. This allows us to calculate a frequency dependent splitting intensity equivalent to a moving band pass filter of the signal.

The window used to extract the phase of interest from the data stream potentially affects the estimated frequency content of the phase (e.g., Teanby et al., 2004; Savage et al., 2010; Wüstefeld et al., 2010; Liu and Gao, 2013). We therefore randomly vary the start and end time of

the window, estimating the dominant frequency of the radial component in this window corresponding to its maximum energy in frequency domain and estimating the corresponding splitting intensity using the dominant frequency as the center frequency. From the different windows, we construct a probability density function for the frequency-dependent splitting intensity. The maximum of the probability density function (PDF) is considered to provide the most robust estimate of the splitting intensity $S_0$, and its corresponding dominant frequency of the wave. In Fig. 2, we illustrate the difference between splitting intensity measurements derived by the classical formulation (Eq. (20), see Fig. 2b) and those derived from the dominant frequency band using our new formulation (Eq. (21), see Fig. 2d). We applied the analysis to a synthetically split waveform, starting from the first derivative of a Gaussian using splitting parameters of $\Delta t = 1$ s and $\varphi = 45°$ (Fig. 2a) and adding low pass filtered Gaussian noise with a corner period of 2 s (Fig. 2b). We varied the window for phase selection 500 times (semi-transparent red lines in Fig. 2a) and applied both the classical and frequency dependent analyses. For both methods, we calculated a probability density function for splitting intensity and dominant period (shown with a color-scaled 2D distribution in Fig. 2b, c). We find that the classical method systematically under-estimates the splitting intensity (Fig. 2b), which has also been suggested in earlier studies (Monteiller and Chevrot, 2010; Hein et al., 2021). However, focusing the formulation of the splitting intensity on the dominant frequency band reduces the influence of the noise and better recovers the expected splitting intensity for the input waveform.

We test the applicability of this method to a real data example (Fig. 3;
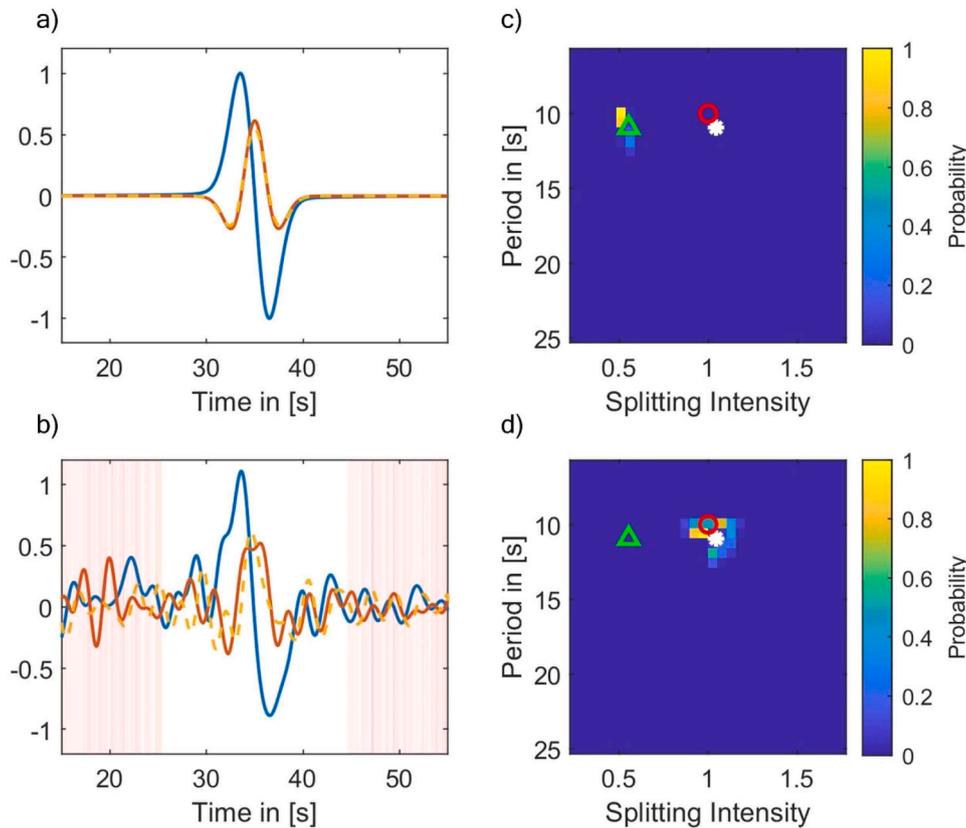


**Fig. 2.** (a) Synthetically split waveform with a nominal splitting intensity of 1.0 for an initial waveform defined by the first derivative of a gaussian with a dominant frequency of 10 s. Blue indicates the radial component, red the transverse component and yellow the time derivative of the radial component multiplied by the splitting intensity. (b) Same waveform as in (a) but with added Gaussian noise of relative amplitude 0.4 to the radial component. The red lines indicate start and end times of the time windows for which the calculation of the splitting intensity has been performed. (c; d) Probability density function of the frequency dependent splitting intensity for the classical formulation of the splitting intensity (c) and the frequency dependent splitting intensity (d). The mean of the distribution for the frequency dependent splitting intensity is marked by a white star, providing the estimate of the robust splitting intensity and corresponding dominant period for the waveform shown in (b). The splitting intensity and period of the noise free waveform are shown as a red circle and the mean of the distribution for the classical formulation of the splitting intensity is marked by a green triangle.
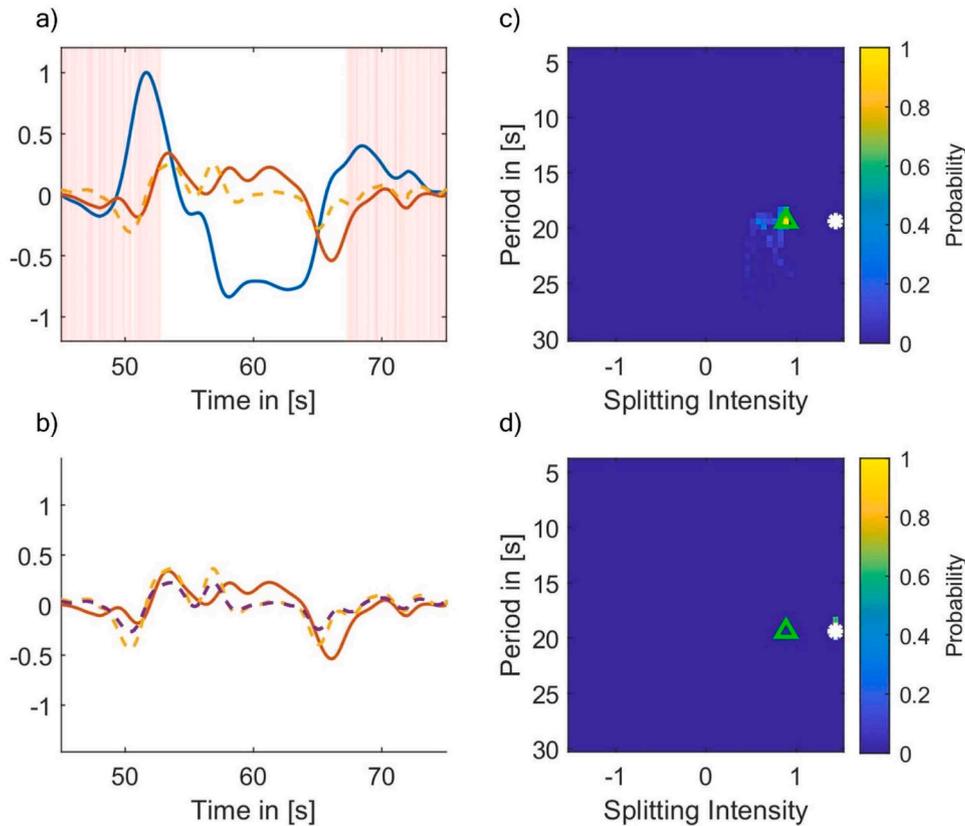
**Fig. 3.** (a) SKS-waveform measured at station D142 of the network ZS (Heit et al. 2017). Blue indicates the radial component, red the transverse component, and the yellow dashed line shows the time derivative of the radial component with a pre-factor of − 0.5. The red lines indicate start and end times of the time windows for which the calculation of the splitting intensity has been performed. (b) Red indicates the Transverse component waveform as in (a). The dashed yellow line and the dashed purple line indicate the time derivative of the radial component scaled by the splitting intensity for the frequency dependent and the classical analysis method, respectively. (c; d) Probability density function of the frequency dependent splitting intensity for the classical formulation of the splitting intensity (c) and the frequency dependent splitting intensity (d). The mean of the distribution for the frequency dependent splitting intensity is marked by a white star providing the estimate of the robust splitting intensity and corresponding dominant period for the recovered waveform shown in (b). The result for the classical formulation of the splitting intensity is marked by a green triangle corresponding to the maximum of the probability density function in (b).

station D142 of the Swath-D network; Heit et al. 2017). We find a splitting intensity of 0.89 for the classical method and 1.43 for the frequency dependent measurement and identify the same dominant frequency of 19 s. Equivalent to the synthetic example, this example from real data indicates that the classical method for a broad frequency band underestimates the splitting intensity. We note that selective band-pass filtering with a narrow frequency band prior to the classical estimate of the splitting intensity leads to the same result as the method we propose here. However, the new method allows a data-adaptive approach that identifies the dominant frequency and estimates the corresponding robust splitting intensity automatically, without the need of extensive data evaluation, as one key element of our approach is the re-evaluation over multiple windows of different length and position.

