Probing depth and lateral variations of upper-mantle seismic anisotropy from full-waveform inversion of teleseismic body-waves

Stephen Beller, Sébastien Chevrot

Geophysical Journal International

Geophys. J. Int. (2020) **222,** 352–387
Advance Access publication 2020 April 20
GJI Seismology

doi: 10.1093/gji/ggaa069

Probing depth and lateral variations of upper-mantle seismic anisotropy from full-waveform inversion of teleseismic body-waves

Stephen Beller* and Sébastien Chevrot

GET, UMR 5563, Observatoire Midi Pyrenees, Université Paul Sabatier, CNRS, IRD, Toulouse, France. E-mail: sbeller@princeton.edu

Accepted 2020 January 23. Received 2020 January 13; in original form 2019 February 18

SUMMARY

While seismic anisotropy can potentially provide crucial insights into mantle dynamics, 3-D imaging of seismic anisotropy is still a challenging problem. Here, we present an extension of our regional full-waveform inversion method to image seismic anisotropy in the lithosphere and asthenosphere from teleseismic *P* and *S* waveforms. The models are parametrized in terms of density and the 21 elastic coefficients of the fourth-order elasticity tensor. The inversion method makes no *a priori* assumptions on the symmetry class or on the orientation of the symmetry axes. Instead, the elasticity tensors in the final models are decomposed with the projection method. This method allows us to determine the orientation of the symmetry axes and to extract the contributions of each symmetry class. From simple synthetic experiments, we demonstrate that our full-waveform inversion method is able to image complex 3-D anisotropic structures. In particular, the method is able to almost perfectly recover the general orientation of the symmetry axis or complex layered anisotropic models, which are both extremely challenging problems. We attribute this success to the joint exploitation of both *P* and *S* teleseismic waves, which constrain different parts of the elasticity tensor. Another key ingredient is the preconditioning of the gradient with an approximate inverse Hessian computed with scattering integrals. The inverse Hessian is crucial for mitigating the artefacts resulting from the uneven (mostly vertical) illumination of teleseismic acquisitions.

Key words: Tomography; Seismic anisotropy; Inversion; Lithosphere; Upper mantle.

1 INTRODUCTION

When mantle deforms, its most abundant constituent, olivine, develops lattice-preferred orientations. The crystallographic *a*-axes of olivine minerals tend to align along the shear direction, producing fabrics that have strong effects on the propagation of seismic waves. This important property of deformed olivine aggregates has been recognized for a long time, and since the pioneering study of azimuthal anisotropy in the Pacific by Raitt *et al.* (1969), a vast amount of literature has accumulated on this topic.

Seismic anisotropy can be studied with a variety of seismic phases, but by far the most popular are Rayleigh waves and SKS waves. Indeed, the variations of the phase velocity of Rayleigh waves with the propagation direction at both global (e.g. Debayle *et al.* 2005; Beghein *et al.* 2006; Marone & Romanowicz 2007) and regional (e.g. Endrun *et al.* 2011; Schaeffer *et al.* 2016) scales constrain large-scale patterns of azimuthal anisotropy. While surface waves may provide valuable constraints on large-scale anisotropic structures under oceanic plates, which are related to seafloor spreading and present-day plate motion (e.g. Wolfe & Silver 1998), imaging seismic anisotropy for example in subduction zones or beneath continental orogens remains challenging with this approach. Splitting of SKS waves has been also extensively used to characterize apparent seismic anisotropy in different tectonic environments (e.g. Silver 1996; Savage 1999; Fouch & Rondenay 2006; Long & Silver 2008; Long & Becker 2010; Long 2013; Long & Wirth 2013). However, owing to the quasi vertical incidence of SKS waves, these studies suffer from a poor vertical resolution.

To overcome this problem, tomographic methods relying on finite-frequency effects on a new seismic observable, the so-called 'splitting intensity' (Chevrot 2000), were developed (Chevrot 2006) and tested on a large data set of SKS waves recorded by the Southern California network (Monteiller & Chevrot 2011). Another study in the same region (Lin *et al.* 2014) relied on sensitivity kernels computed by normal mode summation (Zhao & Chevrot 2011) instead of the semi-analytical approach introduced by Favier *et al.* (2004). These two pionneering tomographic studies were able to obtain 3-D models of seismic anisotropy beneath southern California, but the depth distribution of seismic anisotropy in these models remained poorly constrained. An important limitation of these early

* Now at: Department of Geosciences, Princeton University, Guyot Hall, Princeton, NJ 08544, USA.

tomographic studies is that they relied on splitting intensity, which quantifies the difference in phase delays between the two quasi shear waves that are orthogonally polarized. In other words, even if they properly accounted for finite-frequency effects, they still suffered from the well-known limitations of traveltime (or phase) tomography. In addition, since they only considered shear wave splitting information, they only provided (incomplete) information on shear wave anisotropy since only a small fraction of the complete fourth-order elasticity tensor could be retrieved. For example, splitting intensity tomography only constrains the magnitude of shear wave anisotropy and the projection of the fast direction on the horizontal plane. It has very little sensitivity to its dip (Chevrot & Van Der Hilst 2003), and is also not sensitive to compressional wave anisotropy (Chevrot 2006).

In principle, full waveform inversion (FWI, Lailly 1983; Tarantola 1984) has the potential to dramatically improve the resolution of tomographic images at all scales (Tape *et al.* 2009; Fichtner *et al.* 2010; Zhu *et al.* 2012; Operto *et al.* 2015; Bozdağ *et al.* 2016; Clouzet *et al.* 2018), with a theoretical resolution limit of the order of the shortest propagating wavelength (e.g. Virieux & Operto 2009). This improved resolution potential stems from the exploitation of both the amplitude and phase of seismic waves, but also, in the case of teleseismic imaging, from the contribution of later arrivals that are reflected and/or converted on the main lithospheric discontinuities, such as the Moho (Bostock *et al.* 2001; Pageot *et al.* 2013; Monteiller *et al.* 2015; Beller *et al.* 2018a).

In the last few years, modelling of teleseismic wave propagation in 3-D regional models has become tractable, thanks to the development of new hybrid numerical methods (Roecker *et al.* 2010; Monteiller *et al.* 2013; Masson *et al.* 2013; Tong *et al.* 2014a; Masson & Romanowicz 2017; Beller *et al.* 2018b). The principle of these hybrid methods is to compute the incident teleseismic wavefield only once and to inject this wavefield inside a regional spectral-element grid (Komatitsch & Tromp 1999) in later simulations. The incident wavefield could be computed in a 3-D global earth model, for example with SPECFEM3D_GLOBE (Komatitsch & Tromp 2002) but for computational efficiency, 1-D earth models are usually considered. Available techniques to compute the incident wavefield are the direct solution method (DSM, Kawai *et al.* 2006; Monteiller *et al.* 2013; Wu *et al.* 2018), AxiSEM (Nissen-Meyer *et al.* 2014; Beller *et al.* 2018b) or the frequency–wavenumber method (Tong *et al.* 2014b, a). In that case, the influence of 3-D heterogeneities located outside the regional grid becomes an issue, but it has been shown that they have actually a limited impact on the resulting tomographic images (Masson & Romanowicz 2017). The reason for this is simply that because the regional domain is much smaller than the source-to-receiver distance, each receiver records waves that propagated through the same path outside the grid. Hence, these heterogeneities as well as attenuation affect similarly each seismic record. These common effects in seismic waveforms are thus absorbed by the source wavelet estimation producing a broaden apparent source wavelet. In lithospheric imaging studies, one can thus assume that these effects can be safely neglected. The first applications of FWI on teleseismic *P* wave records in the Pyrenees (Wang *et al.* 2016) and in the Alps (Beller *et al.* 2018b) demonstrated that these effects have indeed a limited imprint on teleseismic waveforms and that FWI outperforms classical passive imaging approaches.

The motivation of this study is to extend this FWI approach, which was so far limited to isotropic media, to the general anisotropic case. Our main purpose is to demonstrate that the same improvement can be expected for imaging seismic anisotropy than the one already achieved in isotropic tomography at the regional scale. The paper is organized as follows. We first present the theoretical aspects of FWI for anisotropic media, with a particular emphasis on the problem of parametrizing seismic anisotropy. We present the new ingredients that were incorporated into FWI. First, we now consider *S* and SKS waveforms into the inversion, in addition to *P* waveforms. The idea is to extend the sensitivity to a larger number of elasticity coefficients, exploiting in particular shear wave splitting and mode conversions (*P*-to-*S* and *S*-to-*P*). Our choice to keep the 21 elasticity coefficients allows us to make no *a priori* assumption regarding the symmetry class of the medium or the orientation of its symmetry axes. These informations are retrieved from the decomposition of the elasticity tensors obtained at each node of the tomographic grid, following the approach of Browaeys & Chevrot (2004). Second, we compute an approximate inverse Hessian with the Born approximation to initialize the *l*-BFGS iterative algorithm. This is a crucial ingredient to equalize the sensitivity to the different elastic parameters. Otherwise, the inversion mainly updates the parameters that describe the vertical propagation of *P* and *S* waves. In other words, the approximate inverse Hessian allows us to deconvolve the results of the inversion from the uneven and anisotropic ray coverage of teleseismic acquisitions. We designed different synthetic experiments to demonstrate that FWI allows us to retrieve both the local orientation of the main symmetry axes and the elasticity coefficients expressed in this local natural reference frame, from teleseismic three-component records of *P*, Sv and SKS waves. An important conclusion of these synthetic tests is the demonstration that FWI can constrain both the azimuth and dip of the symmetry axis and the stratification of seismic anisotropy.

2 ANISOTROPIC FWI

2.1 Full-waveform inversion

Let us define the waveform least-square misfit function $\mathcal{C}(\mathbf{m})$ as the sample-to-sample difference between the recorded seismic traces $\mathbf{d}_s(\mathbf{x}_r, t)$ and their synthetic counterparts $\mathbf{u}_s(\mathbf{x}_r, t; \mathbf{m})$ computed for a given earth model $\mathbf{m}$ (Tarantola 1984):

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \sum_s \sum_r \int_0^T \left|\left| \mathbf{u}_s(\mathbf{x}_r, t; \mathbf{m}) - \mathbf{d}_s(\mathbf{x}_r, t) \right|\right|_2^2 \mathrm{d}t, \tag{1}$$

where subscripts r and s indicate summation over receivers and sources, respectively. The problem consists in finding the model $\mathbf{m}$ that minimizes the objective function (1):

$$\mathbf{m} = \arg\min_{\mathbf{m}} \mathcal{C}(\mathbf{m}). \tag{2}$$

This problem is recast as a local optimization problem by linearizing eq. (1) around a current model $\mathbf{m}_k$ to determine an updated model

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \Delta\mathbf{m}_k \tag{3}$$

that decreases the misfit function. To second order, the model perturbation $\Delta\mathbf{m}_k$ can be derived from the normal equations (Nocedal & Wright 2006; Virieux & Operto 2009) and is given by

$$\Delta\mathbf{m}_k = -\alpha_k \mathcal{H}^{-1}(\mathbf{m}_k)\mathcal{G}(\mathbf{m}_k), \tag{4}$$

with $\mathcal{G}$ and $\mathcal{H}$, respectively, the gradient (first order derivative) and Hessian (second order derivative) of the misfit function with respect to model parameters $\mathbf{m}$, and α a step length. The Newton step (4) is solved using *l*-BFGS (for limited-memory Broyden–Fletcher–Goldfarb–Shanno, Nocedal 1980; Métivier & Brossier 2016), a quasi-Newton optimization algorithm that estimates iteratively the action of the inverse Hessian onto the gradient

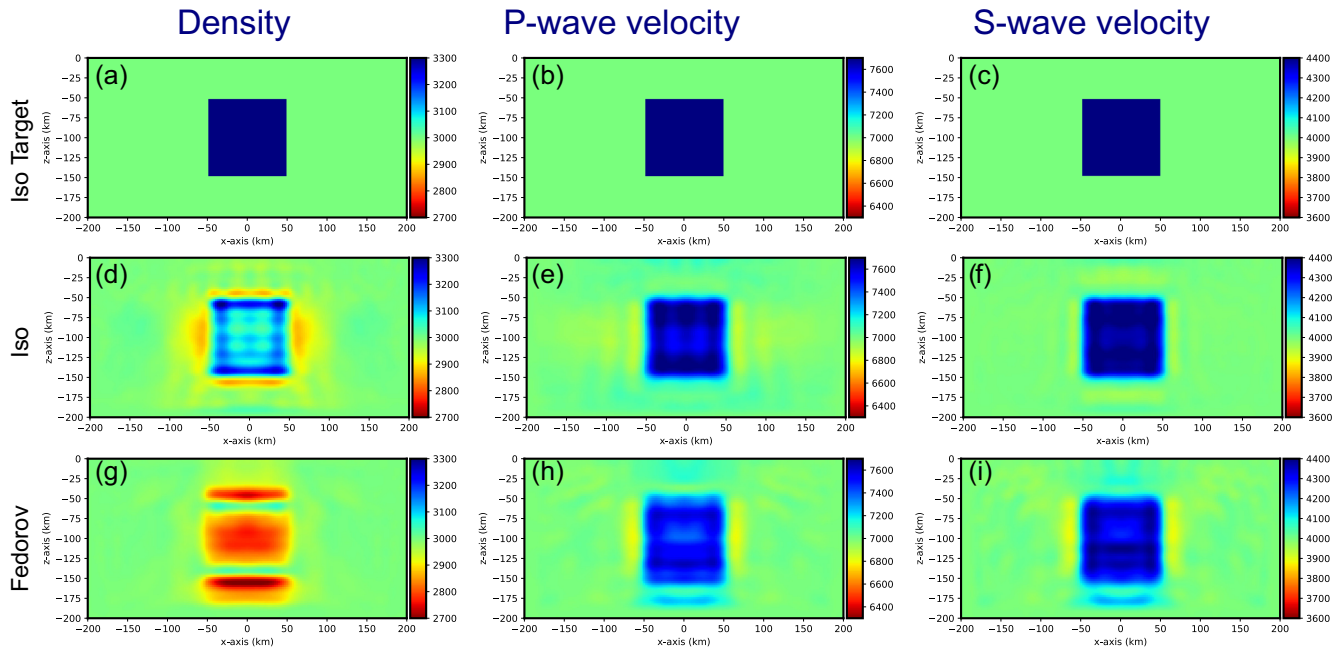

Figure 1. Isotropic FWI models for the isotropic inclusion test. (a–c) Density, P and S wave velocities in the target model. (d–f) Same for FWI models recovered from pure isotropic FWI. (g–i) Fedorov isotropic models obtained from pre-conditioned anisotropic FWI models.

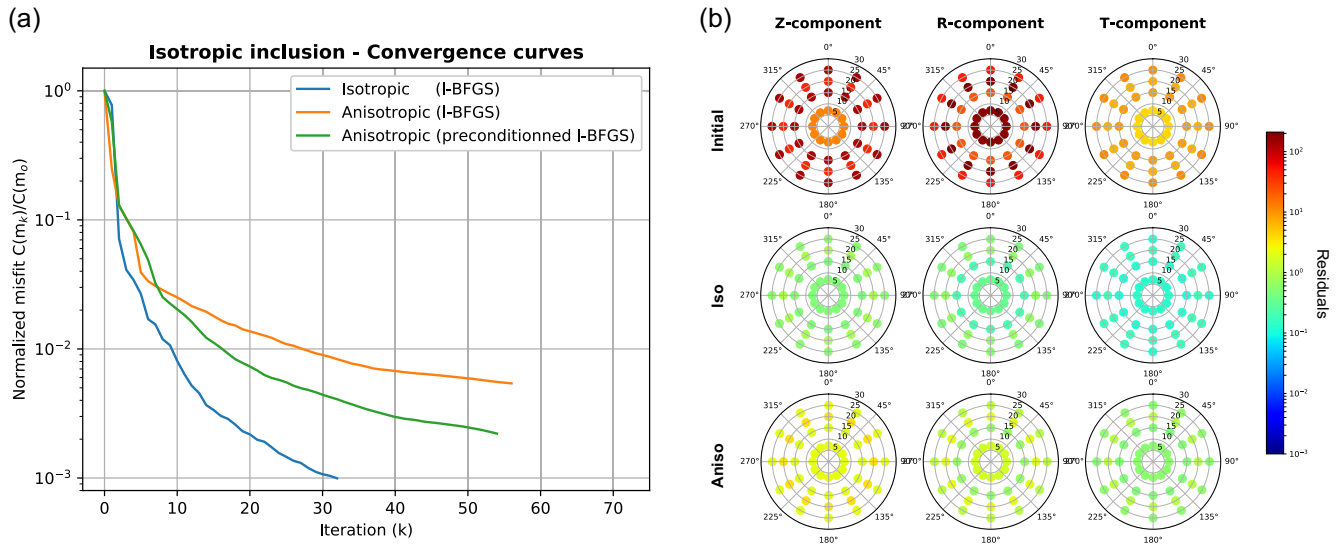

Figure 2. FWI results for the isotropic inclusion test. (a) Convergence curves for isotropic (blue), anisotropic (orange) and pre-conditioned anisotropic (green) FWI. (b) Incidence and azimuthal variation of data residuals: from left to right residuals of vertical, radial and transverse components, from top to bottom residuals in the initial model, final isotropic FWI model and final pre-conditioned anisotropic FWI model.

with a line-search algorithm based on Wolfe's conditions (Wolfe 1969).

2.2 Gradient computation

The gradient $\mathcal{G}(\mathbf{m})$ of the misfit function in (4) is given by

$$\mathcal{G} = \frac{\delta \mathcal{C}}{\delta \mathbf{m}}, \tag{5}$$

where $\delta\mathcal{C}$ the variation of the misfit function is (Tromp *et al.* 2005; Fichtner *et al.* 2006; Plessix 2006; Liu & Tromp 2006)

$$\delta \mathcal{C} = \sum_{s,r} \int_0^T \delta \mathbf{u}_s(\mathbf{x}_r, t; \mathbf{m}) \left[\mathbf{u}_s(\mathbf{x}_r, t; \mathbf{m}) - \mathbf{d}_s(\mathbf{x}_r, t) \right] \mathrm{d}t. \tag{6}$$

In the latter equation, the perturbation of the misfit function relies explicitly on the computation of the Fréchet derivatives $\delta \mathbf{u}_s(\mathbf{x}_r, t; \mathbf{m})$. However, instead of forming them explicitly, the gradient can be efficiently computed with the adjoint state method (Tarantola 1984; Tromp *et al.* 2005; Plessix 2006; Fichtner *et al.* 2006) with two simulations per source only. The first simulation

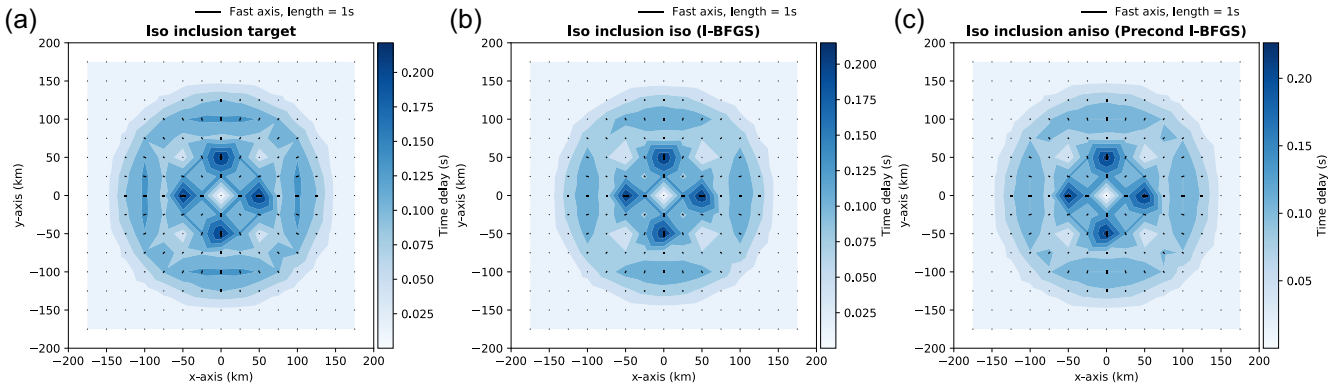

Figure 3. Map of predicted SKS splitting measurements obtained for the target isotropic model (a), the corresponding isotropic FWI model (b) and the anisotropic FWI model (c). The black segments represent the apparent splitting parameters, colourmap and segment lengths are proportional to measured time delays between quasi-Sv and quasi-Sh waves.

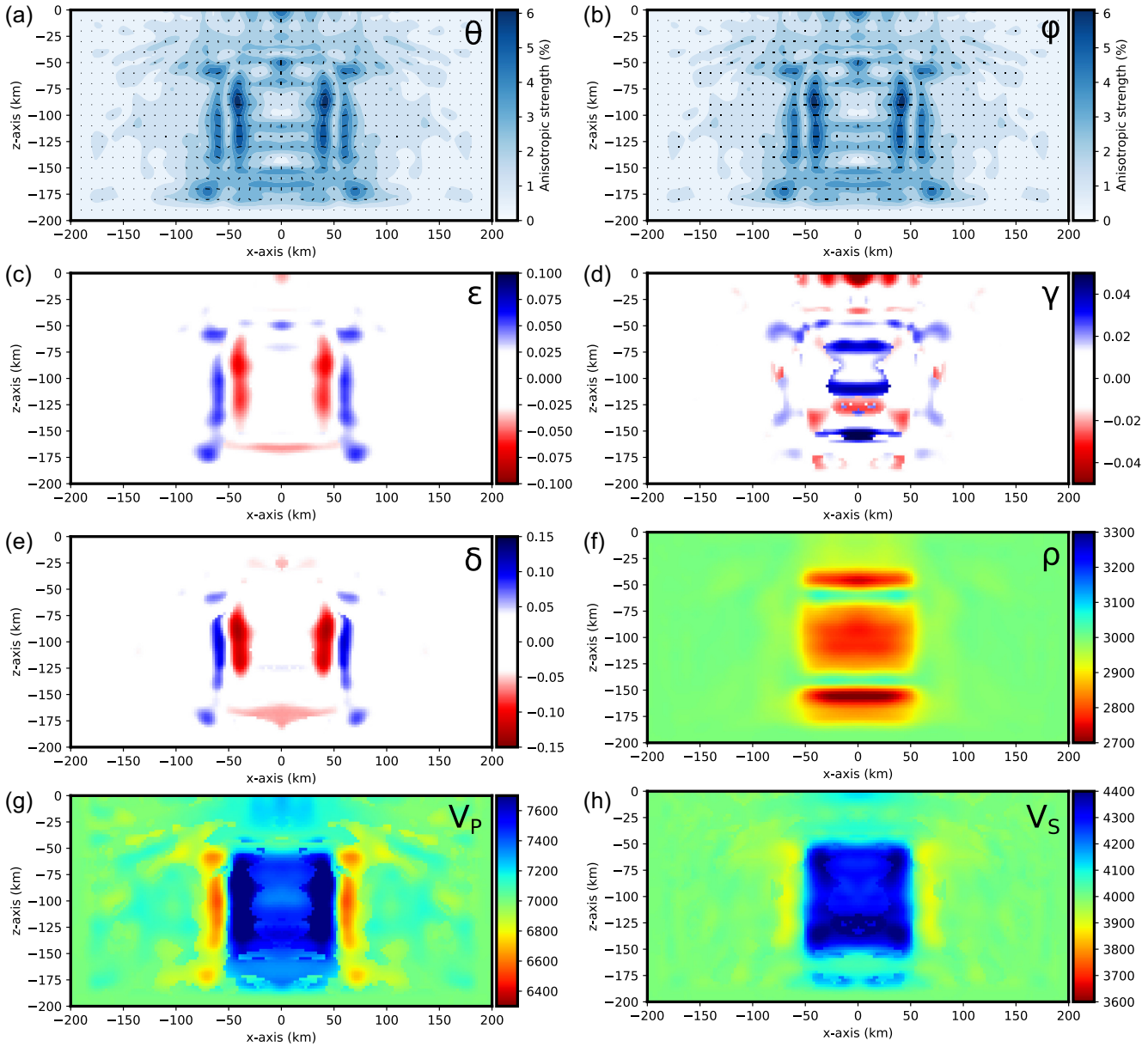

Figure 4. Pre-conditioned anisotropic FWI results for the isotropic inclusion experiment: (a, b) Dip and azimuth angles, (c–e) recovered anisotropic parameters ϵ, γ and δ, (f–h) density and hexagonal reference *P*- and *S*-waves velocities.

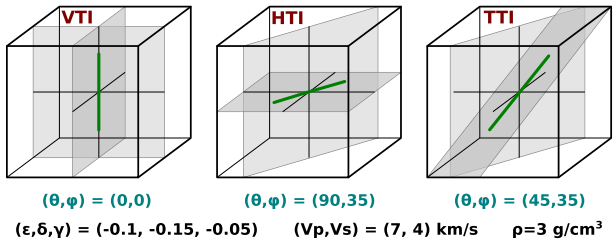

Figure 5. Orientation of hexagonal fast axes for the three anisotropic inclusion models. All inclusions have a 50 km cubic shape, with the same anisotropy parameters ϵ, δ and γ and are embedded in a homogeneous isotropic half-space. Each inclusion differs to the orientation (dip θ and azimuth ϕ) of the symmetry axis. From left to right, inclusions characterize VTI, HTI and TTI anisotropy.

computes the forward wavefield $\mathbf{u}$ by solving the equations of elastodynamics:

$$\rho(\mathbf{x})\partial_t^2\mathbf{u}(\mathbf{x},t) - \nabla\cdot[\mathbf{c}(\mathbf{x}):\nabla\mathbf{u}(\mathbf{x},t)] = f(\mathbf{x}_s,t) \quad \text{s.t.}\ \ \mathbf{u}(\mathbf{x},t=0)=0. \tag{7}$$

The second simulation computes the adjoint wavefield $\mathbf{u}^\dagger$, given by

$$\rho(\mathbf{x})\partial_t^2\mathbf{u}^\dagger(\mathbf{x},t) - \nabla\cdot[\mathbf{c}(\mathbf{x}):\nabla\mathbf{u}^\dagger(\mathbf{x},t)] = f^\dagger(\mathbf{x}_r,t) \quad \text{s.t.}\ \ \mathbf{u}^\dagger(\mathbf{x},t=T)=0, \tag{8}$$

which is equivalent to the previous problem except that the adjoint sources $f^\dagger$ correspond to the time-reversed data residuals injected at receiver locations.

Once eqs (7) and (8) are solved, the gradient is formed by the zero-lag correlation of the forward and adjoint wavefields. In an arbitrary triclinic elastic medium, the gradient with respect to density is obtained by cross-correlating the acceleration of the forward wavefield with the adjoint displacement:

$$\delta\mathcal{C}_\rho = \sum_s\int_t\int_V u_i^\dagger(\mathbf{x},t)\delta\rho(\mathbf{x})\partial_t^2u_i(\mathbf{x},t)\,\mathrm{d}\mathbf{x}^3\mathrm{d}t. \tag{9}$$

The gradients with respect to stiffness tensor moduli are obtained by correlating the state $\boldsymbol{\epsilon}$ and adjoint $\boldsymbol{\epsilon}^\dagger$ strain tensors:

$$\delta\mathcal{C}_\mathbf{c} = \sum_s\int_t\int_V \epsilon_{ij}^\dagger(\mathbf{x},t)\delta c_{ijkl}(\mathbf{x})\epsilon_{kl}(\mathbf{x},t)\,\mathrm{d}\mathbf{x}^3\mathrm{d}t. \tag{10}$$

For the particular case of elastic isotropy, the elasticity tensor can be described with only two independent parameters

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}), \tag{11}$$

where λ is the first Lamé parameter and μ the shear modulus. Inserting (11) into eq. (10) we obtain the gradients with respect to isotropic parameters:

$$\delta\mathcal{C}_\lambda = \sum_s\int_t\int_V \delta\lambda(\mathbf{x})\epsilon_{ii}^\dagger(\mathbf{x},t)\epsilon_{jj}(\mathbf{x},t)\,\mathrm{d}\mathbf{x}^3\mathrm{d}t, \tag{12}$$

and

$$\delta\mathcal{C}_\mu = \sum_s\int_t\int_V \delta\mu(\mathbf{x})2\epsilon_{ij}^\dagger(\mathbf{x},t)\epsilon_{ij}(\mathbf{x},t)\,\mathrm{d}\mathbf{x}^3\mathrm{d}t. \tag{13}$$

2.3 Diagonal Hessian pre-conditioning

We solve the normal eq. (4) with an *l*-BFGS quasi-Newton algorithm which implicitly accounts for the action of the inverse Hessian onto the gradient to determine the optimal search direction. The role of the Hessian is crucial in FWI (Pratt *et al.* 1998; Virieux & Operto 2009; Pan *et al.* 2016). From a pure optimization point of view, the Hessian improves the convergence rate of the algorithm by accounting for the local curvature of the misfit function. Physically, the linear part of the Hessian can be seen as a deconvolution operator which corrects the gradient from limited bandwidth effects related to the source–receiver acquisition (limited wavenumber illumination and geometrical spreading), rescales it to account for parameters dimensionality and reduces interparameter cross-talks. The non-linear part of the Hessian mostly accounts for double scattering effects (Pratt *et al.* 1998).

The *l*-BFGS algorithm approximates the inverse of the true Hessian by finite-differences. However, its first iteration relies on a (sometimes pre-conditioned) steepest descent step, hence possibly driving the solution towards a local attraction basin. The algorithm thus requires a good initial approximation of the inverse Hessian $\mathcal{H}_o^{-1}$ (Nocedal & Wright 2006, p.140). Provided that the Hessian possesses a dominant band-diagonal structure, a good choice for $\mathcal{H}_o^{-1}$ is the inverse of the diagonal terms of the Gauss–Newton Hessian (Jin *et al.* 1992; Alkhalifah & Plessix 2014). In that case, the diagonal terms of $\mathcal{H}_o^{-1}$ correspond to the autocorrelation of the Fréchet derivatives, which can be evaluated via a scattering integral method (e.g. Chen *et al.* 2007):

$$\mathbf{J} = \delta u_i^s(\mathbf{x}_r,t) = -\int_T\int_{-\infty}^{+\infty}(K_\rho + K_\mathbf{c})\,\mathrm{d}\tau\mathrm{d}t, \tag{14}$$

with

$$K_\rho = \delta\rho(\mathbf{x})\,G_{ik}(\mathbf{x}_r,t;\mathbf{x},t-\tau)\,\partial_\tau^2u_k^s(\mathbf{x},\tau) \tag{15}$$

and

$$K_\mathbf{c} = -\partial_kG_{in}(\mathbf{x}_r,t;\mathbf{x},t-\tau)\,\delta c_{knml}(\mathbf{x})\,\partial_mu_l^s(\mathbf{x},\tau), \tag{16}$$

where i corresponds to the receiver component index, $G_{ik}(\mathbf{x}_r,t;\mathbf{x},t-\tau)$ the receiver Green's functions and $u_k^s(\mathbf{x},\tau)$ the incident source wavefield. Computing $\mathbf{J}$ from (14) requires a number of forward modelling simulations given by the number of receiver components times the number of sources. The diagonal preconditioner $\mathcal{H}_o^{-1}$ is then evaluated from:

$$\mathcal{H}_o^{-1} = \mathrm{diag}(\mathbf{J}^t\mathbf{J}+\zeta)^{-1}, \tag{17}$$

where ζ is a small regularization term introduced to stabilize the inversion. To overcome the burden of computing (14) numerically, we consider an approximation of $\mathbf{J}$ obtained by taking the asymptotic form of Green's functions inside an homogeneous half-space (Jin *et al.* 1992).

3 SEISMIC ANISOTROPY PARAMETRIZATION

From now on, and for simplicity, we will consider a right-handed Cartesian coordinate system $(x_1,x_2,x_3)=(E,N,Z)$.

3.1 Problem statement

In general, the number of independent parameters characterizing a particular symmetry class differs from the number of non-zero coefficients of the elasticity tensor expressed in a randomly chosen reference frame (Babuska & Cara 1991). In the highest symmetry class, that is isotropy, the elasticity tensor is described by only two independent parameters, the Lamé parameters λ and μ, which are related to isotropic *P*- and *S*-wave velocities through:

$$v_P = \sqrt{\tfrac{\lambda+2\mu}{\rho}} \tag{18}$$

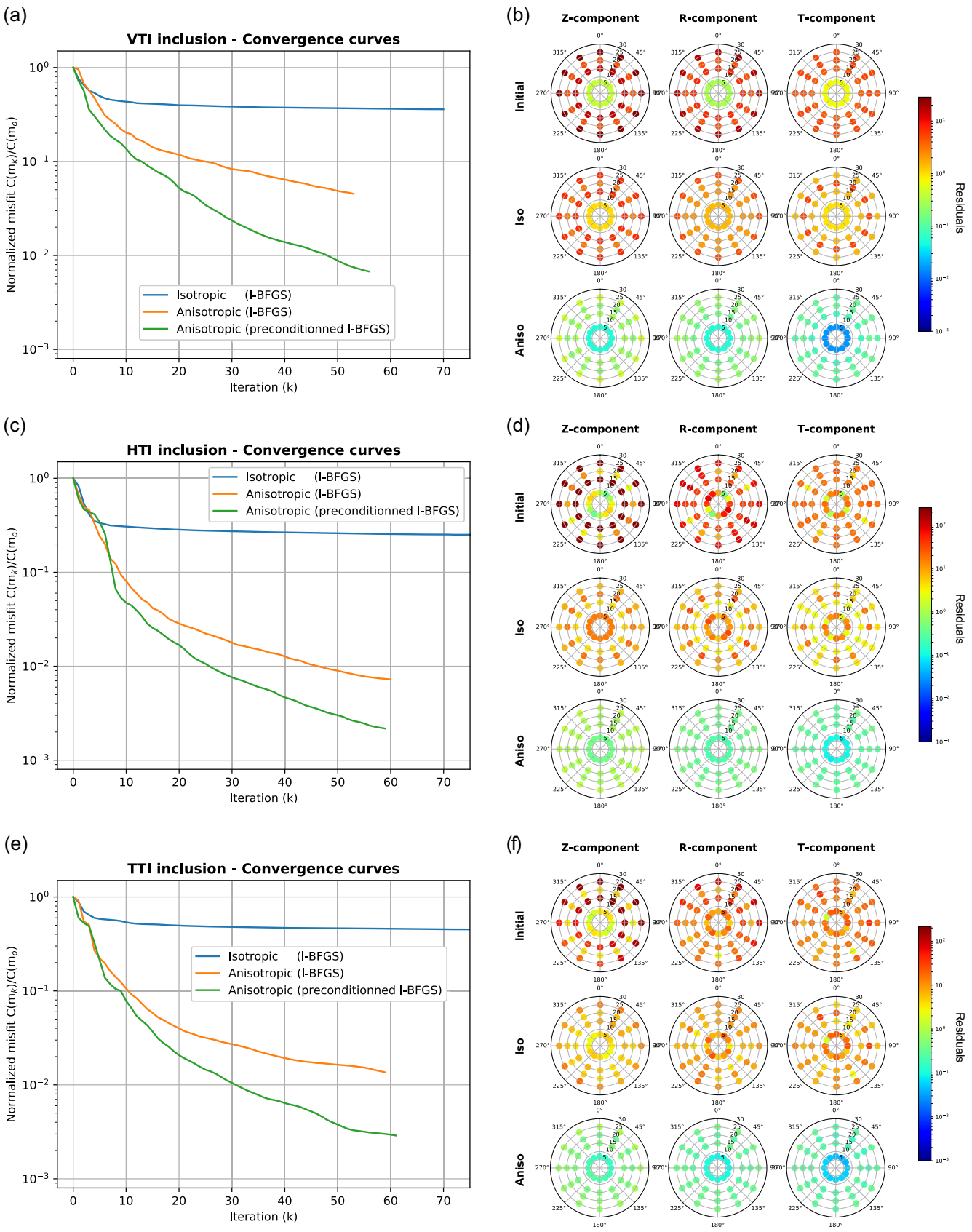

Figure 6. FWI results for hexagonal inclusion tests. Same as Fig. 2 for (a–b) the VTI inclusion experiment, (c–d) the HTI inclusion experiment and (e–f) the TTI inclusion experiments.

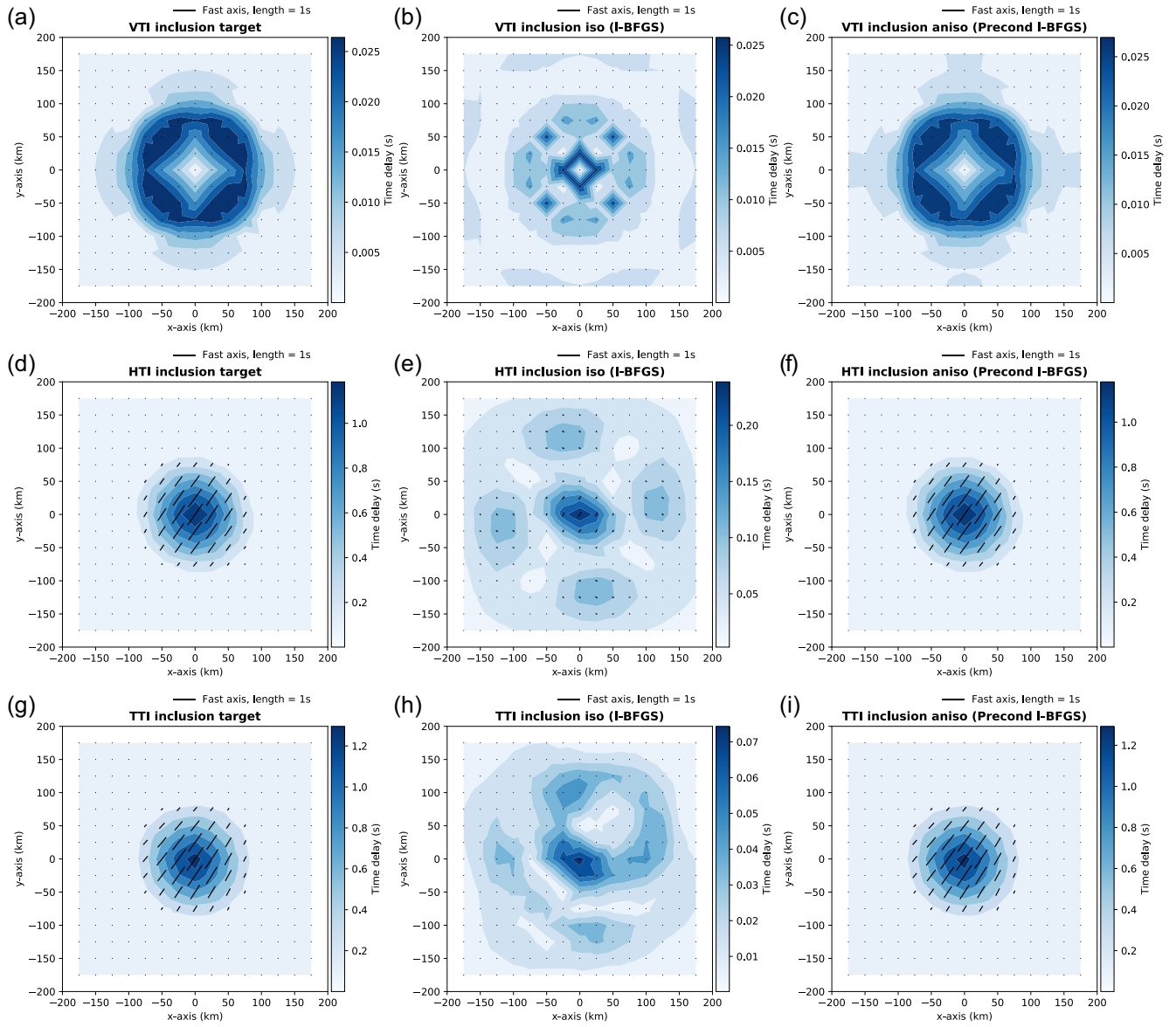

Figure 7. Map of predicted SKS splitting measurements obtained for the target (VTI, HTI and TTI) hexagonal inclusions models (a, d and g), the corresponding isotropic FWI model (b, e and h) and the anisotropic FWI model (c, f and i).

$$v_S = \sqrt{\frac{\mu}{\rho}}. \tag{19}$$

The isotropic elasticity tensor defined by a $\rho = 3\ \mathrm{g\,cm^{-3}}$, $v_P = 7\ \mathrm{km\,s^{-1}}$ and $v_S = 4\ \mathrm{km\,s^{-1}}$ has 9 non-zero components:

$$\mathbf{c}_{\mathrm{iso}} = \begin{pmatrix} 147 & 51 & 51 & 0 & 0 & 0 \\ 51 & 147 & 51 & 0 & 0 & 0 \\ 51 & 51 & 147 & 0 & 0 & 0 \\ 0 & 0 & 0 & 48 & 0 & 0 \\ 0 & 0 & 0 & 0 & 48 & 0 \\ 0 & 0 & 0 & 0 & 0 & 48 \end{pmatrix} \tag{20}$$

with the elasticity coefficients expressed in GPa. Since isotropic tensors are rotationally invariant, the non-zero elasticity coefficients are also rotationally invariant.