## 3. Test of the inversion procedures

### 3.1. Synthetic test setup

We present synthetic tests that are designed to illustrate the strengths and weaknesses of the different inversion procedures. We calculate synthetic splitting intensities for two different input models (model I, II), described further below. For model I, we introduce one version with a variation in the dip of the symmetry axis and one without (models Ib and Ia, respectively). All models vary in two dimensions, while the third dimension simply extends the anisotropic properties from the 2D plane, representing a 2.5D model. This reflects the basic assumptions for an analysis of a seismic profile, which is currently the most common set-up

for a possible application of the presented algorithms, considering the required density of station spacing (e.g., Mondal and Long, 2019, 2020). All models consider a hypothetical upper mantle between 40 and 410 km depth and extend laterally from − 250 to 250 km in x-direction as the variable model dimension and − 250 to 250 km in y-direction, in which the parameters are kept constant. We calculate splitting intensities for 2000 station-event pairs with randomly chosen period varying between 4 and 20 s, randomly selected backazimuth, and randomly chosen ray parameter corresponding to SKS phases at a distance range between 89° and 140°. We assume an incident plane wave, $w(t)$, with a waveform based on the first derivative of a Gaussian, which allows to define a dominant frequency. We assume a simplified (linearly approximated) 1D shear wave velocity model based on ak135 for the calculations (Kennett et al. 1995). We estimate wave propagation based on Snell's law within the model space based on the incidence defined by the ray parameter and azimuth of an event for the radial wave without any effect on its absolute amplitude. The transverse component is estimated from the scattering introduced by the anisotropic medium as described in Section 2.1. The hypothetical stations are located at the surface (z = 0 km) along the x-direction with 5 km spacing. The model dimension is extended in the x-dimension by 250 km in both directions, to allow the calculation of all relevant parts of the sensitivity kernels contributing to the start and end of the hypothetical profile. This area is denoted as a border area, while the remaining area is denoted as the model space. The background isotropic velocity is taken from PREM; we note, however, that the isotropic velocities have no considerable impact on the splitting intensities.
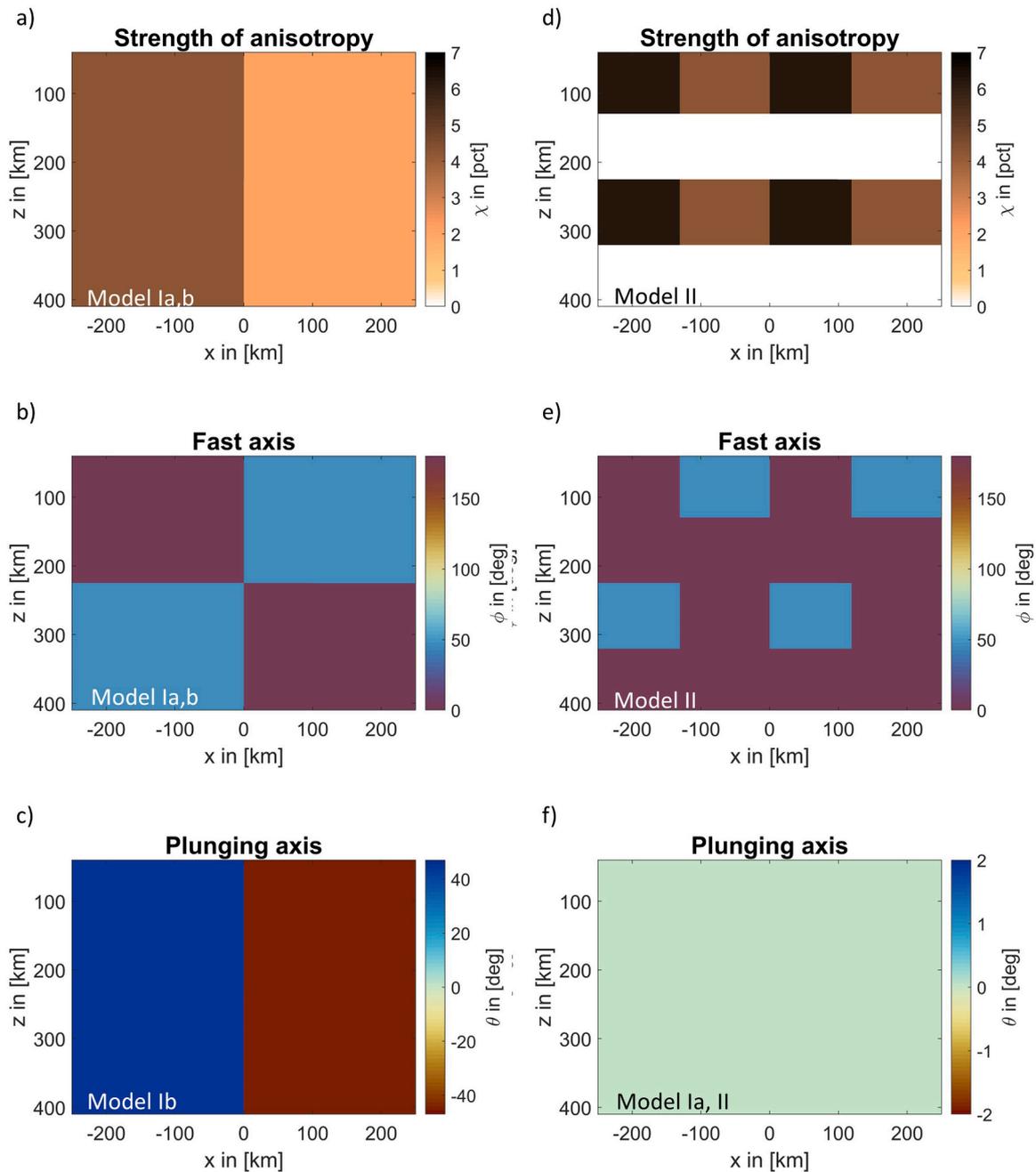
**Fig. 4.** Strength of anisotropy (a, d), fast axis direction (b, e) and plunge of the symmetry axis (c, f) from top to bottom for input models defined by a checkerboard model with (a, b) 4 cells (model I) and (d-e) 16 cells (model II). The plunging axis for model Ia and II has been set to 0 (f), while model Ib is defined by a variation in the plunging axis (c).

The first model is characterized by 4 equally spaced volumes within the model space (Fig. 4a-c), with an anisotropic strength that varies between 4% to 2%, while the fast axis orientation varies in a checkerboard pattern between 0° and 45°. While for model Ia the dip of the symmetry axis is kept at 0°, for model Ib the dip of the symmetry axis varies laterally from 45° to − 45°. Model II is chosen to be more complex, with 16 volumes within the model space (Fig. 4d-f). The strength of anisotropy varies laterally, alternating between 6% and 4%. Vertically, the anisotropic pattern varies in 4 layers. The first and third layers are anisotropic, while the second and fourth layers are isotropic. Within the anisotropic layer, the fast axis varies in a checkerboard pattern between 0° and 45°. For all models, we use a grid of 64 by 32 cells in the calculations.

For the gradient-based inversion procedures, we invert from 50 different starting models using a subset of 1000 observations in each run to test the statistical stability of the results. We expect no direct effect of the reduction of the observations to a subset of the synthetic data; our aim here is to limit the influence of noise in the data on the inversion in real data applications. However, we chose this procedure for consistency with the real data application in the following section. The GD-method was run without a convergence criterion for 50 iterations. In the BFGS-method, we use the infinity norm of the gradient, $\max\left|\frac{\partial X}{\partial m}\right|_{m_o}\right| = 1e^{-6}$, as a convergence criterion and stop the inversion either when convergence is achieved or when the number of iterations reach 50 steps. For the rj-McMC inversion, we run 10 chains with different starting models for 1500 iterations, also using a subset of 1000 observations in each run. We assume an average data uncertainty of $\sigma = 0.2$ for the synthetic splitting intensities. The final model is constructed by drawing random subsets of

1024 Voronoi cells from the best third of Markov chains with highest likelihood. For all methods, a small smoothness parameter of 0.075 is chosen with the aim of avoiding models with extreme changes of symmetry axis direction. This choice is made because neighboring cells with close to 90° offset in symmetry axis orientation would have the effect of destructive interference for the splitting intensity; the method has no direct sensitivity to this configuration. With our choice of small smoothing, we try to avoid these extreme offsets between neighboring cells, while still allowing to resolve relatively sharp changes in the anisotropic parameters.