In contrast, the coefficients of elasticity tensors of lower symmetry classes (e.g. hexagonal or orthorhombic) depend on the orientation of the symmetry planes or symmetry axes. For example, an hexagonal tensor is described by the isotropic P and S velocities, the three Thomsen's parameters (Thomsen 1986; Mensch & Rasolofosaon 1997):

$$\epsilon = \frac{C_{11} - C_{33}}{2C_{33}},\ \gamma = \frac{C_{66} - C_{44}}{2C_{44}} \text{ and } \delta = \frac{C_{13} - C_{33} + 2C_{44}}{C_{33}}, \tag{21}$$

and the orientation of the symmetry axis, characterized by two angles (azimuth and dip). By choosing the reference hexagonal P- and S-waves velocities

$$v_P = \sqrt{\frac{C_{33}}{\rho}},\ v_S = \sqrt{\frac{C_{44}}{\rho}}, \tag{22}$$

respectively equal to 7 and 4 $\mathrm{km\,s^{-1}}$, density ρ to 3 $\mathrm{g\,cm^{-3}}$ and Thomsen's parameter ϵ, δ and γ to -10, -15 and -5 per cent, respectively, the resulting tensor expressed in its intrinsic coordinate

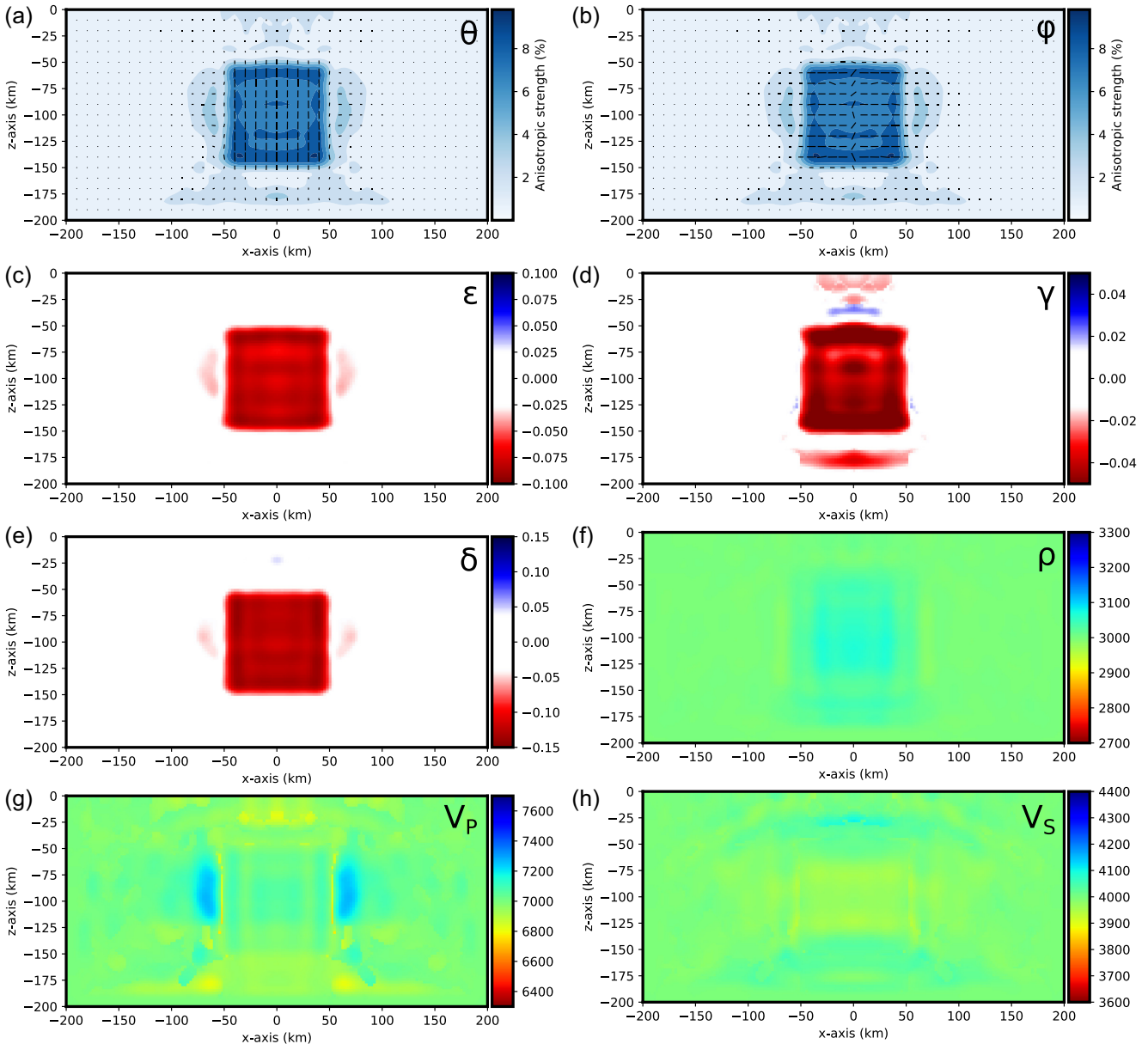

Figure 8. Pre-conditioned anisotropic FWI results for the VTI target inclusion model. (a, b) Dip and azimuth angles. (c–e) Recovered anisotropic parameters ϵ, γ and δ. (f–h) Density and hexagonal reference *P*- and *S*-waves velocities.

system (symmetry axis along x_3) is

$$\mathbf{c}^{\mathrm{VTI}} = \begin{pmatrix} 117.6 & 31.2 & 28.95 & 0 & 0 & 0 \\ 31.2 & 117.6 & 28.95 & 0 & 0 & 0 \\ 28.95 & 28.95 & 147 & 0 & 0 & 0 \\ 0 & 0 & 0 & 48 & 0 & 0 \\ 0 & 0 & 0 & 0 & 48 & 0 \\ 0 & 0 & 0 & 0 & 0 & 43.2 \end{pmatrix}. \tag{23}$$

Note that throughout the paper 'intrinsic' will either refer to the coordinate reference frame that is oriented along the principal symmetry axes of the elasticity tensor or to the elastic coefficients expressed in this intrinsic reference frame. It must not be confused for 'intrinsic anisotropy' as opposed to 'extrinsic anisotropy' as defined for example in Alder *et al.* (2017).

From now on, such hexagonal tensors with a vertical symmetry axis will be referred to as vertical transverse isotropy (VTI). As in the isotropic case, the number of non-zero elasticity coefficients is equal to 9. However, when the symmetry axis is horizontal (Horizontal Transverse Isotropy or HTI), the number of non-zero coefficients increases from 9 to 11. For example, an HTI medium with a fast axis azimuth of 30° of is given by:

$$\mathbf{c}^{\mathrm{HTI}} = \begin{pmatrix} 124.02 & 32.19 & 30.45 & 0 & 0 & 5.72 \\ 32.19 & 134.08 & 29.69 & 0 & 0 & 8.08 \\ 30.45 & 29.69 & 117.6 & 0 & 0 & -1.05 \\ 0 & 0 & 0 & 46.42 & 0 & 0 \\ 0 & 0 & 0 & 0 & 44.77 & 0 \\ 5.72 & 8.08 & -1.05 & 0 & 0 & 51.24 \end{pmatrix}. \tag{24}$$

For an arbitrary orientation of the symmetry axis, hereafter referred to as tilted transverse isotropy (TTI), the elasticity tensor becomes full with 21 non-zero coefficients. For example, considering an

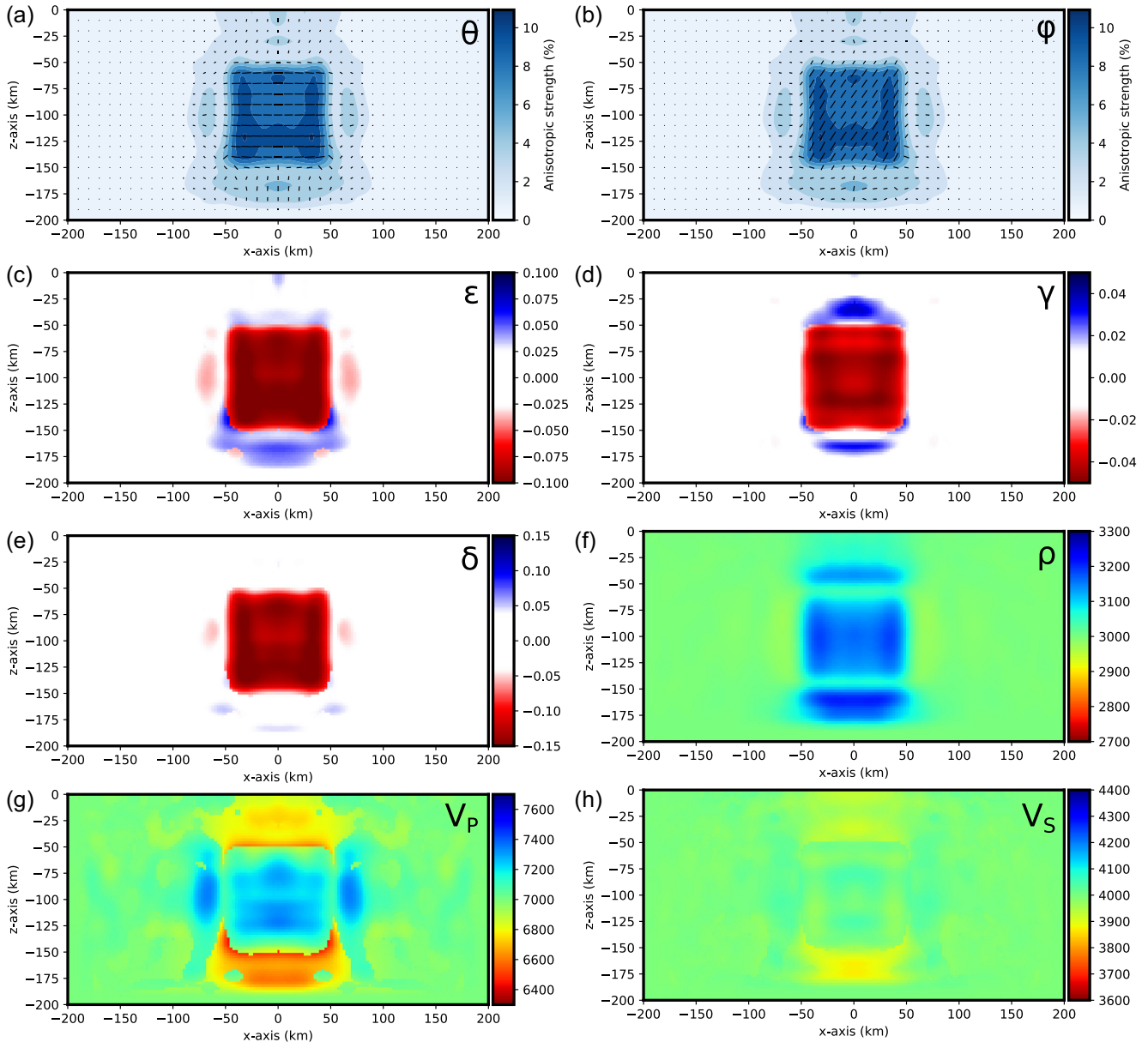

Figure 9. Pre-conditioned anisotropic FWI results for the HTI target inclusion model. (a, b) Dip and azimuth angles. (c–e) Recovered anisotropic parameters ϵ, γ and δ. (f to h) Density and hexagonal reference *P*- and *S*-waves velocities.

azimuth of 30° and a dip of 45° leads to:

$$\mathbf{c}^{\mathrm{TTI}} = \begin{pmatrix} 120.41 & 30.88 & 30.91 & 0.06 & 2.80 & 2.29 \\ 30.88 & 124.18 & 31.78 & 5.03 & 0.76 & 2.88 \\ 30.91 & 31.78 & 128.62 & 6.02 & 4.21 & 1.19 \\ 0.06 & 5.03 & 6.02 & 49.67 & 2.85 & 2.79 \\ 2.80 & 0.76 & 4.21 & 2.85 & 47.59 & 2.95 \\ 2.29 & 2.88 & 1.19 & 2.79 & 2.95 & 46.41 \end{pmatrix}. \quad (25)$$

This simple example demonstrates that the number of non-zero elasticity coefficients depends both on the symmetry class and on the orientation of the elasticity tensor. It also demonstrates that it is necessary to keep the 21 elasticity coefficients to describe even a simple hexagonal tensor with a general orientation of its symmetry axis. This implies that we need to invert for the 21 elasticity coefficients even if we only want to recover the hexagonal part of a 3-D anisotropic medium. Note that different complete sets of elastic parameters could be considered. For example, Rusmanugroho *et al.* (2015) chose to use the parameters defined in Chen & Tromp (2007), which are commonly used in global adjoint tomography (Sieminski *et al.* 2007a, b, 2009). However, Rusmanugroho *et al.* (2017) have shown that choosing either the Chen & Tromp (2007) or the Voigt parameters (the C_{ij} coefficients) leads to similar results. From now on, and for simplicity, we will thus consider a parametrization of seismic anisotropy involving the 21 Voigt parameters.

Many other simpler parametrizations of seismic anisotropy can be found in the literature. They all share the motivation to reduce the number of free parameters involved in the description of seismic anisotropy, and hence the dimensionality of the inverse problem. However, this comes at the cost of making strong *a priori* assumptions on the nature of seismic anisotropy. For example, it is common practice to exploit surface waves to invert for azimuthal anisotropy,

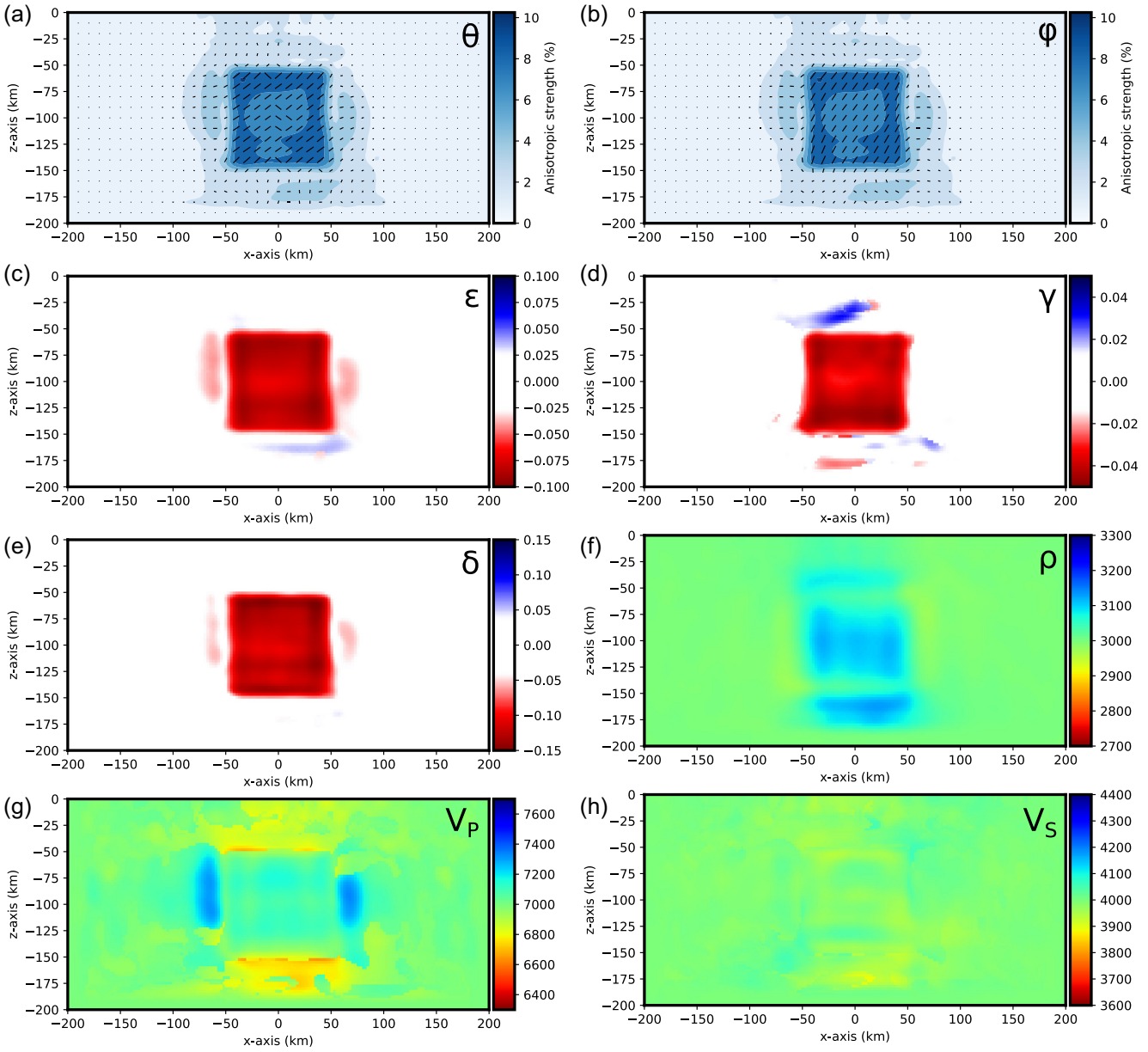

Figure 10. Pre-conditioned anisotropic FWI results for the TTI target inclusion model. (a, b) Dip and azimuth angles. (c–e) Recovered anisotropic parameters ϵ, γ and δ. (f–h) Density and hexagonal reference *P*- and *S*-waves velocities.

a simple type of seismic anisotropy described by a magnitude of anisotropy and an azimuth of fast propagation direction (e.g. Simons *et al.* 2002; Yao *et al.* 2010). It is also possible to enforce the symmetry class (e.g. hexagonal or orthorhombic) and tensor orientation and only invert for the intrinsic—natural—parameters, a common approach in exploration seismology (e.g. Gholami *et al.* 2013; Alkhalifah & Plessix 2014). Note that directly inverting for the angles (azimuth and dip) that describe the orientation of symmetry axes invariably leads to strongly non-linear inverse problems. These problems are untractable with gradient or even quasi-Newton approaches, unless some strong simplifying assumptions are made (e.g. assuming that the symmetry axis lies in the horizontal plane). An alternative would be to systematically explore the model space with a stochastic approach (e.g. Monte Carlo algorithm) as in Mondal & Long (2019). However, such methods can only handle inverse problems with a rather small number of free parameters. To summarize, while apparently simpler alternatives do exist for parametrizing seismic anisotropy, they are not adapted for imaging 3-D general anisotropic media.