### 3.2. Synthetic test results

Results of our synthetic test runs are shown in Figs. 5–7. In our first application (Fig. 5), we only search for the strength of anisotropy and fast axis direction, assuming the fast axis to be horizontal. The application of all three techniques to model Ia shows that the model is well recovered independent of the technique. In particular, the sharp change of model parameters at 0 km is identified by the gradient based methods. The rj-McMC method provides less strong constraints on this lateral transition, which might be result of the interpolation method

used for the model reconstruction from the posteriors and might get better resolved for a longer sampling or a different choice of data uncertainty. However, no methods are capable of resolving the sharp transition in the vertical direction of the fast axis orientation between the top and bottom layers (Fig. 5b, e, f). While the inversion results do approximately match the synthetic models within the model uncertainties (see Fig. S1), this vertical transition appears smooth. This smoothness is strongest for the GD-method, slightly less smooth in the BFGS-method and closest to the synthetic model in the rj-McMC-approach. We identify small but distinct performances in model fit among the three algorithms for this example. The BFGS-method shows the best fit to the data for our choice of convergence criteria, while the GD-method and rj-McMC-method show broader width of the residual distribution and larger RMS. However, the initial model parameters are all resolved within the uncertainty of the model parameters (Fig. S1).

When applying the techniques to model II (Fig. 6), the challenges with vertical resolution become even more apparent. While the fast axis direction is recovered well by all techniques, the strength of anisotropy is significantly obscured. The first layer is well recovered, indicating the strong sensitivity of the splitting tomography at shallower depth (in this case, <100 km). However, while all techniques find a decrease of
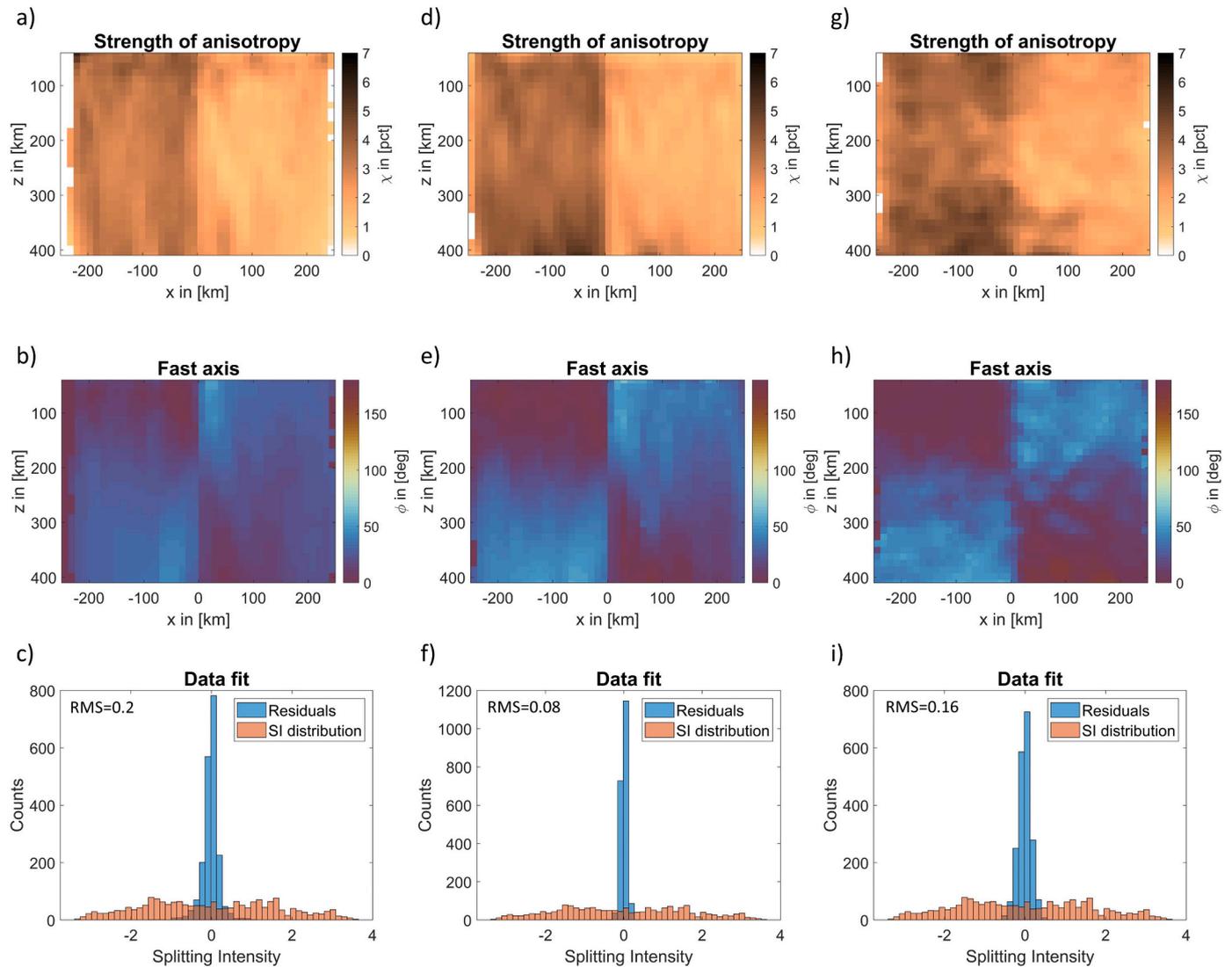


**Fig. 5.** Strength of anisotropy (a, d, g), fast axis direction (b, e, h) and data fit (c, f, i) from top to bottom for an inversion of model Ia based on: (a, b, c) gradient descent; (d, e, f) BFGS algorithm; (g, h, i) stochastic gradient descent with rj-McMC algorithm. (c, f, i) The orange bars show the distribution of the synthetically calculated splitting intensities. The blue bars show the distribution of the residuals after subtracting the forward calculated splitting intensities based on the resulting model parameters shown in the column above (a, b, d, e, g, h).
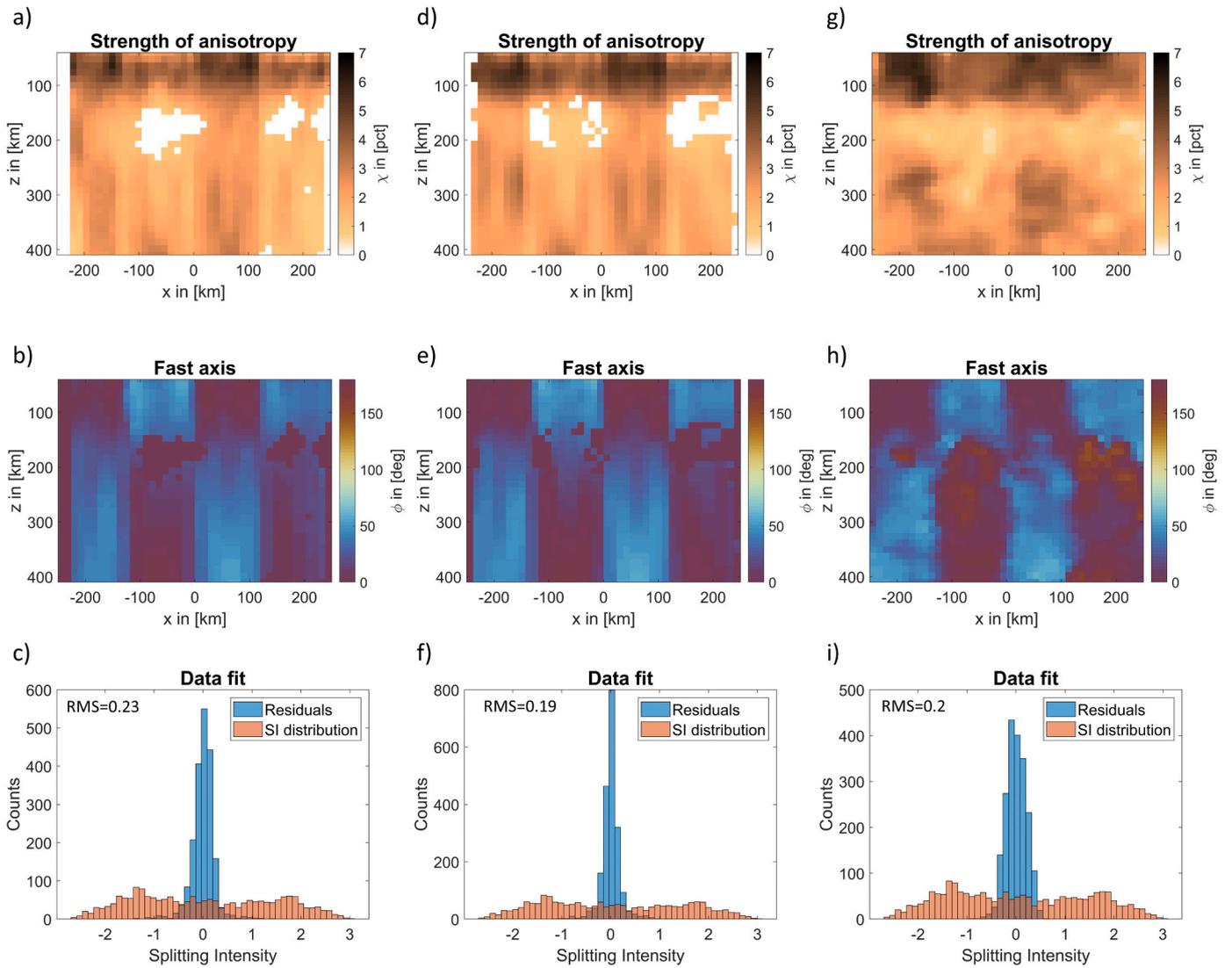
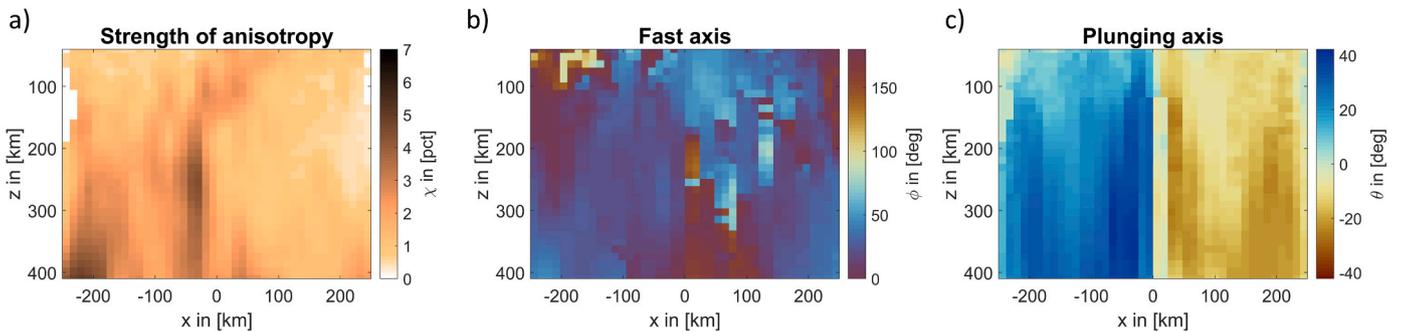**Fig. 6.** Same as Fig. 5, but for an inversion of model II.



**Fig. 7.** Strength of anisotropy (a), fast axis direction (b) and plunge of the symmetry axis (c) from left to right for an inversion of model Ib based on the BFGS algorithm considering a dip of the symmetry axis in the inversion.

anisotropy below the strong anisotropic layer, no technique resolves well the first isotropic layer. In particular, the solution derived from the rj-McMC algorithm does not reflect the isotropic features in the model. In contrast to model Ia, where the synthetic model was within one standard deviation of the inversion results of the rj-McMC, the isotropic property of the second layer remains unresolved, when considering the data uncertainty for model II (see Fig. S2). Qualitatively, in all models

the alternating nature of stronger and weaker anisotropy is found for the third layer, while the strength is strongly underestimated and the anisotropy smears to greater depth. The deepest isotropic layer is not identified by any of the techniques. The data fit for all models is good, although the BFGS-algorithm shows the best performance with the inversion parameters that we used in this particular example.

In the final test of the inversion procedure (Fig. 7), we focus on the

BFGS-algorithm in the application to a more complex model, as it represents the best performing approach in the previous tests for our choice of the convergence and cut-off criteria. Here, we apply the search for the full orientation of the elastic tensor, by considering a dipping symmetry axis. This analysis is performed on the synthetic data set of model Ib. Fig. 7 shows the result of the BFGS-algorithm for the mean of 50 inversion runs with different starting models. Unlike the analysis for model Ia, the strength of anisotropy and fast axis direction are not as well resolved. However, the model volumes showing positive and negative dip are identified well. The magnitude is under-estimated at shallow levels, but more precisely estimated with growing depth. The decrease in resolution of strength and fast axis orientation at shallow levels indicate a strong trade-off between the parameters.

In our synthetic tests, we compared a simple gradient descent (GD), a BFGS-approach, and a gradient-informed reversible jump Markov chain Monte Carlo approach (rj-McMC) in application to synthetic data. For simple models (represented by model I), all three inversion methods recover the model well within their uncertainty. The GD and the rj-McMC algorithms tend to produce lower data fit with the convergence and cut-off criteria we selected for the synthetic tests. While all methods resolve the model well in the lateral direction, the depth resolution is considerably poorer, as expected in XKS splitting tomography due to the nearly vertical incidence of the waves. This results in vertical smearing and in a gradual change between different model volumes, although the input model has a discontinuous jump (seen particularly well in the inversion for model II). Considering the full orientation of the anisotropic tensor, including the plunge of the symmetry axis, adds complexity that introduces ambiguities in the inversion. While the plunge can be well identified, the strength and horizontal orientation are less well resolved by the tomography. We note that the magnitude of the splitting intensities is significantly lower than for the horizontally oriented case, which might explain the poor recovery of the anisotropic strength.

## 4. Application to the dense Swath-D data set

The Swath-D-network provides an ideal real data test case for the application of splitting intensity tomography. This temporary seismic network covers the transition of the Central to the Eastern Alps with 154 densely spaced broadband stations (Heit et al. 2017; AlpArray Seismic Network, 2015; Fig. 8). Previous studies have inferred complex anisotropic characteristics for this area, expressed not only by a sharp lateral change of splitting parameters at ~13°E, but also by strong azimuthal variation of the parameters in the Eastern Alps, indicating anisotropic layering (Qorbani et al. 2015; Link and Rümpker 2021). The classical ray-based technique did not allow for a direct resolution of the depth which produced the anisotropic pattern. This motivates the application of splitting intensity tomography, with the aim of getting at least a first-order idea of the depth distribution of the anisotropy in this area and possibly insights into its origin.

We base our analysis here on the same data set as the previous SKS splitting study by Link and Rümpker (2021) using the Swath-D network, which provides two years of continuous seismic data at 154 stations between 2017 and 2019. This data set is accessible from the GEOFON archive. We also include data from permanent stations with longitudes between 10° − 14.5° and latitudes between 45.5° − 47.5°, following Link and Rümpker (2021). The data is assembled from the networks BW, CH, IV, NI, OE, OX, SI and SL (Fig. 8 and Table S1). Teleseismic events with magnitudes above 5.8 and within the distance range from 89 to 140 degrees were selected. A signal to noise ratio is calculated based on the normalized signal energy in a 25-sec window after the expected arrival time of the phase and the normalized noise energy in a 20-sec window preceding this arrival time. Phases with a signal to noise ratio below 2.5 are discarded. Splitting intensities are calculated for SKS, SKKS, SKIKS, PKS and PKIKS phases with high-quality waveforms based on the method described in Section 2.3; this procedure also provides an estimate of the dominant period of each core-mantle converted phase. The splitting patterns show strong two-dimensional characteristics (Link and Rümpker 2021), suggesting that a 2.5D model is appropriate. We therefore project the station locations on a profile intersecting the network from WSW to ENE (Fig. 8a). We exclude stations at a distance larger than 0.5° from the profile, which results in a total of 2771 observations used in the analysis.