3.2 Stiffness tensor decomposition

Retrieving the symmetry properties of a general fourth-order elasticity tensor is not straightforward. Fedorov (1968) first derived the closest isotropic approximation of an arbitrary elasticity tensor:

$$\kappa = \frac{C_{11} + C_{22} + C_{33} + 2(C_{23} + C_{13} + C_{12})}{9}, \tag{26}$$

$$\mu = \frac{(C_{11} + C_{22} + C_{33} - C_{23} - C_{13} - C_{12}) + 3(C_{44} + C_{55} + C_{66})}{15}, \tag{27}$$

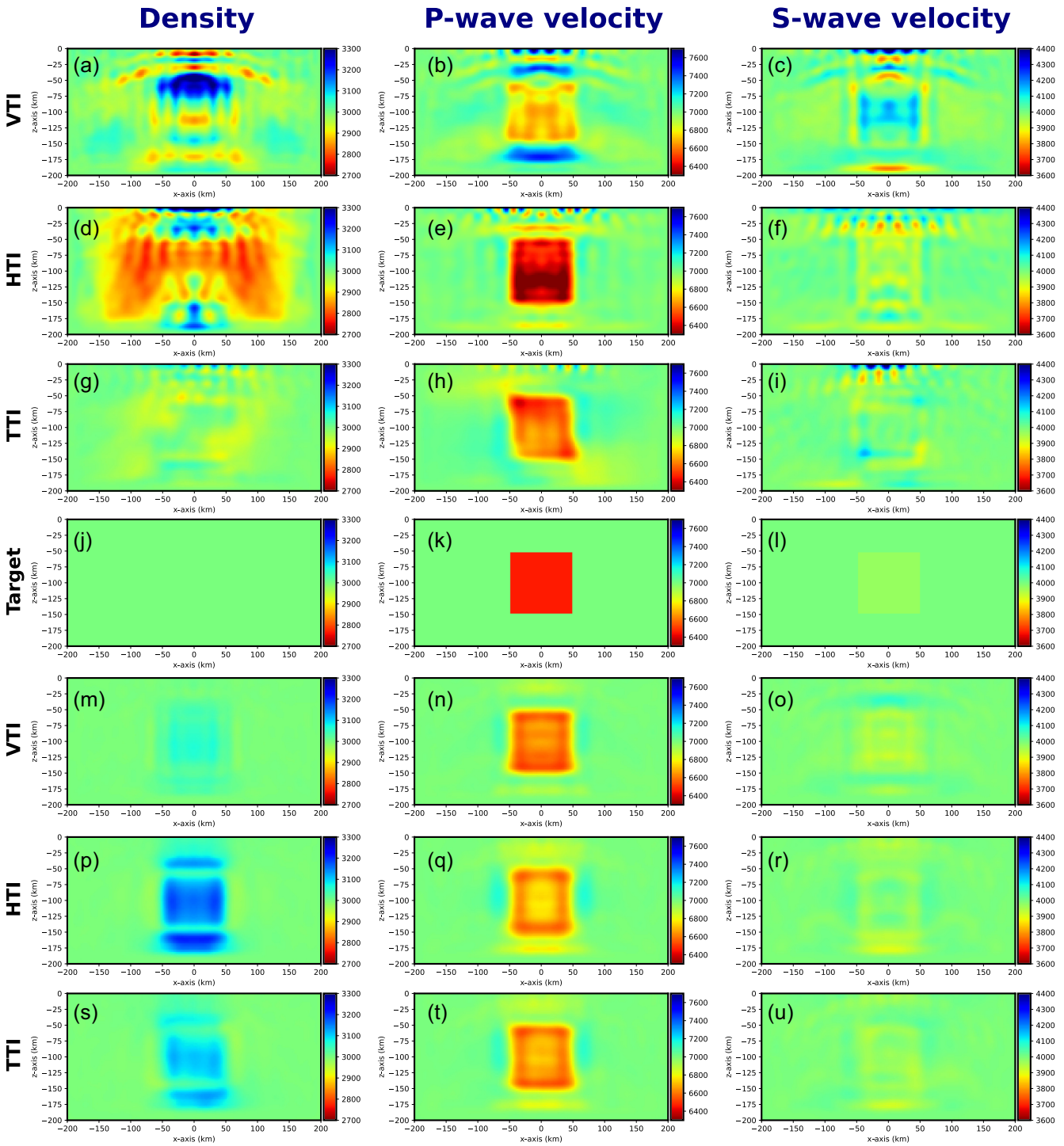

Figure 11. Isotropic FWI models for the hexagonal inclusions experiments. From left to right: density, *P*- and *S*-wave velocities. (a–i), models resulting from pure isotropic FWI of a VTI (a–c), HTI (d–f) and TTI (g–i) hexagonal inclusions. (j–u) Fedorov isotropic averages of hexagonal target model (j–l) and pre-conditioned anisotropic FWI from (m–o) VTI, (p–r) HTI and (s–u) TTI hexagonal inclusion experiments.

where κ and μ are the bulk and shear moduli, respectively. Several authors later generalized this result for lower symmetry classes (Helbig 1995; Browaeys & Chevrot 2004; Moakher & Norris 2006; Norris 2006). The projection method (Browaeys & Chevrot 2004) is equivalent to performing (Voigt) angular average of phase velocities as demonstrated by Norris (2006). Therefore, it seems to be a natural choice in seismic tomography.

The decomposition method (Helbig 1995; Browaeys & Chevrot 2004) starts by recasting the fourth-order elasticity tensor as a vector in a 21-D space. The 21 basis vectors correspond to 21 elementary fourth-order elasticity tensors. In this space, any elasticity tensor can be decomposed into a sum of orthogonal tensors belonging to classes of higher and higher degrees of symmetry (triclinic, monoclinic, orthorhombic, tetragonal, hexagonal and isotropic)

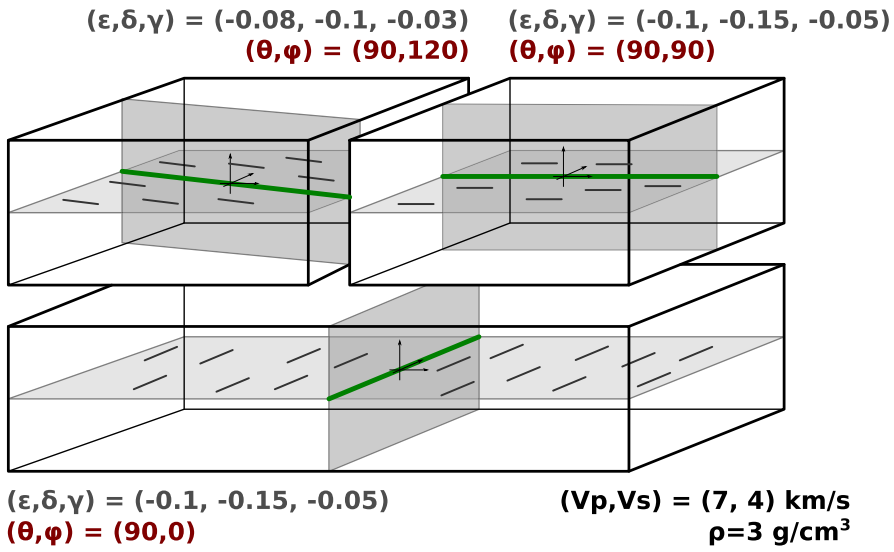

Figure 12. Geometry of the three-block target model. Each block contains an HTI anisotropic perturbation with respect to an homogeneous isotropic background model. The bottom layer is characterized by N–S fast direction. The upper-right block has a fast axis oriented at 90°, that is orthogonal to the one in the bottom layer, and with the same amount of anisotropy. The upper-left block contains a slightly different amount of anisotropy and has a fast axis oriented at 120°.

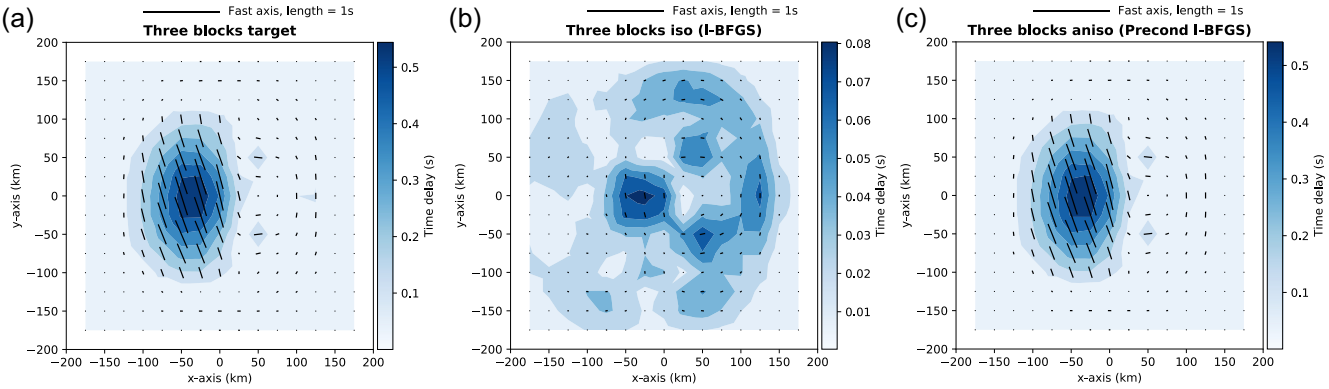

Figure 13. Map of predicted SKS splitting measurements obtained for the three blocks model (a), the corresponding isotropic FWI model (b) and the anisotropic FWI model (c). Black segments represent the apparent splitting parameters. Colour background segment lengths are proportional to measured time delays between quasi-Sv and quasi-Sh waves.

such that

$$\mathbf{C} = \mathbf{C}_{\text{tri}} + \mathbf{C}_{\text{mon}} + \mathbf{C}_{\text{ort}} + \mathbf{C}_{\text{tet}} + \mathbf{C}_{\text{hex}} + \mathbf{C}_{\text{iso}}. \tag{28}$$

Each of these orthogonal tensors belongs to a subspace that is independent (orthogonal) to the others. A higher symmetry approximation is therefore obtained by projecting the original tensor onto the corresponding subspace. Orthogonal projectors and expressions of projection matrices to go from one symmetry class to another are given in Browaeys & Chevrot (2004, Appendix A). Once the different symmetry parts have been obtained from these orthogonal projections, it is possible to compute their respective norms and thus the relative contribution of the different symmetry classes to the full tensor.

However, in order to be able to apply this tensor decomposition, one first needs to express the elasticity tensor in its natural or intrinsic reference frame. In other words, we need to apply a solid rotation to the fourth-order elasticity tensor in order to align the coordinate axes with its symmetry axes. The orientation of the symmetry axes can be determined from the eigenvectors of the dilatational $d_{ij} = c_{ijkk}$ and Voigt $v_{ik} = c_{ijkj}$ tensors which are second-order contractions of the fourth-order elasticity tensor (Cowin & Mehrabadi 1987). For a general anisotropic tensor, the three vectors of the Cartesian coordinate system are chosen as the three bisectors of the closest eigenvectors of the dilatational and Voigt tensors (Browaeys & Chevrot 2004). An example of decomposition for a triclinic elasticity tensor is given in Appendix B.

4 SYNTHETIC EXPERIMENTS

We will now present several synthetic experiments to illustrate the potentials and pitfalls of FWI to image seismic anisotropy in the lithosphere and/or asthenosphere from records of teleseismic *P* and *S* waves.

4.1 FWI experimental setup

Forward modelling is performed using a grid-injection technique (Chevrot *et al.* 2004; Monteiller *et al.* 2013; Beller *et al.* 2018b) to mimic the propagation of an incident teleseismic wavefield within a regional domain. The incident wavefield imposed on the boundaries of the grid is taken as an incident plane-wave propagating in a semi-infinite homogeneous medium (Aki & Richards

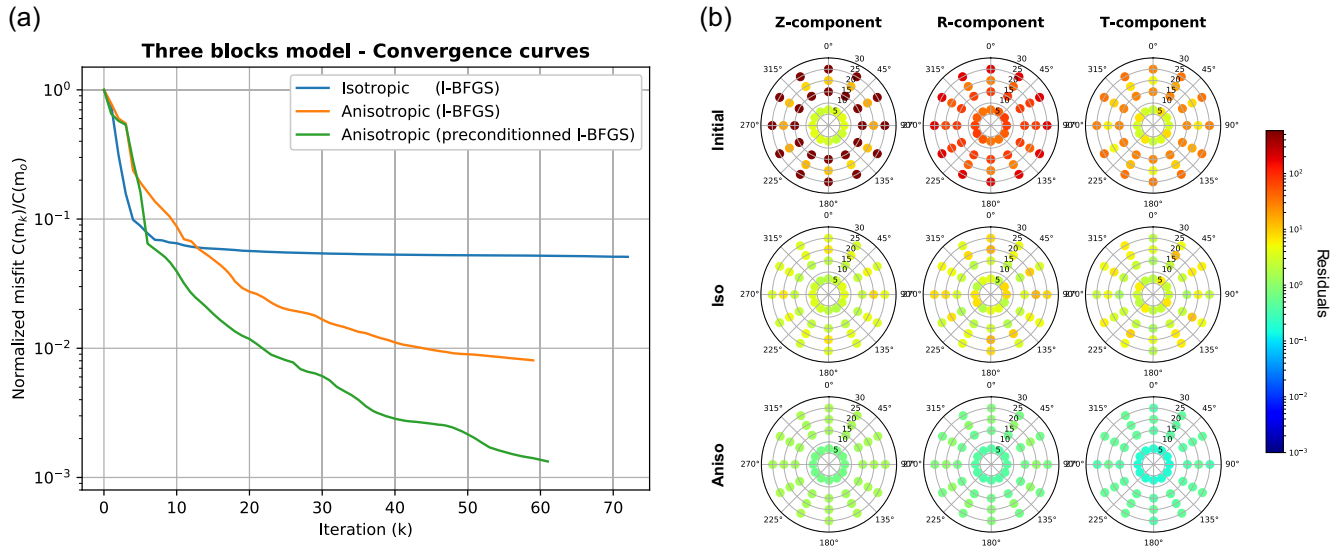

Figure 14. FWI results for the three-block model. (a) Convergence curves for isotropic (blue), anisotropic (orange) and pre-conditioned anisotropic (green) FWI. (b) Incidence and azimuthal variation of data residuals: from left to right residuals of vertical, radial and transverse components, from top to bottom residuals in the initial model, final isotropic FWI model and final pre-conditioned anisotropic FWI model.

2002). The wavefield in the regional domain is computed with a spectral-element method (Komatitsch & Tromp 1999) in a mesh containing 32 000 P5 hexahedral elements of 10 km size, leading to a total of 4×10^6 degrees of freedom. The overall size of the mesh is 400×400×200 km along the *x*-, *y*- and *z*-axes, respectively. The duration of one simulation is 150 s for a time step of 0.05 s.

In our experiments the data set is composed of *P* waves with incidence angles of 15° and 25°, Sv wave with incidence angles of 20°, and SKS waves with incidence angle of 7°, coming from 12 azimuths taken evenly from 0 to 330° azimuth. These waves are recorded by 225 three-component sensors placed on a regular 2-D grid with a spacing of 25 km, which covers a square area of 175×175 km.

Each experiment is performed starting from the same homogeneous isotropic model defined by a density of 3 g cm^{-3}, a *P*-wave velocity of 7 km s^{-1} and an *S*-wave velocity of 4 km s^{-1}. We tested inversions with and without Hessian pre-conditioning. For each target model we perform three different types of FWI. First, we apply a regular isotropic FWI, updating for density, *P*- and *S*-wave velocities using an *l*-BFGS optimization algorithm. Second, we apply a FWI updating for density and the 21 elastic coefficients expressed in the local Cartesian coordinate system. Finally, we apply another anisotropic FWI this time pre-conditioned with the inverse of the diagonal elements of the approximate Hessian. For anisotropic FWI, we will mainly focus on the hexagonal part of the elasticity tensors, which is extracted with the projection method. At each iteration, we consider the gradients and models from the four previous iterations in the *l*-BFGS algorithm for the estimation of the inverse Hessian. To keep the interpretations simple and the comparisons relevant, note that we do not include any regularization term in the cost function. For the same reason, we do not apply a hierarchical inversion from low to high frequencies (Bunks *et al.* 1995; Sirgue & Pratt 2004) and invert bandpass filtered data in a 5–20 s period band in all the synthetic tests. Convergence criteria are chosen so that inversions stop when the misfit attains one per mil of the initial cost function, when the line-search algorithm fails to find a new iterate, or simply when the computation wall-time is reached, that is after roughly 70 iterations for isotropic FWI and 60 iterations for anisotropic FWI.

4.2 Isotropic inclusion model

Because our formulation of ainversions are also compared to pure isotropic FWI tonisotropic FWI involves the complete set of the 21 elasticity coefficients, the dimensionality of the model space is largely increased compared to the classical isotropic case. This naturally raises concerns on our ability to constrain such a large number of parameters from teleseismic records and on potential trade-offs between the different parameters. To investigate the global behavior of anisotropic FWI, we carry out a first simple synthetic experiment which consists in recovering a purely isotropic target model. This model is designed as a 100-km-wide isotropic cubic-shaped inclusion embedded in an isotropic homogeneous half-space. We impose a density of 3 g cm^{-3}, and *P*- and *S*-wave velocities of 7 and 4 km s^{-1}, respectively in the background model and perturbations of +10 per cent on these parameters within the cubic inclusion (Figs 1a–c).

We first perform an isotropic FWI updating for density, *P*- and *S*-wave velocities. The isotropic model obtained from this isotropic inversion is shown in Figs 1(d)–(f). As expected for a typical teleseismic acquisition, dominated by the transmission regime, both short and long wavelengths of *P*- and *S*-wave velocity models are nicely reconstructed. On the other hand, only the top and bottom sides of the inclusion, thus the short wavelengths, are recovered in the density model (Fig. 1d). This results from the choice of the (ρ, v_P, v_S) parametrization for which the reflections control the density reconstruction (Tarantola 1986; Beller *et al.* 2018a). The overall convergence of FWI is fast, with a misfit reduced to 1‰ after 32 iterations (Fig. 2a, blue curve).

We then perform two anisotropic FWI using *l*-BFGS and pre-conditioned *l*-BFGS. To quantify the influence of seismic anisotropy, we measure the splitting intensity on the synthetic SKS waveforms by projecting the transverse component onto the time

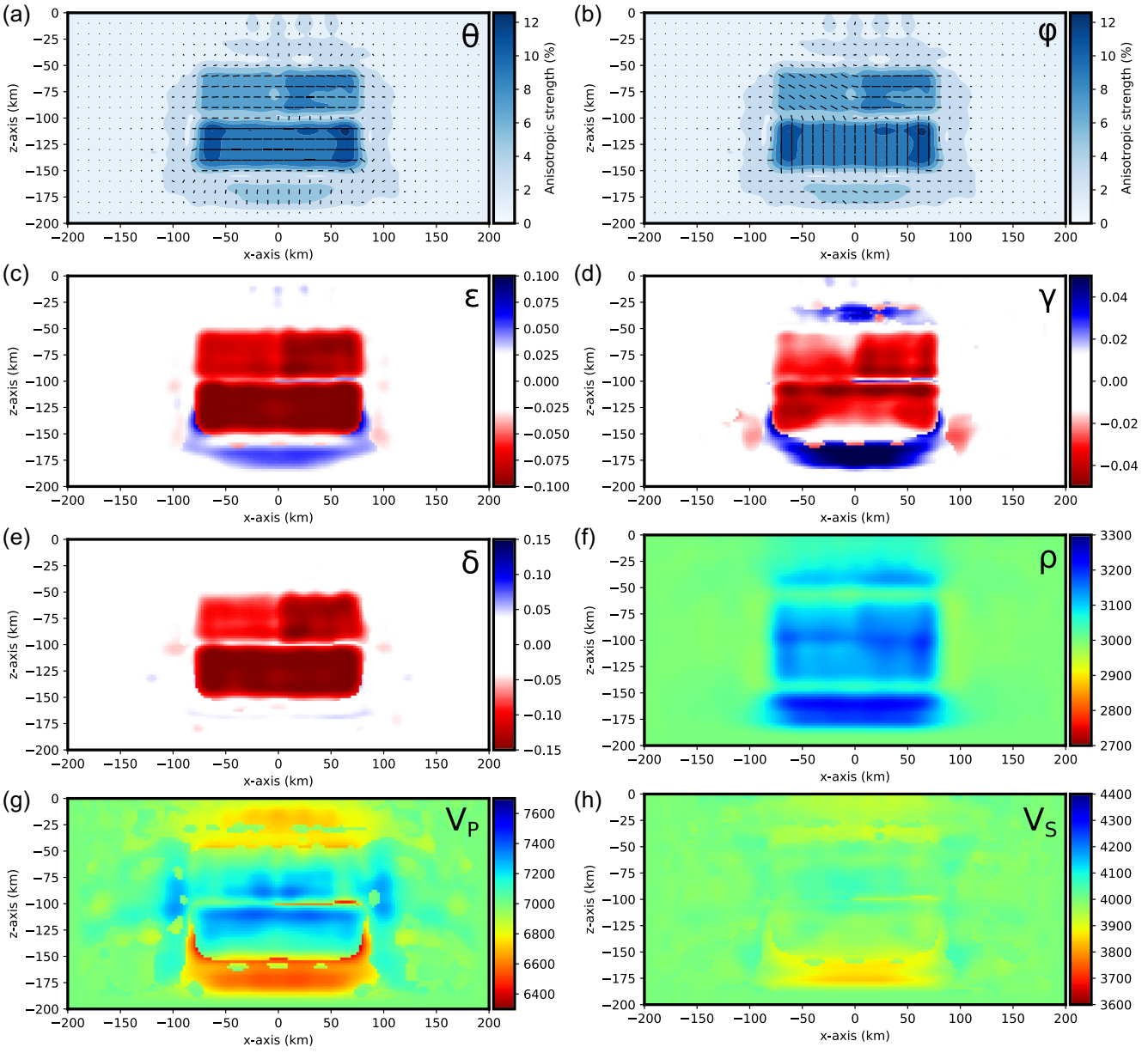

Figure 15. Vertical cross-sections in the three-block model obtained after 61 iterations using *l*-BFGS with the approximate hessian pre-conditioner. Recovered dip and azimuth angles (a and b) and reconstructed values of Thomsen's parameters ϵ, γ and δ (c–e). The three last panels (f–h) show density and reference hexagonal *P*- and *S*-wave velocities.

derivative of the radial component (Chevrot 2000; Sieminski *et al.* 2008). The splitting parameters (delay time and fast-axis orientation) are determined from the amplitude and phase of the sinusoid that best fits the azimuthal variations of splitting intensity:

$$S = \frac{1}{2}\Delta t \sin 2(\phi - \phi_0), \tag{29}$$

where Δt is the time delay between the fast and slow quasi-shear waves, ϕ_0 the azimuth of the fast axis and ϕ the source backazimuth or polarization of the incoming SKS wave. The comparison of SKS splitting predicted by the target, isotropic and anisotropic FWI models (Fig. 3) shows that neither the isotropic nor the anisotropic FWI models produce significant SKS splitting, as expected. However, we still note that weak SKS splitting are still produced by isotropic models. This apparent splitting is caused by wave scattering on the vertical edges of the inclusion which transfers energy on the transverse component.

As shown in Fig. 2(a) (orange and green curves), convergence rates are slower than those corresponding to pure isotropic FWI. The misfit reaches 5.4‰ and 2.2‰ after 56 and 54 iterations, respectively. This suggests that increasing the number of degrees of freedom does not improve the final fit. Instead, the larger dimensionality of the model space to explore prevents the iterative algorithm to locate the global minimum of the misfit function which is almost flat in many different directions. In short, it is easier to retrieve an isotropic model from isotropic FWI than with anisotropic FWI. Nevertheless, the recovered anisotropic model (Fig. 4) shows a relatively good recovery of both *P*- and *S*-wave velocities. In contrast, the density model (Fig. 4f) is poorly retrieved. This is expected from the radiation patterns of a density perturbation, which radiates

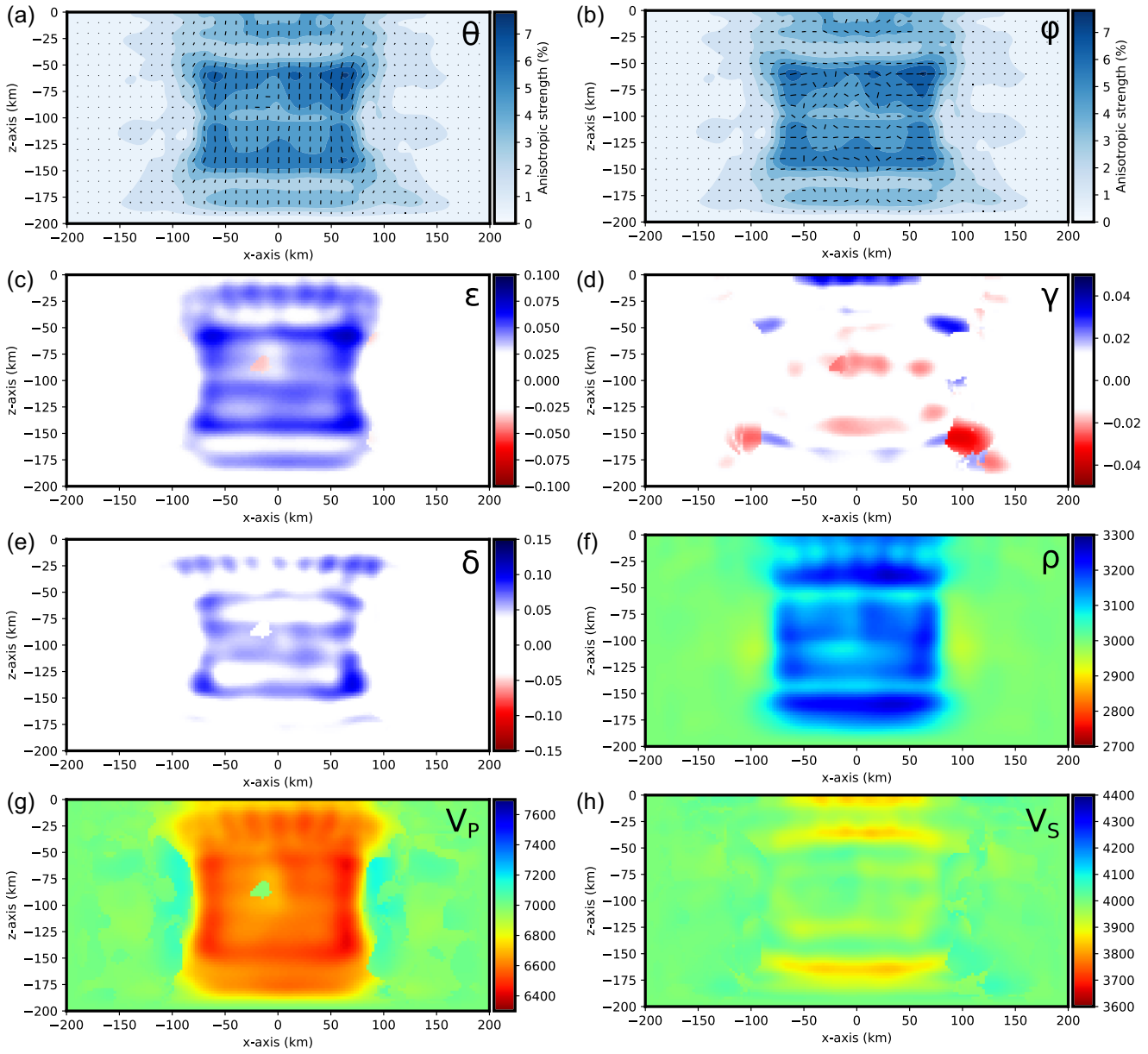

Figure 16. Vertical cross-sections in the three-block model obtained after 61 iterations using *l*-BFGS without the approximate hessian pre-conditioner. Recovered dip and azimuth angles (a and b) and reconstructed values of Thomsen's parameters ϵ, γ and δ (c–e). The three last panels (f–h) show density and reference hexagonal *P*- and *S*-wave velocities.

energy equally in the forward and backward directions (Tarantola 1986; Beller *et al.* 2018a). However, the scattering coefficient is negative, because an increase of wave velocity can be achieved equally well by increasing elastic parameters or decreasing density. The averaged deviations from isotropy (Fig. 4) are weak, except along the lateral boundaries of the inclusion where the anisotropic strength reaches up to 6 per cent of the norm of the elasticity tensor (Figs 4a and b). This suggests potential trade-offs between ρ, ϵ and δ at intermediate angles, as already pointed out by Alkhalifah & Plessix (2014). Finally, Fedorov averaged *P*- and *S*-wave velocity models (Figs 1g–i) present a similar resolution to those obtained with pure isotropic FWI but display slower *P*- and *S*-wave velocities.

4.3 Anisotropic inclusion models

In the next synthetic experiments, we seek to assess the ability of anisotropic FWI to recover the anisotropic parameters of hexagonal symmetry (ρ, V_p, V_s, ϵ, δ, γ) as well as the azimuth ϕ and dip θ of the symmetry axis, so a total of 8 independent parameters. These anisotropic inversions are also compared to pure isotropic FWI to

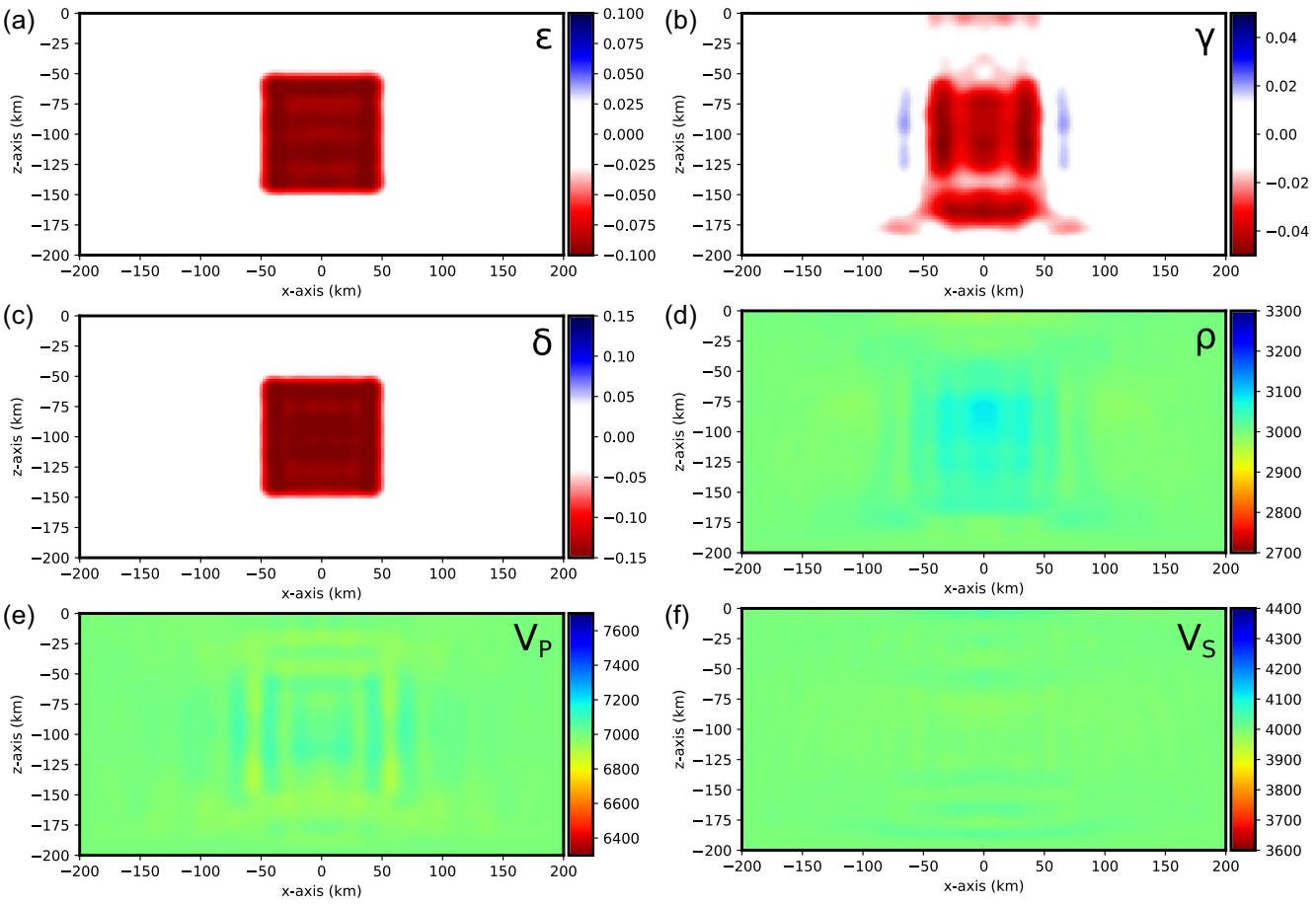

Figure 17. FWI model for the VTI inclusion using a reduced VTI parametrization.

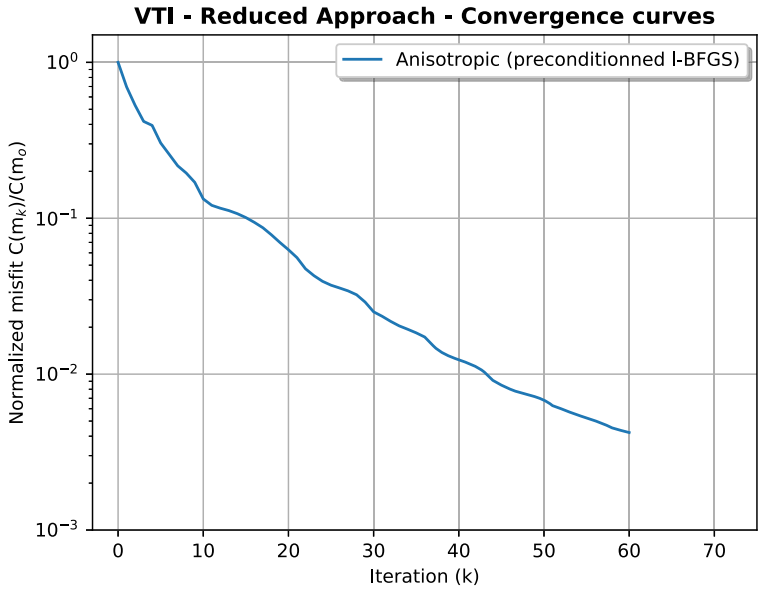

Figure 18. Convergence of anisotropic FWI using a reduced VTI parametrization.

investigate the effects and possible biases of seismic anisotropy in isotropic tomography. For this purpose, we consider three new synthetic target models that all contain a 50 km cubic-shaped hexagonal inclusion located in the middle of an homogeneous isotropic half-space at 100 km depth. Inside these anisotropic domains, the hexagonal parameters are $\epsilon = -0.1$, $\delta = -0.15$ and $\gamma = -0.05$, which represent the anisotropy of a typical olivine and pyroxene aggregate in the upper mantle (Becker *et al.* 2006, their eqs 8 and

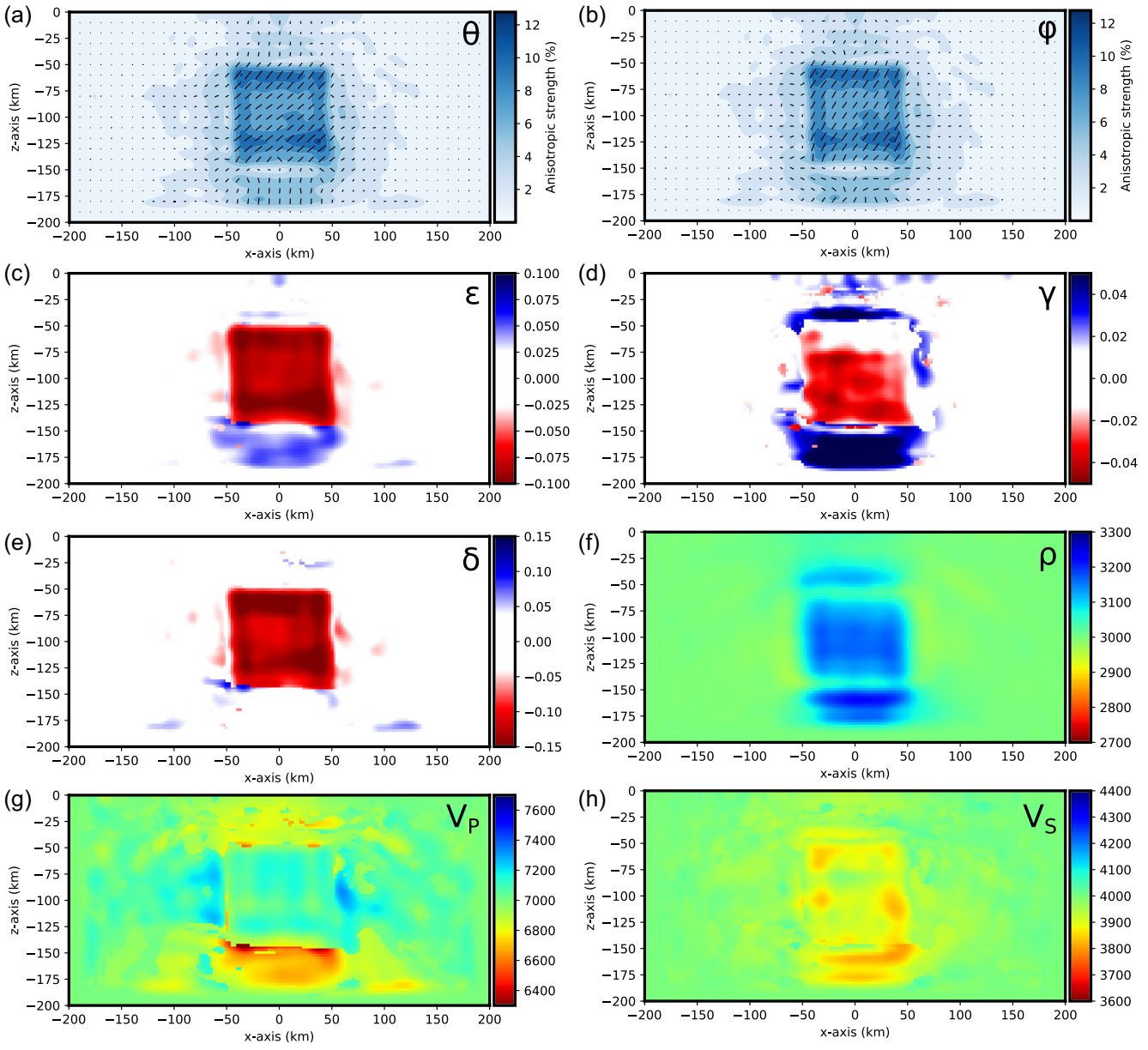

Figure 19. FWI model for the TTI inclusion experiment using incident *P* waves only.

9). The three models are presented in Fig. 5. They only differ by the orientation of the hexagonal fast symmetry axis. In the first experiment, we consider a VTI inclusion with a vertical fast symmetry axis ($\theta = 0^\circ$ and $\phi = 0^\circ$). In the second experiment, we consider an HTI inclusion with an horizontal fast symmetry axis ($\theta = 90$, $\phi = 35^\circ$). And finally, we consider a TTI inclusion with a tilted symmetry axis ($\theta = 45^\circ$ and $\phi = 35^\circ$).