Our synthetic tests showed the superior performance of the BFGS-algorithm in the comparison to the other two algorithms shown in this study. We therefore choose the BFGS-algorithm as basis for our application of the splitting intensity tomography to a real data set. The model is divided in 64 by 32 cells, which corresponds to ~10 km steps in z and ~5 km in x direction. In general, densifying the grid allows a higher accuracy in the calculation of the contributions to the splitting
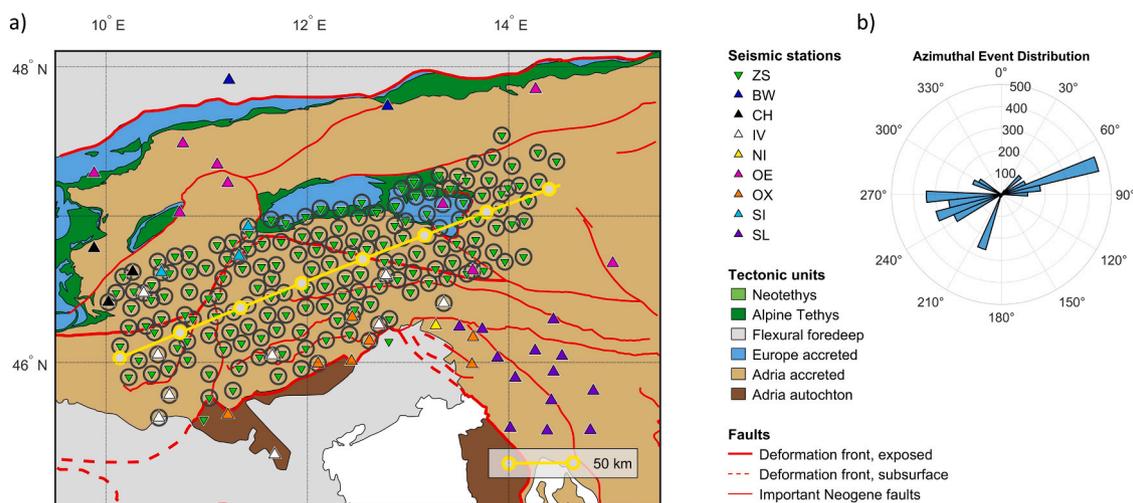


**Fig. 8.** a) Station distribution of the dense Swath-D seismic network (ZS) and permanent stations of national networks in the same region. The background shows the geological map compiled by M.R. Handy with units and major lineaments simplified from Schmid et al. (2013, 2008); Handy et al. (2010, 2015, 2019), Bigi et al. (1992), Froitzheim et al. (1996) and Bousquet et al. (2012). Selected profile (yellow line) with 50 km markings (gray circles with yellow outlines). This line with markers is also shown in Fig. 9 as reference. Stations located within 0.5° distance from the profile (black circles) are used in this analysis. b) Azimuthal distribution of the event-station pairs used the analysis. Most events arrive from two opposite directions aligned with the strike of the profile, which results in a good overlap of the corresponding sensitivity kernels for a successful depth resolution.

intensities and sensitivity kernels. However, at the same time larger computational resources become necessary. We find that a lateral spacing of about 5 km allows sufficient accuracy, while such a low grid spacing in depth has less impact on the forward calculation. We therefore use a less dense spacing of 10 km at depth. We apply the inversion in two steps. First, we only consider a horizontal symmetry axis, solving for strength of anisotropy and fast axis direction (relative to north). In a second step, we use the model found in the first step as initial model and re-run the analysis, searching additionally for the dip of the symmetry axis. Similar to the previous applications, we use 50 different starting models in the first step using an arbitrary orientation of the horizontal fast axis, and use for both steps different subsets of the data in the inversion for the 50 different runs. This procedure allows us to investigate the uncertainties of the data and its impact on the inferred model parameters. At the same time, we avoid dominating effects of outliers in the data.

Fig. 9a-d shows the mean result of the 50 inversion runs (based on different initial models and event selections as described above) without

consideration of a dipping symmetry axis. From these inversion results, we identify a sharp change of fast axis orientation from around 60° in the west to about 110° in the east (see also Fig. 10 in the discussion). This change is found at around 250 km, which coincides with 13° longitude. A shallow volume with increased anisotropic strength is found in the east at depth between 70 and 190 km (see structure A in Fig. 10). West of 13° a fast axis of 60° dominates, corresponding to an anisotropic volume beneath 100 km depth. An increased level of anisotropic strength appears to dip downwards from the west to larger depth closer to the transition in the east, where it reaches the lower boundary of the model space (see structure B in Fig. 10). The second step of the inversion procedure allows the dip of the symmetry axis to vary as well (see Fig. 9e-h). In this second inversion result, the strength of anisotropy and fast axis direction remain stable, and we identify a slight dip of the symmetry axis, between $-8°$ and $+6°$ from horizontal. The axis dipping downward to the east is found in the area west of 13° longitude, while an axis dipping upward to the southeast is found for the shallow anisotropy in the east. Fig. 9(g; h) shows the data fit of the two inversions, which
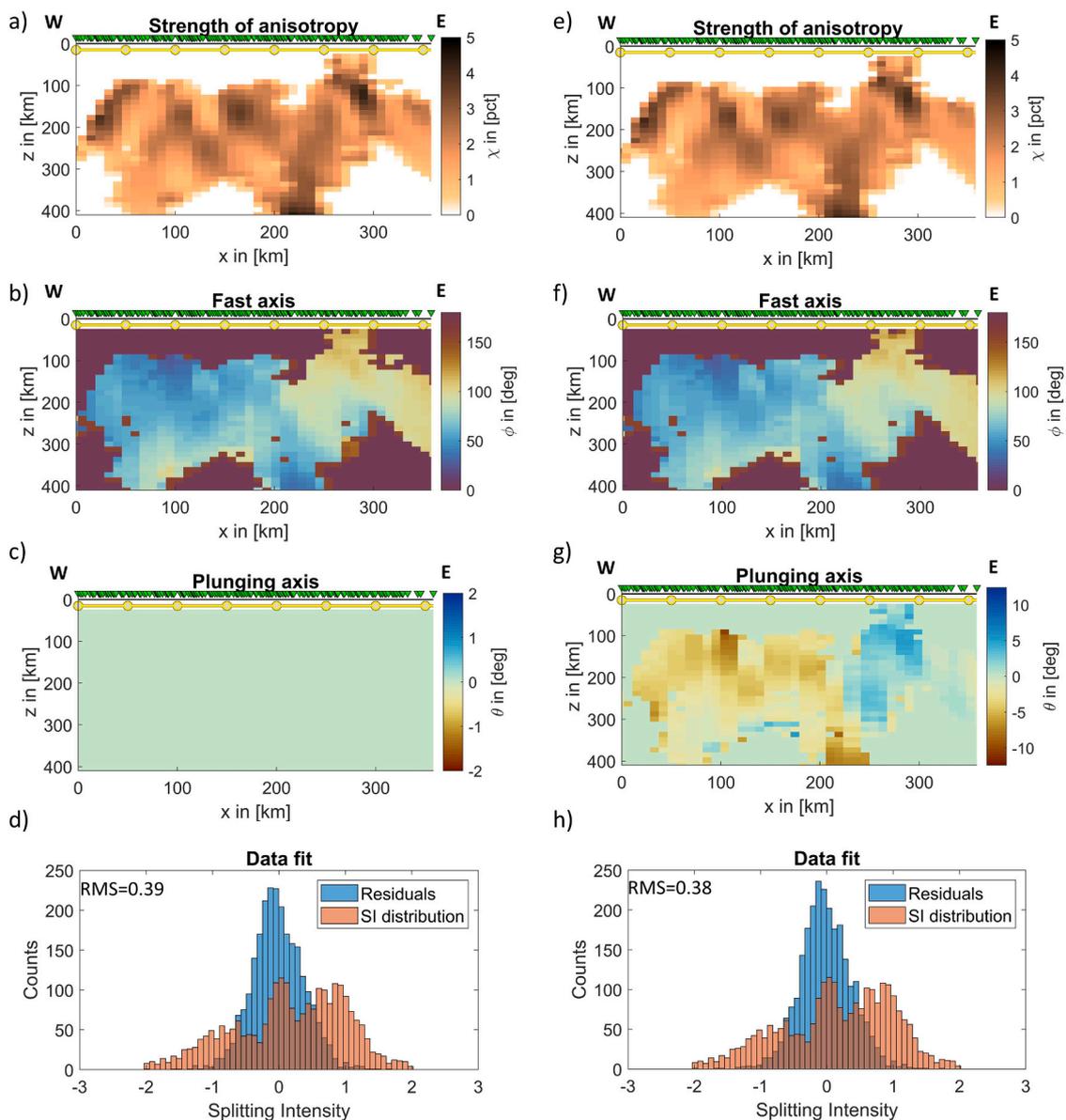


**Fig. 9.** Strength of anisotropy (a, e), fast axis direction (b, f), plunge of the symmetry axis (c, g) and data fit (d, h) from top to bottom for an inversion of the Swath-D data set based on the BFGS algorithm. We only consider a horizontal orientation of the symmetry axis in a first inversion procedure (a, b, c, d) and allow for the full orientation of the tensor (e, f, g, h) in a second inversion step that is based on the initial findings shown in (a, b, c). The yellow lines in (a, b, c, e, f, g) represent the profile line in Fig. 8 with yellow/gray dots at 50 km spacing.
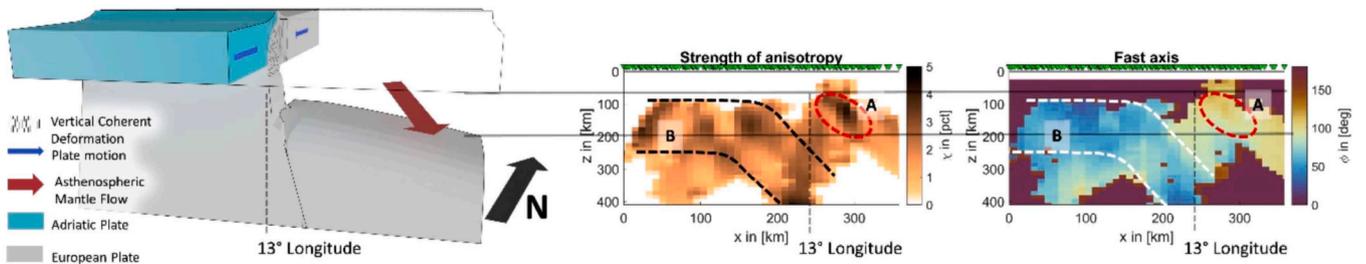
**Fig. 10.** Schematic figure of the findings for the anisotropic tomography in the Eastern Alps. The diagram on the left shows suggested slab break-off in the Eastern Alps, which allows an opening for asthenospheric flow through the corresponding gap (red arrow). We identify an increase of anisotropy corresponding to a distinct change of fast axis at 13° Longitude at a depth between 80 and 190 km (compare red dashed line, structure A). The increased level of anisotropy west of 13° longitude tends to decline to larger depth to the bottom of the model space from west to east (structure B).

reflects the uncertainty based on the measurement error. The standard deviation of the first inversion is 0.39 s while it is 0.38 s for the second inversion. Allowing the symmetry axis dip to vary results in only a minor improvement, which is statistically not significant.

## 5. Discussion and summary

We present a new toolbox for the MATLAB environment that enables the tomographic inversion of splitting intensity measurements for anisotropy beneath dense seismic networks. The method aims to overcome the limited depth resolution of classical shear-wave splitting techniques that results from the ray-based approximation of the nearly vertically incident core-mantle converted phases. With the formulation of finite-frequency sensitivity kernels, we obtain increased sensitivity at depth by accounting for the three-dimensional volume contributing to the shear wave splitting observations. The toolbox provides a new frequency-dependent estimate of splitting intensities for core-mantle converted phases that can be imported from a dataset previously processed with the SplitRacer software (Reiss and Rümpker 2017; Link et al. 2022). This estimate of the splitting intensity identifies the dominant frequency band of a waveform and allows a more stable measurement estimate. The inversion approach is based on the forward calculation of splitting intensities and gradients of the cost function. Our sensitivity kernel computations allow an analytic formulation (Chevrot, 2006; Mondal and Long, 2019), which increases its computational efficiency compared to adjoint methods often used for estimating sensitivity kernels (e.g., Tarantola, 1987; Tromp et al. 2005; Sieminski et al., 2009). The formalism allows us to introduce any effective hexagonal elasticity tensor; however, we focus on olivine LPO with the model parameterized by the strength of the anisotropic contribution relative to its isotropic approximation (Mondal and Long, 2019) and the orientation of the fast anisotropic symmetry axis in space. Extending the previous work by Mondal & Long (2019), we also allow for the formulation of the anisotropic tensor from an anisotropic perturbation of isotropic velocities (Farra et al. 1991), which provides directly an expression for strength of anisotropy. While Mondal and Long (2019) only considered the far-field term and focused on vertically incident plane waves, we allow here for non-vertical incidence and include the near- and mid-field terms (similar to Chevrot, 2006), resulting in higher precision of the forward calculation. To stabilize the inversion, we introduce the Laplacian operator as regularization in the cost function to reduce the roughness of the model (Chevrot, 2006). A new feature in this study compared to previous work on splitting intensity tomography is the implementation of three inversion strategies; previously, Mondal and Long (2019) implemented a stochastic Gibbs sampler and Chevrot (2006) implemented a deterministic solver based on LSQR algorithm): Here we introduce a simple gradient descent (GD), a BFGS-approach, which approximates the Hessian and stabilizes the inversion, and a gradient-informed reversible jump Markov chain Monte Carlo approach (rj-McMC). The new algorithms allow fast convergence, due to the consideration of the gradient in all techniques, which allows to explore a larger model space with

similar computational resources as the technique presented by Mondal and Long (2019). Furthermore, the toolbox is flexible, such that users can choose the algorithm and parameter choices that are best suited to the needs of the problem under study.

We compare the three techniques in application to synthetic data, illustrating the strengths and weaknesses of SKS-splitting intensity tomography and of the three different inversion strategies. For simple models, all three inversion methods recover the model well, but the BFGS-algorithm leads to the best fit to the data. The GD and the rj-McMC algorithms result in a lower data fit in our tests. We emphasize that this might be related to our choice of convergence and cut-off criteria for the inversion procedures; and a different choice of convergence criteria or extended inversion steps might lead to a similar data fit as the BFGS-algorithm. However, trends of the RMS-error of the splitting intensities indicate fast convergence for all Markov chains (Fig. S3). Only the larger scatter of the number of Voronoi cells might indicate that extended iterations might be required. In any case, users of the toolbox may select the convergence and cut-off criteria that make the most sense for their individual applications.

While all methods resolve the model well in the lateral direction, the depth resolution is considerably poorer. This results in vertical smearing and in a gradual change between different model volumes, even though the input model has a discontinuous jump. This is not surprising; while depth sensitivity is introduced by the consideration of overlapping sensitivity kernels at depth, we still rely on core-mantle converted phases with almost vertical incidence. While this leads to the excellent lateral resolution of structure, the similar shape and large overlap of the sensitivity kernels at depth has an averaging effect. As a result, the vertical resolution decreases with depth. While the fast axis orientation in the horizontal plane is generally very well resolved, the dip of the symmetry axis has a complex trade-off with the strength of anisotropy. The poor recovery of these parameters can be explained by the fact that a considerable dip leads to a decrease of the magnitude of the splitting intensities and therefore a decrease of the overall sensitivity for the anisotropic structure. We expect a similarly good resolution of those parameters, comparable to models without a dipping symmetry axis, for cases with stronger anisotropy, leading to splitting intensities at similar magnitude even for a dipping anisotropy case.

The computational costs of the various algorithms are determined by their inherent characteristics. While the BFGS- and the GD-methods are comparably highly efficient, due to a rapid convergence based on the purely gradient driven model improvement, the rj-McMC algorithm requires many more iterations. The additional iterations (and slower convergence) of this algorithm are required by its stochastic element, which allows us to estimate uncertainties. At the same time, the rj-McMC is not based on a fixed dense model grid, as are the purely gradient based methods, but aims to estimate the actual number of model parameters required to fit the data.

Despite the challenges, the application of SKS-splitting intensity tomography to the dense Swath-D network (Heit et al. 2017) shows the potential of the method for real data sets. Previous studies found

complex anisotropy beneath the Eastern Alps (Qorbani et al. 2015; Link and Rümpker 2021). For the Swath-D data set, two significant anisotropic layers were previously identified: one layer with constant fast axis direction of about 60° over the whole area, with a possible east-dipping symmetry axis and a shallower layer east of 13° longitude with a fast axis direction of about 115°. This general observation agrees well with our inversion results (Figs. 9 and 10). Two possible geodynamic scenarios had been previously suggested for this region, both of which invoked a slab break-off in the eastern part of the study area, allowing asthenospheric flow through its gap (Link and Rümpker, 2021). The anisotropic structure west of 13° longitude was interpreted in the past as either frozen-in anisotropy with lithospheric origin or an asthenospheric flow evading the retreat of the European slab that would precede the break-off event.

Our models (Fig. 10) show that the shallower anisotropic structure beneath Swath-D extends from about 70 km to around 190 km depth (labeled as structure A in Fig. 10). The fast axis orientation of this shallow structure indicates likely asthenospheric flow, as it cross-cuts geologic structures visible at the surface and it likely dominates over this entire depth range. The finding that this anisotropic volume extends to the very shallowest mantle (~70 km) indicates an almost complete delamination of the lithospheric mantle during the slab-break off. While our synthetic tests highlight the fact that the depth resolution of XKS splitting intensity tomography is limited, and vertical smearing of anisotropic structures occur, the resolution of the upper 150 km appears acceptable for this model (compare Fig. 6a, b). We therefore assume that the estimated depth and anisotropic parameters for the shallow structure east of 13° longitude is robust. The distribution of the anisotropic structure with ~60° fast orientation starts at a deeper level of 100 km (structure B in Fig. 10), which is still at a depth with considerable resolution. This would be consistent with either an isotropic lithosphere in the overriding Adriatic plate and lithospheric anisotropy in the downgoing Eurasian slab, or with the predictions of an asthenospheric evasive flow model. We suggest that there are two aspects of our model that might indicate a lithospheric origin for this structure. First, the increased anisotropic strength west of 13° longitude tends to decline to larger depth from west to east, reaching the bottom of the model space at 13° longitude. However, we caution that the resolution at this depth is very limited, as shown in the synthetic tests. Second, there is a slight indication of a downward dip of the symmetry axis in this volume (see Fig. 9), which would tend to support the idea of lithospheric anisotropy that is altered by the down- and eastward-directed tear in the slab break-off event. We caution, however, that the dip of the anisotropy does not result in a statistically significant improvement of the data fit; an over-interpretation of this feature should therefore be avoided. We also note the relatively large width of the residuals in the real data application compared to the synthetics (Fig. 9), which suggests that the splitting intensity estimates are strongly contaminated by noise, leading to a large data uncertainty. In addition, errors may be introduced by 3D heterogeneities in anisotropic structure, which are not considered in our 2D projection of the measurements to a profile. A more detailed interpretation for the anisotropy west of 13° longitude is therefore not possible, and an evasive asthenospheric flow model, as previously suggested (Barruol et al. 2011; Petrescu et al. 2020; Hein et al. 2021; Link and Rümpker 2021; Link and Rümpker 2023), cannot be excluded. A more conclusive answer about the geologic interpretation might potentially be found in the further analysis of the anisotropic distribution beneath the crosscutting seismic profiles of the TRANSALP (Lüschen et al. 2004) and EASI (Hetényi et al. 2018) experiments in future applications of this method.