The convergence curves and data residuals of isotropic and anisotropic FWI for the hexagonal anisotropy experiments are shown in Fig. 6. Regardless the orientation of the fast symmetry axes (VTI, HTI and TTI), the behaviour of FWI is always the same. In isotropic inversions (blue curves), FWI faces difficulties to achieve a good misfit reduction. The misfit reaches 35.9, 25.1 and 45.1 per cent of the initial misfit after 70, 75 and 75 iterations for the VTI, HTI and TTI cases respectively. In contrast, anisotropic FWI—with (green curves) or without (orange curves) pre-conditioning—is characterized by better convergence rates and better misfit reductions that are one to two orders of magnitude below those of isotropic inversions. Introducing the Hessian diagonal pre-conditioner in anisotropic FWI systematically improves the convergence. In the VTI experiment (Fig. 6a), the misfit reaches 4.53 per cent and 6.73‰ of its initial value after 53 and 56 iterations for anisotropic FWI and pre-conditioned anisotropic FWI, respectively. In the HTI experiment (Fig. 6c), the misfit reaches 7.24 per cent and 2.17‰ of its initial value after 60 and 59 iterations for anisotropic FWI and pre-conditioned anisotropic FWI, respectively. In the TTI experiment (Fig. 6e), the misfit reaches 1.36 per cent and 2.89‰ of its initial value after 61 and 59 iterations for anisotropic FWI and pre-conditioned anisotropic FWI, respectively.

The data misfits in Figs 6(b), (d) and (f) reveal the overall effects of anisotropy on seismic waveforms and the degree to which FWI succeeds in explaining these effects. In the VTI experiment (Fig. 6b), initial data residuals are almost constant over the full

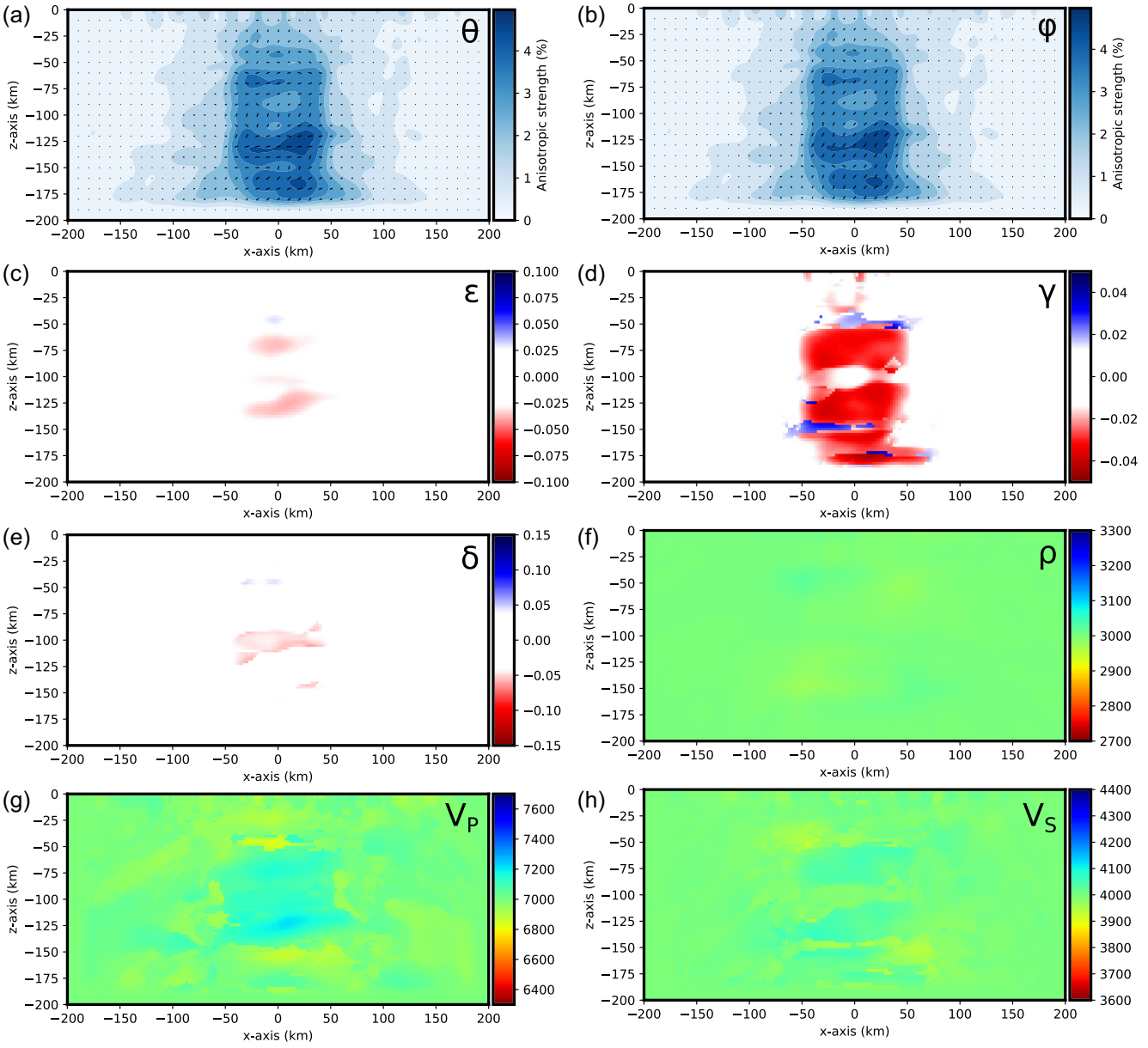

Figure 20. FWI model for the TTI inclusion experiment using incident SKS waves only.

azimuthal range. They are predominant on the vertical and radial components and increase with incidence angles. Note that they are more important for incident *P* waves. On the other hand, in the HTI experiment (Fig. 6d), the incident *S* wave initial misfit displays strong azimuthal variations, in contrast to *P* wave misfits. Specifically, we observe on the transverse component a $\pi/2$-periodic pattern while the vertical and radial components show a π-periodic pattern. Moreover, for the TTI experiment (Fig. 6f), the *P*-wave vertical and radial components also present a 2π-periodic azimuthal variation. Note that this 2π periodicity also affect to some degree the *S* waves. The data residuals computed in the final isotropic and anisotropic models present strong differences. The most striking feature is that isotropic inversion does not succeed in explaining anisotropic effects since the final residuals still exhibit the same azimuthal variations as in the initial model. This is particularly true for the transverse component of SKS waves in the HTI experiment (Fig. 6d) where the isotropic inversion failed to reduce the residuals. On the other hand, anisotropic FWI recovers a model that is able to reproduce the effects of anisotropy since azimuthal variations are no longer observed on the final data residuals. Furthermore, these residuals are 10–100 times weaker than those obtained with isotropic FWI.

Fig. 7 presents predicted SKS splitting for the hexagonal inclusion models. The VTI inclusion model (panel a) does not produce significant shear wave splitting. The HTI and VTI target models (panels d and g) produce a significant and very similar shear wave splitting. Note that models obtained from isotropic inversions (panels b, e and h) never reproduce observed shear wave splitting patterns in contrast to models obtained from anisotropic inversions (panels c, f and i). This strongly suggests that purely isotropic models cannot reproduce shear wave splitting effects.

Results of pre-conditioned anisotropic FWI are presented in Figs 8 for VTI, 9 for HTI and 10 for TTI cases. For these three

experiments, the application of pre-conditioned anisotropic FWI succeeds in recovering both the shape of the anomaly and the amplitudes of Thomsen's parameters (panels c, d and e). The inversion also recovers the dip and azimuth angles within the anisotropic inclusions (panels a and b). Although the target model only contains perturbations of anisotropic parameters, pronounced perturbations of density and reference velocities (panels f, g and h) are visible depending on the orientation of the fast symmetry axis. Indeed, while there is almost no perturbation of these parameters in the VTI case (Figs 8f–h), the TTI (Figs 10 f–h) and especially the HTI cases (Figs 9f–h) present strong artefacts, especially for density.

Results of purely isotropic inversions are presented in Fig. 11 and are compared with Fedorov isotropic averages derived from anisotropic FWI models. Isotropic models resulting from isotropic FWI present strong artefacts mostly affecting the density model (Figs 11a and d), and the amplitude and shape of *P*- and *S*-wave velocity anomalies (Fig. 11h). These imprints of seismic anisotropy on the tomographic models are clearly varying with the orientation of the symmetry axis. Interestingly, Fedorov isotropic averages recover isotropic models that are less biased than those obtained from isotropic FWI (Figs 11m–u). These models are very close to the Fedorov isotropic averages of the target models (Figs 11j–i).

4.4 Three-block model

To determine the potential of anisotropic FWI for imaging the depth and lateral variations of seismic anisotropy, we performed a last synthetic experiment. The target anisotropic model we aim to recover is composed of three distinct anisotropic blocks containing homogeneous HTI perturbations embedded in an homogeneous isotropic half-space (Fig. 12). To mimic a layered anisotropic structure, we first introduce a 50 km high and 150 km wide block located at 125 km depth representing the asthenospheric anisotropy. Inside this layer, the intrinsic anisotropic parameters ϵ, δ and γ are -0.1, -0.15 and -0.05, respectively, and the fast axes are oriented along the north–south direction ($\theta = 0^\circ$ and $\phi = 0^\circ$). On top of this layer, we put two distinct lithospheric anisotropic blocks. These blocks are 50 km high and 75 km wide and have their center located at 75 km depth. In the western block $\epsilon = -0.08$, $\delta = -0.1$ and $\gamma = -0.03$ and the fast axes are oriented at 120° from North. In the eastern block, the anisotropy is the same as in the asthenospheric layer but the fast symmetry axes are orthogonal ($\theta = 0^\circ$ and $\phi = 90^\circ$). This challenging test is designed to investigate the ability of FWI to constrain seismic anisotropy in a case where apparent SKS splitting is very weak (Fig. 13a).

Fig. 14(a) shows the evolution of the misfit function over several iterations for the isotropic (blue), anisotropic (orange) and pre-conditioned anisotropic (green) FWI. First, we note that isotropic FWI quickly reaches a plateau after 10 iterations, with a final misfit function of about 5 per cent of its initial value after 72 iterations. In contrast, anisotropic FWI shows a slower convergence rate but reaches a much smaller misfit value of 8.1‰ after 59 iterations. Pre-conditioned anisotropic FWI gives the best misfit function reduction going down to 1.3‰ of the initial misfit function after 61 iterations. Data residuals (Fig. 14b) present azimuthal variations similar to those obtained for the HTI inclusion, with a 2π-periodicity that is coherent with predicted SKS splitting in the western block.

Anisotropic FWI models obtained with the diagonal inverse Hessian pre-conditioner are shown in Fig. 15. For comparison we also show in Fig. 16 the FWI anisotropic model without pre-conditioning. When using the pre-conditioner (Fig. 15), anisotropic FWI perfectly retrieves the depth and lateral variations of the fast axis directions (panels a and b). Intrinsic anisotropic parameters (panels c, d and e) are also well recovered except for γ which exhibits significant artefacts. The shape of the anomalies remains close to the target model, but the hexagonal reference velocities and density models present strong artefacts (see panels f, g and h).

Anisotropic FWI models obtained without pre-conditioning (Fig. 16) are dramatically different. In this case, the inversion completely fails to recover the fast symmetry axis orientation (panels a and b) and the depth and lateral variation of anisotropy. The recovered medium is mainly VTI (panel a) making the azimuthal orientation of the fast axis uninterpretable (panel b). The γ parameter is not constrained at all and presents strong artefacts. The ϵ and δ parameters are also poorly resolved and more importantly have opposite sign to the target model. The reference hexagonal *P*-wave velocities are highly perturbed and present strong negative anomalies (panel g) while *S*-wave velocities are weakly perturbed but present visible artefacts. Again, the density is strongly affected.

It is clear from these experiments that, despite a fairly good convergence rate, anisotropic FWI without pre-conditioner gives poor results. The diagonal Hessian pre-conditioner is thus a key ingredient of anisotropic FWI.

5 DISCUSSION

5.1 The importance of the pre-conditioner

Anisotropic FWI without pre-conditioning failed to recover the three-block target model despite a good reduction of the misfit function, which emphasizes the importance of the diagonal Hessian pre-conditioner. Because we invert the whole 21 elastic coefficients, the model space is drastically expanded compared to the isotropic case, which exacerbates the non-uniqueness of the inverse problem. A good reduction of the objective functional can thus be achieved while converging towards a biased model of the subsurface.

The model obtained from non-pre-conditioned FWI (Fig. 16) differs from the expected target model but the results of the inversion can be understood with simple physical arguments. Indeed, because the target model is composed of HTI perturbations, characterized by fast horizontal symmetry axes, *P* waves propagating at near vertical incidence mostly sample slow propagation directions. Since teleseismic *P* waves are mainly sensitive to the C_{33} elastic coefficient, anisotropic FWI will principally update C_{33}, keeping C_{11} and C_{22} close to their original values in the starting model. Therefore, C_{33} will be smaller than C_{11} and C_{22} in the final model, leading to an equivalent VTI medium, characterized by a slow vertical symmetry axis. Since the reference hexagonal velocities are defined for propagation along the symmetry axis, in that case the vertical direction, ϵ, δ and γ become positive. This explains the strong imprint of a teleseismic acquisition on the results of FWI, with strong negative perturbations of *P*-wave velocities (Fig. 16g), and the recovery of a slow vertical symmetry axis (Fig. 16a) with positive ϵ and δ parameters (Fig. 16c). The γ parameter is mainly constrained by the birefringence of S and SKS waves. However, in the three-block model the apparent splitting is weak, and thus seismic anisotropy is mainly expressed on *P*-wave records (Fig. 14b), which explains the poor reconstruction of γ.

In contrast, when the diagonal Hessian pre-conditioner is used (Fig. 15) the inversion almost perfectly recovers the target model. We recall from Section 2.3 that the Hessian accounts for the limited bandwidth effects of the source–receiver acquisition and reduces

artefacts resulting from the partial (mostly vertical) illumination offered by teleseismic waves. This means that for the first iterations of FWI, application of the Hessian reduces the sensitivity to C_{33} and increases the sensitivity to C_{11} and C_{22}. The search direction is no longer constrained in the vicinity of an equivalent VTI medium. The quality of the pre-conditioner is critical for anisotropy imaging as it removes this fast-slow axis ambiguity resulting from the teleseismic acquisition geometry. Another benefit of this pre-conditioner is that it reduces the high sensitivity close to receiver locations and helps to better focus perturbations at depth. The action of the inverse Hessian depends on the chosen water-level parameter ζ, which we took as the value of the Hessian at the center of the model. Increasing ζ will enhance perturbations close to the surface, while decreasing ζ will enhance perturbations at depth. From our experience, FWI is not critically dependent on the water-level because the inverse Hessian obtained by *l*-BFGS quickly improves after only a few iterations. Hence, the pre-conditioner is important only during the first iterations, and especially during the initial pre-conditioned steepest descent step.

5.2 Resolution potential of anisotropic FWI for constraining the layering of seismic anisotropy and the dip of the symmetry axis

To some extent, anisotropic FWI can be compared to vectorial tomography (Chevrot 2006; Monteiller & Chevrot 2011) and adjoint tomography (Sieminski *et al.* 2008) based on splitting intensity measurements (Chevrot 2000). This observable is homogeneous to the finite-frequency traveltime measurements (Marquering *et al.* 1999) commonly used in adjoint tomography (Tromp *et al.* 2005). The resolution improvement of anisotropic FWI compared to splitting intensity tomography is thus analogous to the resolution increase between full-waveform inversion and adjoint phase tomography (Monteiller *et al.* 2015).

Apparent SKS splitting patterns in the HTI and TTI models are very similar (Fig. 7). This stems from the poor sensitivity of SKS splitting to the dip of the fast symmetry axis (Chevrot & Van Der Hilst 2003). In other words, from SKS splitting alone it is not possible to discriminate tilted anisotropy from pure horizontal anisotropy. However, a striking feature of anisotropic FWI is its ability to constrain both the azimuth and dip angles of anisotropy (Figs 8, 9, 10 and 15). This remarkable achievement comes from the exploitation of both *P* and *S* teleseismic waves in our synthetic tests.

To get further insights into the behaviour of anisotropic FWI, let us examine in detail the initial waveform residuals. A 2π periodicity of *P* wave vertical and radial component residuals is clearly observed in the TTI case (Fig. 6c). SKS wave residuals are characterized by a π-periodicity on the vertical and radial components, and a quasi π-periodicity on the transverse component. On the other hand, in the HTI case (Fig. 6b), no periodicity is visible on the *P* wave residuals while the three components of SKS waves present a 2π-periodicity. These azimuthal variations suggest that the sensitivity to dip is mainly carried by *P* waves.

We performed a new TTI experiment by considering only *P* waves (Fig. 19) or SKS-waves (Fig. 20). When only SKS waves are considered, anisotropic FWI fails to retrieve the anisotropic parameters: only γ is recovered with a relatively poor spatial resolution. A closer look at recovered dip and azimuth angles shows that despite the sensitivity of SKS waves to the azimuth of the symmetry axis, we recover a mean azimuth of 30°, and fail to recover the mean dip of the symmetry axis. In contrast, considering only incoming *P* waves allows us to recover both the dip and azimuth angles of the fast axis. In this particular case, ϵ and δ are fairly well recovered while γ is only partially reconstructed with strong artefacts. This last synthetic experiment therefore gives further support to the key role played by *P* waves in the inversion, which should be kept in mind for future applications on real data.

5.3 Can we resolve the whole elasticity tensor with teleseismic acquisitions?

The choice of keeping the 21 coefficients of the elasticity tensor is appealing as it greatly simplifies the formulation of the inverse problem for arbitrary anisotropy. However, this raises the question of our ability to resolve them all as required in the general triclinic case (Köhn *et al.* 2015), especially with a coarse teleseismic source-receiver acquisition (Beller *et al.* 2018a).

To check if our approach would recover a full triclinic tensor, we apply pre-conditioned anisotropic FWI to the triclinic tensor presented in Section 3.3. Results of anisotropic FWI (not shown) indicate that FWI fails to retrieve a triclinic medium using only teleseismic waves. Perhaps further refinements of the technique could improve the results of FWI for triclinic media. For example, adding Sh or surface waves into the inversion may provide additional independent information to better constrain seismic anisotropy, but these extensions are left for future studies.

Another possible issue is related to the posterior decomposition of the elasticity tensor. A potential pitfall is to mix parameters resolved with a different spatial resolution, hence estimating new parameters that are potentially biased. It seems that in our synthetic tests, this problem, if present, has a rather limited impact on the estimation of the hexagonal part of the elasticity tensors. To limit these effects, simple solutions would be to smooth the gradients, add some regularization terms to the cost function, or smooth the model parameters just before performing the decomposition. Again, this was not necessary for the synthetic tests that we have considered in this study.

To investigate the potential detrimental effects of projecting a full anisotropic tensor on its hexagonal subspace, we compare the results of the VTI inclusion experiment (Fig. 8) to the same experiment performed using a reduced VTI parameterization (Fig. 17). The reduced approach only involves reference *P*- and *S*-waves velocities, density, and ϵ, δ and γ, since the symmetry axis is kept fixed along the vertical direction. The convergence rate of the reduced approach is similar to the one of the decomposition approach (Fig. 18). The two approaches also give similar results: ϵ and δ are quite well recovered while γ presents artefacts, especially with the reduced approach. The *P*-wave velocity model obtained with the decomposition method presents obvious trade-offs with ϵ and γ, whereas this is not the case with the reduced method. In any case, this test demonstrates the capacity of the decomposition method compared to the reduced method to retrieve anisotropy models with hexagonal symmetry.

Since the fabrics of olivine and pyroxene aggregates in the upper-mantle have a predominant hexagonal component (Becker *et al.* 2006), we thus believe that our FWI method should be sufficient to obtain robust and finely resolved images of upper-mantle seismic anisotropy, at least in regions that are covered by dense seismic arrays.

5.4 On the trade-off between isotropy and anisotropy

As already stressed by Bezada *et al.* (2016), anisotropy-induced artefacts on regional traveltime tomography are common and important enough to significantly alter the interpretation of tomographic images. Our study provides further insights into (1) the potential biases induced by applying isotropic FWI on data contaminated by anisotropy and (2) the potential trade-offs between isotropic and anisotropic parameters emerging from anisotropic inversion of isotropic structures.