There are several potential strategies to overcome some of the shortcomings of the SKS splitting intensity tomography method, particularly its lack of depth resolution. One possible future improvement is to approximate the full radial and transverse wavefield by allowing a change of the incident radial waveform due to the anisotropic structure as well. This requires the estimation of the sensitivity kernels in an adjoint approach and would allow for the use of waveforms instead of splitting intensities. Previous studies have shown that waveforms contain additional information about vertical structure not contained in splitting intensities, which allow vertical resolution only by the overlap of the sensitivity kernels (Rümpker et al. 2023). Another potential strategy is to incorporate additional, complementary data sources such as splitting of S-phases from local earthquakes (e.g., Abt et al. 2010), anisotropy of surface waves (e.g., Yuan and Romanowicz, 2010; Wagner and Long, 2013), or anisotropy-aware receiver function constraints (e.g., Yuan and Levin, 2014; Luo et al., 2023). Each of these analysis strategies provides additional information about the vertical distribution of anisotropy and improves the resolution of the orientation of the elastic tensor as a function of depth. On another note, while we only allow for one estimate of robust splitting intensity from our new frequency dependent approach, real waveforms potentially exhibit multiple sets of dominant frequencies. In future applications of this method, these multiple frequencies can be considered and exploited, making multiple, frequency-dependent estimates of splitting intensities for each waveform. This might provide additional information about the sampled anisotropic structure and enhance the capabilities of the inversion procedures. More generally, future applications of joint inversions of SKS splitting data with other types of complementary seismic and geologic information might allow a more detailed insight into the 3D-distribution and orientation of seismic anisotropy and the corresponding geodynamic and tectonic implications.

## CRediT authorship contribution statement

**Long Maureen D.:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Link Frederik:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

We have shared the link to the code presented in this paper within the manuscript (see Computer Code Availability section). The data example is publicly available alongside the cited publication (Link & Rümpker, 2021).

## Acknowledgements

*Computer Code Availability*

The code SItomo for the tomographic inversion of Splitting Intensity using finite frequency sensitivity kernels was developed by Frederik Link (Department of Earth and Planetary Sciences, Yale University, New Haven, Connecticut, United States, email: frederik.link@yale.edu). The software package is available at the GitHub repository: https://github.com/seismolink/SItomo. The required storage space is 4 Mb. The Code is written for the MATLAB-environment and compatible with the version 2019b or newer. Supporting information for the installation of the Software package can be found in the User Guide included in the package.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jog.2024.102018.

## References

Abt, D.L., Fischer, K.M., Abers, G.A., Protti, M., González, V., Strauch, W., 2010. Constraints on upper mantle anisotropy surrounding the Cocos slab from SK (K) S splitting. J. Geophys. Res.: Solid Earth 115 (B6).

Aki, K., & Richards, P.G. (2002). Quantitative seismology.

Alsina, D., Snieder, R., 1995. Small-scale sublithospheric continental mantle deformation: constraints from SKS splitting observations. Geophys. J. Int. 123 (2), 431–448.

AlpArray Seismic Network, 2015, AlpArray Seismic Network (AASN) temporary component, AlpArray Working Group. Other/Seismic Network.

Backus, G.E., 1962. Long-wave elastic anisotropy produced by horizontal layering. J. Geophys. Res. 67 (11), 4427–4440.

Barruol, G., Bonnin, M., Pedersen, H., Bokelmann, G.H., Tiberi, C., 2011. Belt-parallel mantle flow beneath a halted continental collision: The Western Alps. Earth Planet. Sci. Lett. 302 (3-4), 429–438.

Bigi, G., Cosentino, D., Parotto, M., Sartori, R., & Scandone, P., 1992, Structural model of Italy, 1:500.000, CNR Progetto Finalizzato Geodinamica, 114(3).

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

Bonnans, J.F., Gilbert, J.C., Lemaréchal, C., Sagastizábal, C.A., 2006. Numerical optimization: theoretical and practical aspects. Springer Science & Business Media,.

Bousquet, R., Schmid, S.M., Zeilinger, G., Oberhänsli, R., Rosenberg, C., Molli, G., Robert, C., Wiederkehr, M., & Rossi, P., 2012. Tectonic framework of the Alps, CCGM/CGMW.

Bowman, J.R., Ando, M., 1987. Shear-wave splitting in the upper-mantle wedge above the Tonga subduction zone. Geophys. J. Int. 88 (1), 25–41.

Boyd, S.P., Vandenberghe, L., 2004. Convex Optimization. Cambridge university press,.

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

Broyden, C.G., 1967. Quasi-Newton methods and their application to function minimisation. Math. Comput. 21 (99), 368–381.

Chevrot, S., 2000. Multichannel analysis of shear wave splitting. J. Geophys. Res.: Solid Earth 105 (B9), 21579–21590.

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

Christensen, N.I., 1966. Shear wave velocities in metamorphic rocks at pressures to 10 kilobars. J. Geophys. Res. 71 (14), 3549–3556.

Crampin, S., 1987. The basis for earthquake prediction. Geophys. J. Int. 91 (2), 331–347.

Farra, V., Vinnik, L.P., Romanowicz, B., Kosarev, G.L., Kind, R., 1991. Inversion of teleseismic S particle motion for azimuthal anisotropy in the upper mantle: a feasibility study. Geophys. J. Int. 106 (2), 421–431.

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

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

Fernández-Martínez, J.L., Pallero, J.L.G., Fernández-Muñiz, Z., Pedruelo-González, L.M., 2014. The effect of noise and Tikhonov's regularization in inverse problems. Part I: The linear case. J. Appl. Geophys. 108, 176–185.

Fichtner, A., Trampert, J., 2011. Hessian kernels of seismic data functionals based upon adjoint techniques. Geophys. J. Int. 185 (2), 775–798.

Fletcher, R., Watson, G.A., 1980. First and second order conditions for a class of nondifferentiable optimization problems. Math. Program. 18, 291–307.

Froitzheim, N., Schmid, S.M., Frey, M., 1996. Mesozoic paleogeography and the timing of eclogite-facies metamorphism in the Alps: A working hypothesis. Eclogae Geol. Helv. 89 (1), 81.

Gao, S.S., Liu, K.H., 2012. AnisDep: A FORTRAN program for the estimation of the depth of anisotropy using spatial coherency of shear-wave splitting parameters. Comput. Geosci. 49, 330–333.

Goldfarb, D., 1976. Matrix factorizations in optimization of nonlinear functions subject to linear constraints. Math. Program. 10 (1), 1–31.

Handy, M.R., Schmid, S.M., Bousquet, R., Kissling, E., Bernoulli, D., 2010. Reconciling plate-tectonic reconstructions of Alpine Tethys with the geological–geophysical record of spreading and subduction in the Alps. Earth Sci. Rev. 102, 121–158.

Handy, M.R., Ustaszewski, K., Kissling, E., 2015. Reconstructing the Alps–Carpathians–Dinarides as a key to understanding switches in subduction polarity, slab gaps and surface motion. Int J. Earth Sci. 104, 1–26.

Handy, M.R., Giese, J., Schmid, S.M., Pleuger, J., Spakman, W., Onuzi, K., Ustaszewski, K., 2019. Coupled Crust–Mantle Response to Slab Tearing, Bending, and Rollback Along the Dinaride–Hellenide Orogen. Tectonics 38, 2803–2828.

Hein, G., Kolínský, P., Bianchi, I., Bokelmann, G., 2021. Shear wave splitting in the Alpine region. Geophys. J. Int. 227 (3), 1996–2015.

Heit, B., Weber, M., Tilmann, F., Haberland, C., Jia, Y., & Pesaresi, D. (2017). The Swath-D seismic network in Italy and Austria.

Hetényi, G., Plomerová, J., Bianchi, I., Exnerová, H.K., Bokelmann, G., Handy, M.R., AlpArray-EASI Working Group, 2018. From mountain summits to roots: Crustal structure of the Eastern Alps and Bohemian Massif along longitude 13.3 E. Tectonophysics 744, 239–255.

Huang, Z., Chevrot, S., 2021. Mantle dynamics in the SE Tibetan Plateau revealed by teleseismic shear-wave splitting analysis. Phys. Earth Planet. Inter. 313, 106687.

Karato, S.I., Jung, H., Katayama, I., Skemer, P., 2008. Geodynamic significance of seismic anisotropy of the upper mantle: New insights from laboratory studies. Annu. Rev. Earth Planet. Sci. 36, 59–95.

Kennett, B.L.N., Engdahl, E.R., Bulland, R., 1995. Constraints on Seismic Velocities in the Earth from Traveltimes. Geophys. J. Int. 122, 108–124. https://doi.org/10.1111/j.1365-246X.1995.tb03540.x.

Lin, Y.-P., Zhao, L., Hung, S.-H., 2014. Full-wave multiscale anisotropy tomography in Southern California. Geophys. Res. Lett. 41, 8809–8817. https://doi.org/10.1002/2014GL061855.

Link, F., Reiss, M.C., Rümpker, G., 2022. An automatized XKS-splitting procedure for large data sets: Extension package for SplitRacer and application to the USArray. Comput. Geosci. 158, 104961.

Link, F., Rümpker, G., 2021. Resolving Seismic Anisotropy of the Lithosphere–Asthenosphere in the Central/Eastern Alps Beneath the SWATH-D Network. Front. Earth Sci. 9, 679887.

Link, F., Rümpker, G., 2023. Shear-Wave Splitting Reveals Layered-Anisotropy Beneath the European Alps in Response to Mediterranean Subduction. J. Geophys. Res.: Solid Earth 128 (9) e2023JB027192.

Liu, K.H., Gao, S.S., 2013. Making reliable shear-wave splitting measurements. Bull. Seismol. Soc. Am. 103 (5), 2680–2693. https://doi.org/10.1785/0120120355.

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

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

Luo, Y., Long, M.D., Link, F., Karabinos, P., Kuiper, Y.D., 2023. Insights from layered anisotropy beneath southern New England: From ancient tectonism to present-day mantle flow (in press,). Geochem. Geophys. Geosyst.. https://doi.org/10.1029/2023GC011118.

Lüschen, E., Lammerer, B., Gebrande, H., Millahn, K., Nicolich, R., TRANSALP Working Group, 2004. Orogenic structure of the Eastern Alps, Europe, from TRANSALP deep seismic reflection profiling. Tectonophysics 388 (1-4), 85–102.

Mensch, T., Rasolofosaon, P., 1997. Elastic-wave velocities in anisotropic media of arbitrary symmetry—generalization of Thomsen's parameters ε, δ and γ. Geophys. J. Int. 128 (1), 43–64.

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

Mondal, P., Long, M.D., 2020. Strong seismic anisotropy in the deep upper mantle beneath the Cascadia backarc: Constraints from probabilistic finite-frequency SKS splitting intensity tomography. Earth Planet. Sci. Lett. 539, 116172.

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

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

Nur, A., Simmons, G., 1969. Stress-induced velocity anisotropy in rock: An experimental study. J. Geophys. Res. 74 (27), 6667–6674.

Petrescu, L., Pondrelli, S., Salimbeni, S., Faccenda, M., 2020. Mantle flow below the central and greater Alpine region: Insights from SKS anisotropy analysis at AlpArray and permanent stations. Solid Earth Discuss. 2020, 1–26.

Qorbani, E., Bianchi, I., Bokelmann, G., 2015. Slab detachment under the Eastern Alps seen by seismic anisotropy. Earth Planet. Sci. Lett. 409, 96–108.

Reiss, M.C., Rümpker, G., 2017. SplitRacer: MATLAB code and GUI for semiautomated analysis and interpretation of teleseismic shear-wave splitting. Seismol. Res. Lett. 88 (2A), 392–409.

Roberts, G.O., Tweedie, R.L., 1996. Exponential convergence of Langevin distributions and their discrete approximations. Bernoulli 2 (4), 341–363. https://doi.org/10.2307/3318418.

Rümpker, G., Ryberg, T., 2000. New "Fresnel-zone" estimates for shear-wave splitting observations from finite-difference modeling. Geophys. Res. Lett. 27 (13), 2005–2008.

Rümpker, G., Kaviani A, A., Link, F., Reiss, M.C., Komeazi, A., 2023. Testing observables for teleseismic shear-wave splitting inversions: ambiguities of intensities, parameters, and waveforms. Ann. Geophys. 66 this volume, doi:10.4401/ag-8870.

Rümpker, G., Silver, P.G., 1998. Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy. Geophys. J. Int. Volume 135 (Issue 3), 790–800. https://doi.org/10.1046/j.1365-246X.1998.00660.x.

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

Savage, M., Wessell, A., Teanby, N., Hurst, A., 2010. Automatic measurement of shear wave splitting and applications to time varying anisotropy at Mount Ruapehu volcano, New Zealand. J. Geophys. Res. 115 https://doi.org/10.1029/2010JB007722.

Schmid, S.M., Bernoulli, D., Fügenschuh, B., Matenco, L., Schefer, S., Schuster, R., Tischler, M., Ustaszewski, K., 2008. The Alpine-Carpathian-Dinaridic orogenic system: correlation and evolution of tectonic units. Swiss J. Geosci. 101, 139–183.

Schmid, S.M., Scharf, A., Handy, M.R., Rosenberg, C.L., 2013. The Tauern Window (Eastern Alps, Austria): a new tectonic map, with cross-sections and a tectonometamorphic synthesis. Swiss J. Geosci. 106, 1–32.

Shanno, D.F., 1978. Conjugate gradient methods with inexact searches. Math. Oper. Res. 3 (3), 244–256.

Shepard, D. (1968, January). A two-dimensional interpolation function for irregularly-spaced data. In Proceedings of the 1968 23rd ACM national conference (pp. 517–524).

Sieminski, A., Trampert, J., Tromp, J., 2009. Principal component analysis of anisotropic finite-frequency sensitivity kernels. Geophys. J. Int. 179 (2), 1186–1198.

Silver, P.G., Chan, W.W., 1991. Shear wave splitting and subcontinental mantle deformation. J. Geophys. Res.: Solid Earth 96 (B10), 16429–16454.

Silver, P.G., Savage, M.K., 1994. The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers. Geophys. J. Int. 119 (3), 949–963.

Silver, P.G., 1996. Seismic anisotropy beneath the continents: Probing the depths of geology. Annu. Rev. earth Planet. Sci. 24 (1), 385–432.

Skemer, P., Hansen, L.N., 2016. Inferring upper-mantle flow from seismic anisotropy: an experimental perspective. Tectonophysics 668, 1–14.

Tarantola, A., 1987. Inversion of travel times and seismic waveforms. Seism. Tomogr.: Appl. Glob. Seismol. Explor. Geophys. 135–157.

Teanby, N., Kendall, J., van der Baan, M., 2004. Automation of shear-wave splitting measurements using cluster analysis. Bull. Seismol. Soc. Am. 94, 453–463. https://doi.org/10.1785/0120030123.

Tromp, J., Tape, C., Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. Geophys. J. Int. 160 (1), 195–216.

Vidal, A.F., De Bortoli, V., Pereyra, M., Durmus, A., 2020. Maximum likelihood estimation of regularization parameters in high-dimensional inverse problems: An empirical bayesian approach part i: Methodology and experiments. SIAM J. Imaging Sci. 13 (4), 1945–1989. https://doi.org/10.1137/20M1339829.

Wagner, L.S., Long, M.D., 2013. Distinctive upper mantle anisotropy beneath the High Lava Plains and Eastern Snake River Plain, Pacific Northwest, USA. Geochem. Geophys. Geosyst. 14, 4647–4666. https://doi.org/10.1002/ggge.20275.

Wüstefeld, A., Al-Harrasi, O., Verdon, J.P., Wookey, J., Kendall, J.M., 2010. A strategy for automated analysis of passive microseismic data to image seismic anisotropy and fracture characteristics. Geophys. Prospect. 58, 755–773. https://doi.org/10.1111/j.1365-2478.2010.00891.x.

Yousef, B.M., Angus, D.A., 2016. When do fractured media become seismically anisotropic? Some implications on quantifying fracture properties. Earth Planet. Sci. Lett. 444, 150–159.

Yuan, H., Romanowicz, B., 2010. Depth dependent azimuthal anisotropy in the western US upper mantle. Earth Planet. Sci. Lett. 300 (3-4), 385–394.

Yuan, H., Levin, V., 2014. Stratified seismic anisotropy and the lithosphere-asthenosphere boundary beneath eastern North America. J. Geophys. Res.: Solid Earth 119 (4), 3096–3114.

Zhu, W., Xu, K., Darve, E., Beroza, G.C., 2021. A general approach to seismic inversion with automatic differentiation. Comput. Geosci. 151, 104751.