The inversion results for the anisotropic inclusion indeed suggest that the contribution of hexagonal anisotropy can potentially have a strong signature in isotropic tomographic models. Recovered images are highly distorted (e.g. Fig. 11h), and show strong amplitude variations and artefacts on all parameters (Figs 11a–i). Since the target models only differ with respect to the orientation of the symmetry axis, it is clear that anisotropy is responsible for these biases. This also demonstrates the inability of isotropic FWI in imaging anisotropic media and the potential pitfalls of interpreting the results of isotropic tomography, in particular in terms of thermal anomalies in the upper-mantle. However, the isotropic part of the anisotropic models obtained by FWI properly recovers the Fedorov isotropic average of the target models (Figs 11j–u). This suggests that it is better to apply anisotropic FWI with projection onto the isotropic part than pure isotropic FWI in the presence of seismic anisotropy, which we expect to be relatively ubiquitously distributed in the lithospheric and asthenospheric mantle. A similar idea has been proposed in the context of homogenization theory by Capdeville & Métivier (2018). Indeed, homogenization aims at recovering the anisotropic up-scaled equivalent model of the subsurface able to reproduce the data that would have been modelled from a fine-scale isotropic model. Capdeville & Métivier (2018) showed that inverting for the 21 elastic coefficients allows to recover the homogenized isotropic version of the subsurface model by extracting the isotropic part of the resulting complete elasticity tensors.

On the other hand, when considering a purely isotropic target model, isotropic FWI gives better results than anisotropic FWI. Obviously, in that case, allowing anisotropic parameters to vary unnecessarily increases the dimensionality of the parameter space, and as can be seen in Fig. 4, anisotropic FWI does not perfectly recover the isotropic model. But the overall deviation from isotropy remains weak (less than 2 per cent).

A drawback of anisotropic FWI is that it produces strong artefacts in the density models which result from strong apparent trade-offs with velocity models (Fig. 4g) as already discussed by Pan *et al.* (2018). Since we update simultaneously density and elastic moduli, an increase of elastic wave velocities can be obtained by either an increase of elastic moduli or a decrease of density. Further inspection of Fig. 4 indicates another possible trade-off between density and Thomsen's parameters as suggested by Alkhalifah & Plessix (2014). Indeed, the scattering patterns of a density perturbation (Tarantola 1986; Beller *et al.* 2018a) radiates energy both in the forward and backward directions. This explains the broad reconstruction of density within the inclusion and the short wavelength reconstruction of the inclusion at the top and bottom boundaries. Therefore, waves scattered at intermediate scattering angles on the vertical boundaries cannot be explained by density perturbations, and FWI tries to explain these scattered waves by perturbing *P*-wave velocities and anisotropic parameters. Perhaps a possible workaround to overcome these tradeoffs would be to consider density normalized elastic parameters (Sieminski *et al.* 2007a; Rusmanugroho *et al.* 2017).

Finally, it is interesting to note the similarity between the weak convergence of isotropic FWI obtained when trying to reconstruct anisotropic models (Fig. 6) and the weak convergence of previous applications of FWI on real data. Indeed, in Wang *et al.* (2016) and Beller *et al.* (2018b), misfit functions converged after a small number of iterations to reach a plateau, a typical behavior that we observed when we tried to image an anisotropic target model with isotropic FWI. Previously, this weak convergence on real data sets was attributed to incoherent noise. In the Earth's upper-mantle, since anisotropy is the rule (as testified by numerous shear wave splitting and surface waves studies), our results strongly advocate for moving from isotropic to anisotropic FWI to improve the results of seismic tomography.

5.5 Remaining barriers to apply FWI on real data

While our anisotropic FWI method gave very promising results on the different synthetic cases that we considered, we realize that substantial efforts are still needed before considering real data applications.

First, the expected resolution will depend on the distribution of seismic sources and the geometry of the acquisition. While several recent deployments of seismic sensors are comparable to the acquisition considered in our synthetic study, the uneven distribution of earthquakes will generally lead to uncompleted azimuthal illumination. Impact of uneven illumination has not been investigated in this study. Nevertheless, for dense seismic network that were operational over a long periodof time, we believe that this should not be an issue.

In practice, the application of teleseismic FWI requires an estimation and/or refinement of the apparent source wavelet corresponding to the incoming teleseismic wavefield prior to inversion. The wavelet estimation is a critical ingredient of teleseismic FWI since it permits to absorb many propagation effects that occur between the source and the regional domain. Examples of such effects are amplitude and traveltime anomalies that may result from global-scale seismic structures, seismic source complexities (origin time, directivity, moment magnitude) and wavelet broadening due to anisotropy and attenuation. Since the particular motion of teleseismic *P* waves is polarized along the quasi vertical propagation direction the array-averaged vertical component of the *P* wave provides a good proxy of the incoming wavelet. In contrast, the polarization of *S* waves is mainly controlled by the source focal mechanism, but it can be perturbed by noise on the horizontal components, complexity of the source, anisotropy in the source and receiver regions, and 3-D heterogeneities. The estimation of the incoming shear wavefield reaching the boundaries of the 3-D regional grid will thus require special care. In principle, using similar approaches as in *S* receiver function studies (e.g. Farra & Vinnik 2000) should allow us to estimate both the polarization and the source wavelet of an incoming teleseismic *S* wave. Updating the source wavelets after each iteration of FWI should also help us to better separate the effects of 3-D heterogeneities inside and outside the regional grid. While straightforward, this has not been implemented yet. Another route for improvement could also come from the introduction of source-free waveform misfit functions as proposed in several recent seismic imaging studies (Choi & Alkhalifah 2011; Warner & Guasch 2016), or considering waveform reconstruction based inversions (van Leeuwen & Herrmann 2013; Aghamiry *et al.* 2019).

6 CONCLUSIONS

We have presented a new method to tackle the problem of 3-DXXXXX imaging of upper-mantle anisotropy from full-waveform inversion of teleseismic P and S waves. Our method follows a two-step strategy. The 21 components of the elasticity tensors are first inverted by FWI, using a pre-conditioned quasi-Newton approach. The recovered elasticity tensors obtained at each node of the tomographic grid are then analyzed to determine the orientation of the symmetry axes. After rotating the elasticity tensors to their natural coordinate systems, they are then decomposed by projections to extract the contributions corresponding to the different symmetry classes. From the hexagonal symmetry part, we can determine both the azimuth and dip of the symmetry axis, and also the values of the generalized Thomsen's parameters ϵ, δ and γ. Using rather simple synthetic experiments, we have demonstrated the efficiency of our method to completely characterize seismic anisotropy at least for hexagonal media, with a good lateral and (more remarkably) vertical resolution. Using an approximate Hessian turned out to be a key ingredient to improving the convergence rate and properly recovering seismic anisotropy. Our study also emphasized the importance of considering both P and S waves in the inversion.

ACKNOWLEDGEMENTS

We are very grateful to the editor Ludovic Métivier and the two anonymous reviewers for their thoughtful reviews. We gratefully acknowledge support from the Centre National de la Recherche Scientifique (CNRS), the Bureau de Recherches Géologiques et Minières (BRGM) and TOTAL through the OROGEN project. This study was granted access to the HPC resources of CALMIP supercomputing center under the allocation 2018-P1138, and the HPC resources of CINES/IDRIS under the allocation A0040410394 made by GENCI.

REFERENCES

Aghamiry, H.S., Gholami, A. & Operto, S., 2019. Improving full-waveform inversion by wavefield reconstruction with the alternating direction method of multipliers, *Geophysics,* **84**(1), R139–R162.

Aki, K. & Richards, P.G., 2002, *Quantitative Seismology,* University Science Books.

Alder, C., Bodin, T., Ricard, Y., Capdeville, Y., Debayle, E. & Montagner, J., 2017. Quantifying seismic anisotropy induced by small-scale chemical heterogeneities, *Geophys. J. Int.,* **211**(3), 1585–1600.

Alkhalifah, T. & Plessix, R.-É., 2014. A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study, *Geophysics,* **79**(3), R91–R101.

Babuska, V. & Cara, M., 1991. *Seismic Anisotropy in the Earth,* Vol. **10,** Springer Science & Business Media.

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

Beghein, C., Trampert, J. & Van Heijst, H., 2006. Radial anisotropy in seismic reference models of the mantle, *J. geophys. Res.,* **111**(B2), doi:10.1029/2005JB003728.

Beller, S., Monteiller, V., Combe, L., Operto, S. & Nolet, G., 2018a. On the sensitivity of teleseismic full-waveform inversion to earth parametrization, initial model and acquisition design, *Geophys. J. Int.,* **212**(2), 1344–1368.

Beller, S., Monteiller, V., Operto, S., Nolet, G., Paul, A. & Zhao, L., 2018b. Lithospheric architecture of the South-Western Alps revealed by multi-parameter teleseismic full-waveform inversion, *Geophys. J. Int.,* **212**(2), 1369–1388.

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

Bostock, M., Rondenay, S. & Shragge, J., 2001. Multiparameter two-dimensional inversion of scattered teleseismic body waves 1. Theory for oblique incidence, *J. geophys. Res.,* **106**(B12), 30 771–30 782.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N. & Pugmire, D., 2016. Global adjoint tomography: first-generation model, *Geophys. J. Int.,* **207**(3), 1739–1766.

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

Bunks, C., Saleck, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics,* **60**(5), 1457–1473.

Capdeville, Y. & Métivier, L., 2018. Elastic full waveform inversion based on the homogenization method: theoretical framework and 2-d numerical illustrations, *Geophys. J. Int.,* **213**(2), 1093–1112.

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

Chen, P., Jordan, T.H. & Zhao, L., 2007. Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods, *Geophys. J. Int.,* **170**(1), 175–181.

Chevrot, S., 2000. Multichannel analysis of shear wave splitting, *J. geophys. Res.,* **105**(B9), 21 579–21590.

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.

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

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

Choi, Y. & Alkhalifah, T., 2011. Source-independent time-domain waveform inversion using convolved wavefields: Application to the encoded multisource waveform inversion, *Geophysics,* **76**(5), R125–R134.

Clouzet, P., Masson, Y. & Romanowicz, B., 2018. Box tomography: first application to the imaging of upper-mantle shear velocity and radial anisotropy structure beneath the North American continent, *Geophys. J. Int.,* **213**(3), 1849–1875.

Cowin, S.C. & Mehrabadi, M.M., 1987. On the identification of material symmetry for anisotropic elastic materials, *Quart. J. Mech. Appl. Math.,* **40**(4), 451–476.

Debayle, E., Kennett, B. & Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of australia, *Nature,* **433**(7025), 509.

Endrun, B., Lebedev, S., Meier, T., Tirel, C. & Friederich, W., 2011. Complex layered deformation within the Aegean crust and mantle revealed by seismic anisotropy, *Nat. Geosci.,* **4**(3), 203.

Farra, V. & Vinnik, L., 2000. Upper mantle stratification by P and S receiver functions, *Geophys. J. Int.,* **141**(3), 699–712.

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.

Fedorov, F.I., 1968. *Theory of Elastic Waves in Crystals,* Springer.

Fichtner, A., Bunge, H.-P. & Igel, H., 2006. The adjoint method in seismology: I. Theory, *Phys. Earth planet. Inter.,* **157**(1–2), 86–104.

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

Fouch, M.J. & Rondenay, S., 2006. Seismic anisotropy beneath stable continental interiors, *Phys. Earth planet. Inter.,* **158**(2), 292–320.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. & Virieux, J., 2013. Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: sensitivity and trade-off analysis, *Geophysics,* **78**(2), R81–R105.

Helbig, K., 1995. Representation and approximation of elastic tensors, in *Society of Exploration Geophysicists,* eds , Fjær , E., Holt, R.M & Rathore, J.S., pp. 37–75, GeoScienceWorld.

Igel, H., Mora, P. & Riollet, B., 1995. Anisotropic wave propagation through finite-difference grids, *Geophysics,* **60**(4), 1203–1216.

Jin, S., Madariaga, R., Virieux, J. & Lambaré, G., 1992. Two-dimensional asymptotic iterative elastic inversion, *Geophys. J. Int.,* **108**(2), 575–588.

Kawai, K., Takeuchi, N. & Geller, R.J., 2006. Complete synthetic seismograms up to 2 Hz for transversely isotropic spherically symmetric media, *Geophys. J. Int.,* **164**(2), 411–424.

Köhn, D., Hellwig, O., De Nil, D. & Rabbel, W., 2015. Waveform inversion in triclinic anisotropic media – a resolution study, *Geophys. J. Int.,* **201**(3), 1642–1656.

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

Komatitsch, D. & Tromp, J., 2002. Spectral-element simulations of global seismic wave propagation. validation, *Geophys. J. Int.,* **149**(2), 390–412.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations, in *Conference on iNverse Scattering: Theory and Application,* pp. 206–220, SIAM.

Lin, Y.-P., Zhao, L. & Hung, S.-H., 2014. Full-wave multiscale anisotropy tomography in Southern California, *Geophys. Res. Lett.,* **41**(24), 8809–8817.

Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. seism. Soc. Am.,* **96**(6), 2383–2397.

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

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

Long, M.D. & Silver, P.G., 2008. The subduction zone flow field from seismic anisotropy: a global view, *Science,* **319**(5861), 315–318.

Long, M.D. & Wirth, E.A., 2013. Mantle flow in subduction systems: the mantle wedge flow field and implications for wedge processes, *J. geophys. Res.,* **118**(2), 583–606.

Marone, F. & Romanowicz, B., 2007. The depth distribution of azimuthal anisotropy in the continental upper mantle, *Nature,* **447**(7141), 198.

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

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

Masson, Y., Cupillard, P., Capdeville, Y. & Romanowicz, B., 2013. On the numerical implementation of time-reversal mirrors for tomographic imaging, *Geophys. J. Int.,* **196**(3), 1580–1599.

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.

Métivier, L. & Brossier, R., 2016. The SEISCOPE optimization toolbox: a large-scale nonlinear optimization library based on reverse communication, *Geophysics,* **81** (2), doi:10.1190/geo2015–0031.1.

Moakher, M. & Norris, A.N., 2006. The closest elastic tensor of arbitrary symmetry to an elasticity tensor of lower symmetry, *J. Elast.,* **85**(3), 215–263.

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.

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

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

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

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

Nocedal, J., 1980. Updating quasi-newton matrices with limited storage, *Math. Comput.,* **35**(151), 773–782.

Nocedal, J. & Wright, S.J., 2006, *Numerical Optimization,* 2nd edn, Springer.

Norris, A.N., 2006. Elastic moduli approximation of higher symmetry for the acoustical properties of an anisotropic material, *J. acoust. Soc. Am.,* **119**(4), 2114–2121.

Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., Ribodetti, A. & Virieux, J., 2015. Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation, *Geophys. J. Int.,* **202**(2), 1362–1391.

Pageot, D., Operto, S., Vallée, M., Brossier, R. & Virieux, J., 2013. A parametric analysis of two-dimensional elastic full waveform inversion of teleseismic data for lithospheric imaging, *Geophys. J. Int.,* **193**(3), 1479–1505.

Pan, W., Innanen, K.A., Margrave, G.F., Fehler, M.C., Fang, X. & Li, J., 2016. Estimation of elastic constants for HTI media using Gauss-Newton and full-newton multiparameter full-waveform inversion, *Geophysics,* **81**(5), R275–R291.

Pan, W., Geng, Y. & Innanen, K.A., 2018. Interparameter trade-off quantification and reduction in isotropic-elastic full-waveform inversion: synthetic experiments and hussar land data set application, *Geophys. J. Int.,* **213**(2), 1305–1333.

Plessix, R.-E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.,* **167**(2), 495–503.

Pratt, R.G., Shin, C. & Hick, G., 1998. Gauss–Newton and full newton methods in frequency–space seismic waveform inversion, *Geophys. J. Int.,* **133**(2), 341–362.

Raitt, R., Shor, G., Francis, T. & Morris, G., 1969. Anisotropy of the pacific upper mantle, *J. geophys. Res.,* **74**(12), 3095–3109.

Roecker, S., Baker, B. & McLaughlin, J., 2010. A finite-difference algorithm for full waveform teleseismic tomography, *Geophys. J. Int.,* **181**(2), 1017–1040.

Rusmanugroho, H., Modrak, R. & Tromp, J., 2015. Anisotropic imaging with fast recovery of tilt and azimuthal angles, in *SEG Technical Program Expanded Abstracts 2015,* pp. 4008–4012, Society of Exploration Geophysicists.

Rusmanugroho, H., Modrak, R. & Tromp, J., 2017. Anisotropic full-waveform inversion with tilt-angle recovery, *Geophysics*.

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

Schaeffer, A., Lebedev, S. & Becker, T., 2016. Azimuthal seismic anisotropy in the earth's upper mantle and the thickness of tectonic plates, *Geophys. Suppl. Mon. Not. R. Astron. Soc.,* **207**(2), 901–933.

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

Sieminski, A., Liu, Q., Trampert, J. & Tromp, J., 2007b. Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods, *Geophys. J. Int.,* **168**(3), 1153–1174.

Sieminski, A., Paulssen, H., Trampert, J. & Tromp, J., 2008. Finite-frequency SKS splitting: measurement and sensitivity kernels, *Bull. seism. Soc. Am.,* **98**(4), 1797–1810.

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., 1996. Seismic anisotropy beneath the continents: probing the depths of geology, *Ann. Rev. Earth planet. Sci.,* **24**(1), 385–432.

Simons, F.J., Van Der Hilst, R.D., Montagner, J.-P. & Zielhuis, A., 2002. Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle, *Geophys. J. Int.,* **151**(3), 738–754.

Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics,* **69**(1), 231–248.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the Southern California crust, *Science,* **325**(5943), 988–992.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49**(8), 1259–1266.

Tarantola, A., 1986. A strategy for nonlinear elastic inversion of seismic reflection data, *Geophysics,* **51**(10), 1893–1903.

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

Tong, P.,Chen,C.-w., Komatitsch, D., Basini, P. & Liu, Q., 2014a. High-resolution seismic array imaging based on an SEM-FK hybrid method, *Geophys. J. Int.,* **197**(1), 369–395.

Tong, P., Komatitsch, D., Tseng, T.-L., Hung, S.-H., Chen, C.-W., Basini, P. & Liu, Q., 2014b. A 3-D spectral-element and frequency-wave number hybrid method for high-resolution seismic array imaging, *Geophys. Res. Lett.,* **41**(20), 7025–7034.

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

van Leeuwen, T. & Herrmann, F.J., 2013. Mitigating local minima in full-waveform inversion by expanding the search space, *Geophys. J. Int.,* **195**(1), 661–667.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC1–WCC26.

Wang, Y. *et al.*, 2016. The deep roots of the western Pyrenees revealed by full waveform inversion of teleseismic p waves, *Geology,* **44**(6), 475–478.

Warner, M. & Guasch, L., 2016. Adaptive waveform inversion: theory, *Geophysics,* **81**(6), R429–R445.

Wolfe, C.J. & Silver, P.G., 1998. Seismic anisotropy of oceanic upper mantle: shear wave splitting methodologies and observations, *J. geophys. Res.,* **103**(B1), 749–771.

Wolfe, P., 1969. Convergence conditions for ascent methods, *SIAM Rev.,* **11**(2), 226–235.

Wu, W., Ni, S., Zhan, Z. & Wei, S., 2018. An SEM-DSM three-dimensional hybrid method for modelling teleseismic waves with complicated source-side structures, *Geophys. J. Int.,* **215**(1), 133–154.

Yao, H., Van Der Hilst, R.D. & Montagner, J.-P., 2010. Heterogeneity and anisotropy of the lithosphere of SE Tibet from surface wave array tomography, *J. geophys. Res.,* **115**(B12), doi:10.1029/2009JB007142.

Zhao, L. & Chevrot, S., 2011. An efficient and flexible approach to the calculation of three-dimensional full-wave Fréchet kernels for seismic tomography. theory, *Geophys. J. Int.,* **185**(2), 922–938.

Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.,* **5**(7), 493.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.
2018_BELLER_ANISOTROPICFWI_suplementary.pdf

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

APPENDIX A: SUPPLEMENTARY MATERIALS

In this supplementary material, we provide waveforms plots for the various synthetic experiments performed in this study.

A1 Sensitivity to noise

We perform a synthetic experiment considering noisy data. We add a 5 per cent Gaussian white noise to the synthetic data computed in the TTI inclusion model and invert the noisy data with a pre-conditioned anisotropic FWI. To regularize the inversion, we simply convolve the gradients with a 3-km Gaussian filter.

The convergence rate (Fig. A1) is slower than in the noise-free case. After 57 iterations, the misfit function is only reduced to 16.9 per cent of its initial value. The hexagonal part of the final model is shown in Fig. A2. The model looks noisier than than the one obtained in the noise-free case, but the TTI inclusion parameters (orientation angles, ϵ and γ) are properly recovered. This result suggests that anisotropic FWI is quite robust to the presence of noise in the data.

A2 VTI waveforms

P-wave propagation is only weakly affected by anisotropic structures in the VTI experiment (Figs A3 and A4), with weak azimuthal variations of waveforms and weak scattered wavefields. For large incidence angles (Fig A4), a weak perturbation of the transverse component amplitudes (red curves) and scattered wavefields (black curves) are still distinguishable (around 100° for $x = -50$ km and around 270° at $x = +50$ km).

For *S* waves with small incidence angle, the waveforms do not vary as a function of azimuth (Fig. A5), as expected, and show very little amount of shear wave splitting (Fig. A7). At larger incidence angles (Fig. A6), shear wave splitting becomes significant.

These results emphasize the importance of large incidence angles for both *P* and *S* waves for imaging seismic anisotropy with near vertical symmetry axis.

A3 HTI waveforms

Considering *P* waves only, the HTI model produces significant wave scattering. This is visible on Figs A8 and A9, where significant secondary arrivals occur right after the main *P* wave arrival. In addition, the three components exhibit azimuthal variations of the amplitude of the first *P* wave arrival.

As for *S* waves, the HTI model produces important shear wave splitting as indicated by the strong amplitude on transverse component waveforms (red curves on Figs A10 and A11). Fig. A12 displays the comparison between the transverse components and the time derivatives of the radial components, which show a clear $\pi/2$-periodicity.

Since both *P* and *S* waves carry information on seismic anisotropy, it is important to jointly consider *P* and *S* waves for lithospheric imaging.

A4 TTI waveforms

Waveforms for the TTI experiment are only shown for the sake of completeness since they basically present the same features as in the VTI and HTI experiments.

A5 Three-block model waveforms

The waveforms for the three-block experiment are also shown here, for the sake of completeness. Fig. A22 compares the transverse

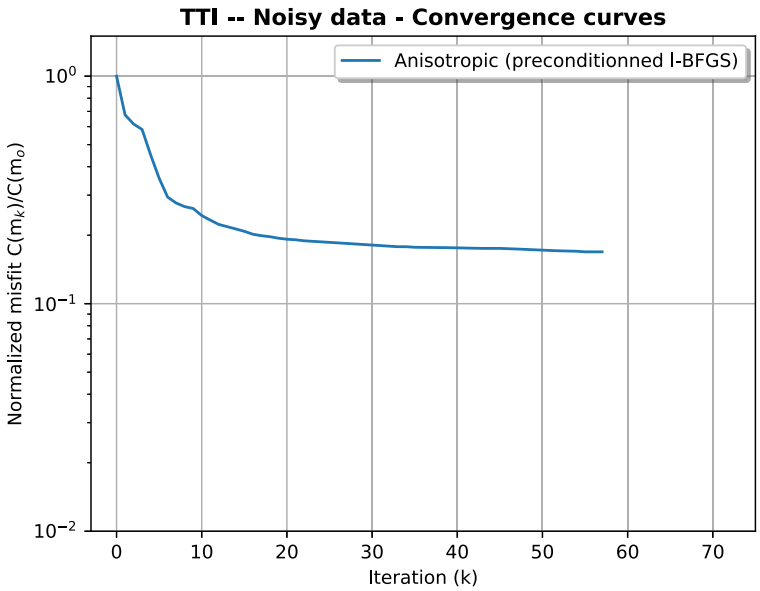

Figure A1. Convergence of anisotropic FWI for noise data.

component and the derivative of the radial component, which quantifies SKS splitting. We clearly observe that apparent splitting is extremely weak at $x = 50$ km while it becomes significant at $x = -50$ and $x = 0$ km.

The ability of FWI to recover layering of anisotropy at $x = +50$ km certainly comes from the rather complicated P waveforms (Figs A18 and A19).

APPENDIX B: TRICLINIC TENSOR DECOMPOSITION

We provide an example of elastic tensor decomposition using the triclinic tensor of Igel *et al.* (1995):

$$\mathbf{C}=\begin{pmatrix} 7.26 & 0.65 & 0.35 & 1.20 & 1.10 & 1.11 \\ 0.65 & 9.28 & 5.70 & -0.21 & -2.72 & -1.40 \\ 0.35 & 5.70 & 9.03 & -1.41 & -0.24 & -1.50 \\ 1.20 & -0.21 & -1.41 & 3.67 & 0.95 & -0.58 \\ 1.10 & -2.72 & -0.24 & 0.95 & 3.04 & 1.54 \\ 1.11 & -1.40 & -1.50 & -0.58 & 1.54 & 4.49 \end{pmatrix}. \tag{B1}$$

After decomposition, we find that its isotropic part represents 74.16 per cent of its norm:

$$\mathbf{C}_{\text{iso}}=\begin{pmatrix} 9 & 2 & 2 & 0 & 0 & 0 \\ 2 & 9 & 2 & 0 & 0 & 0 \\ 2 & 2 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3.5 \end{pmatrix}. \tag{B2}$$

Hexagonal part (2.1647 per cent):

$$\mathbf{C}_{\text{hex}}=\begin{pmatrix} -0.38 & -1.68 & 1.02 & 0 & 0 & 0 \\ -1.68 & -0.38 & 1.02 & 0 & 0 & 0 \\ 1.02 & 1.02 & 0.03 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.14 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.65 \end{pmatrix}. \tag{B3}$$

Tetragonal part (0.1651 per cent):

$$\mathbf{C}_{\text{tet}}=\begin{pmatrix} -0.33 & 0.33 & 0 & 0 & 0 & 0 \\ 0.33 & -0.33 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.33 \end{pmatrix}. \tag{B4}$$

Orthorhombic part (5.6265 per cent):

$$\mathbf{C}_{\text{ort}}=\begin{pmatrix} -1.01 & 0 & -2.67 & 0 & 0 & 0 \\ 0 & 1.01 & 2.67 & 0 & 0 & 0 \\ -2.67 & 2.67 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.31 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.31 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{B5}$$

Monoclinic part (5.2492 per cent):

$$\mathbf{C}_{\text{mon}}=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1.11 \\ 0 & 0 & 0 & 0 & 0 & -1.40 \\ 0 & 0 & 0 & 0 & 0 & -1.50 \\ 0 & 0 & 0 & 0 & 0.95 & 0 \\ 0 & 0 & 0 & 0.95 & 0 & 0 \\ 1.11 & -1.40 & -1.50 & 0 & 0 & 0 \end{pmatrix}. \tag{B6}$$

Triclinic part (12.6335 per cent):

$$\mathbf{C}_{\text{tri}}=\begin{pmatrix} 0 & 0 & 0 & 1.20 & 1.10 & 0 \\ 0 & 0 & 0 & -0.21 & -2.72 & 0 \\ 0 & 0 & 0 & -1.41 & -0.24 & 0 \\ 1.20 & -0.21 & -1.41 & 0 & 0 & -0.58 \\ 1.10 & -2.72 & -0.24 & 0 & 0 & 1.54 \\ 0 & 0 & 0 & -0.58 & 1.54 & 0 \end{pmatrix}. \tag{B7}$$

The tensor's natural or intrinsic coordinate system is:

$$\begin{pmatrix} -0.23 & 0.29 & 0.92 \\ 0.95 & 0.26 & 0.15 \\ -0.19 & 0.91 & -0.34 \end{pmatrix}. \tag{B8}$$

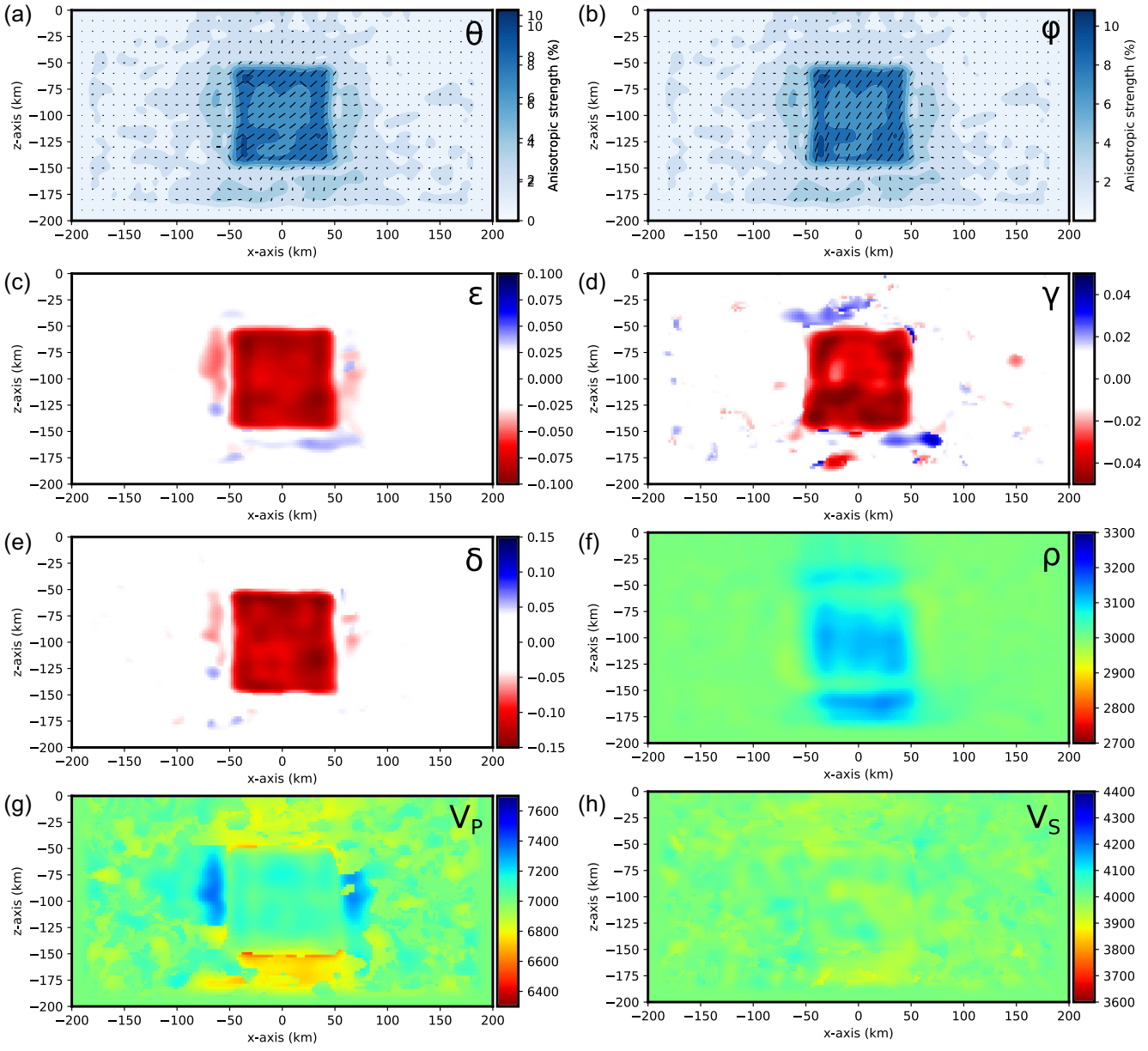

Figure A2. Anisotropic FWI sensitivity to noise. FWI model for the TTI inclusion when data are contaminated with a 5 per cent level of Gaussian white noise.

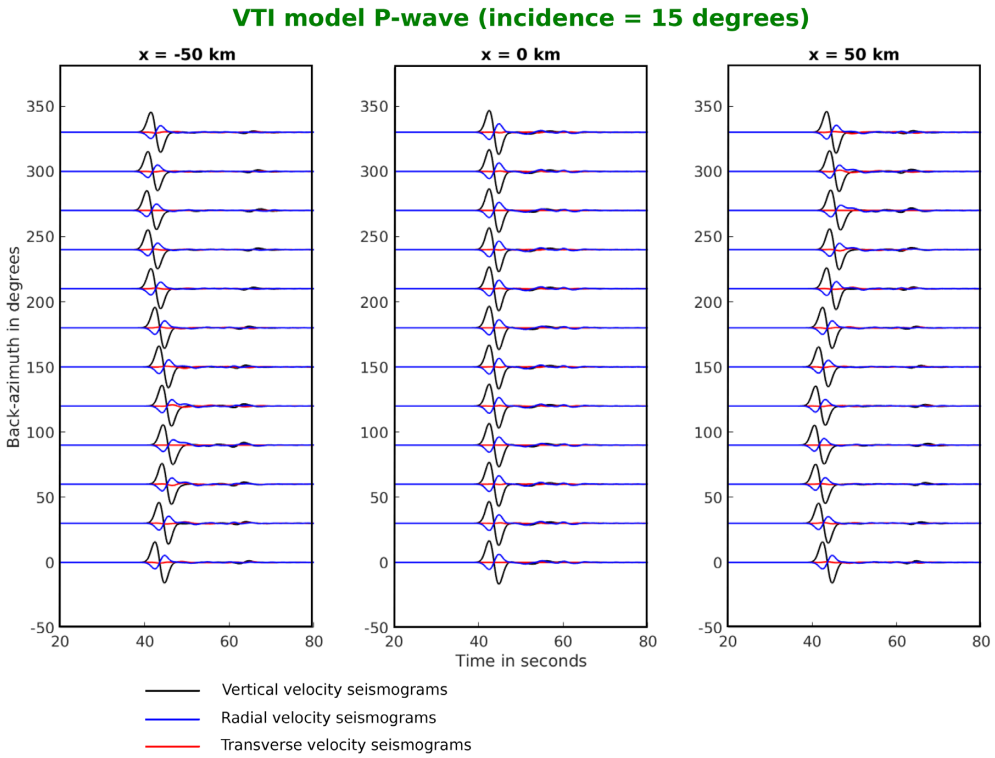

Figure A3. Synthetic waveforms computed in the VTI inclusion model for a *P* wave with 15° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively.

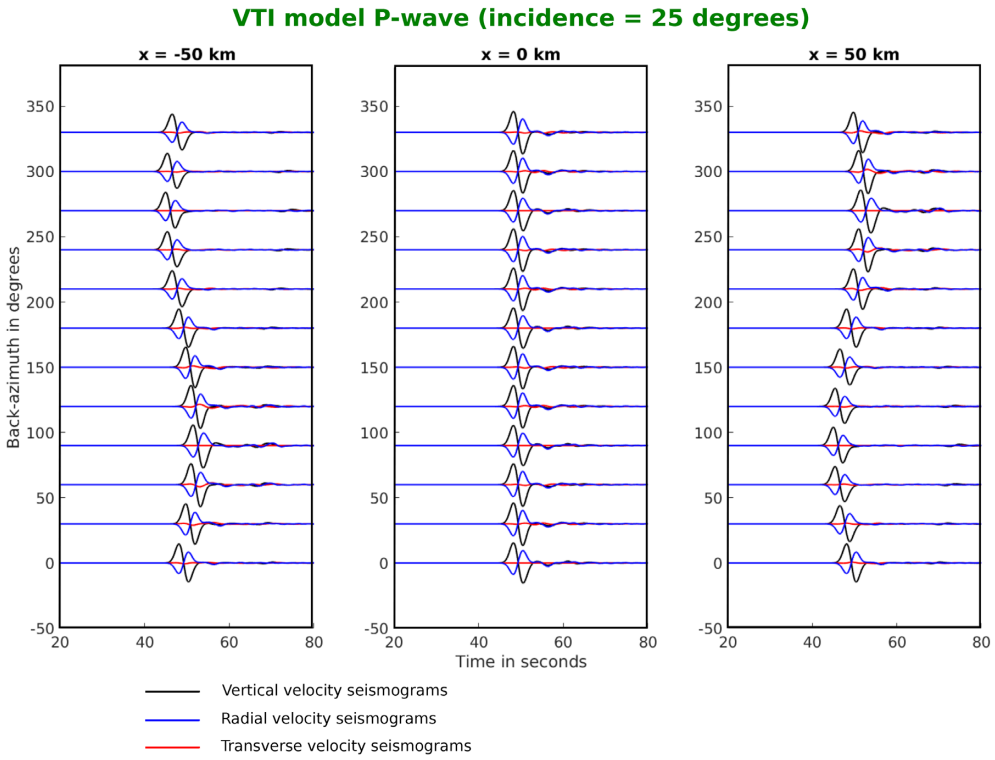

Figure A4. Synthetic waveforms computed in the VTI inclusion model for a *P* wave with 25° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively.

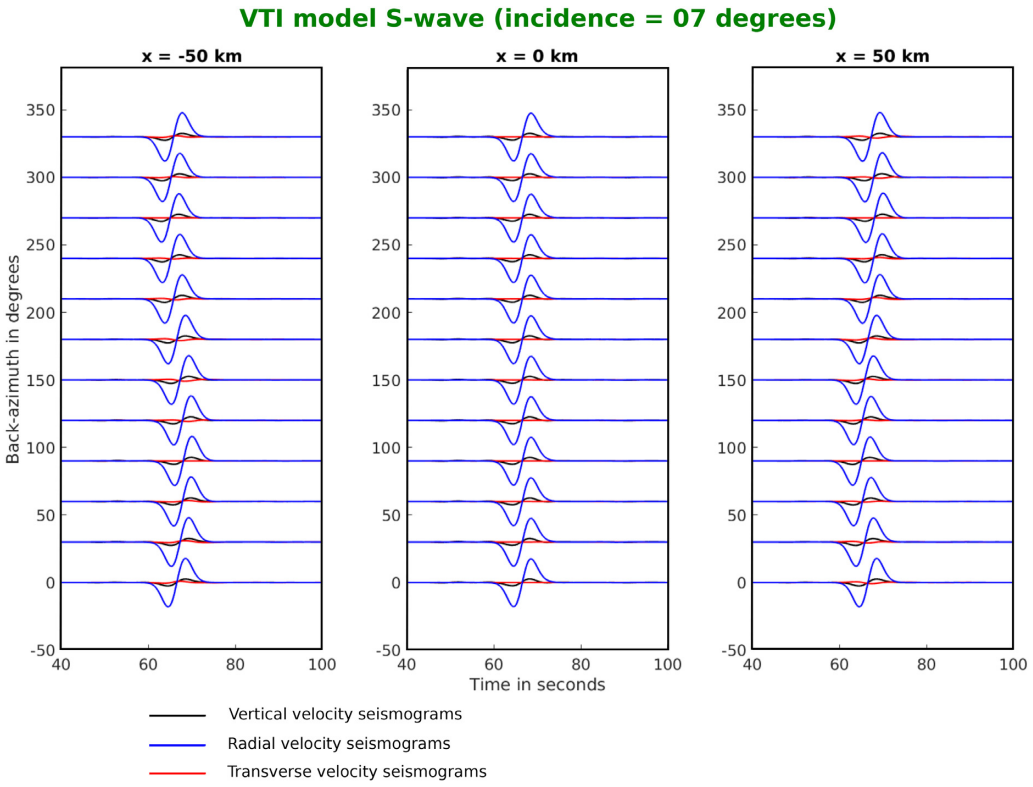

Figure A5. Synthetic waveforms computed in the VTI inclusion model for a S wave with 7° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

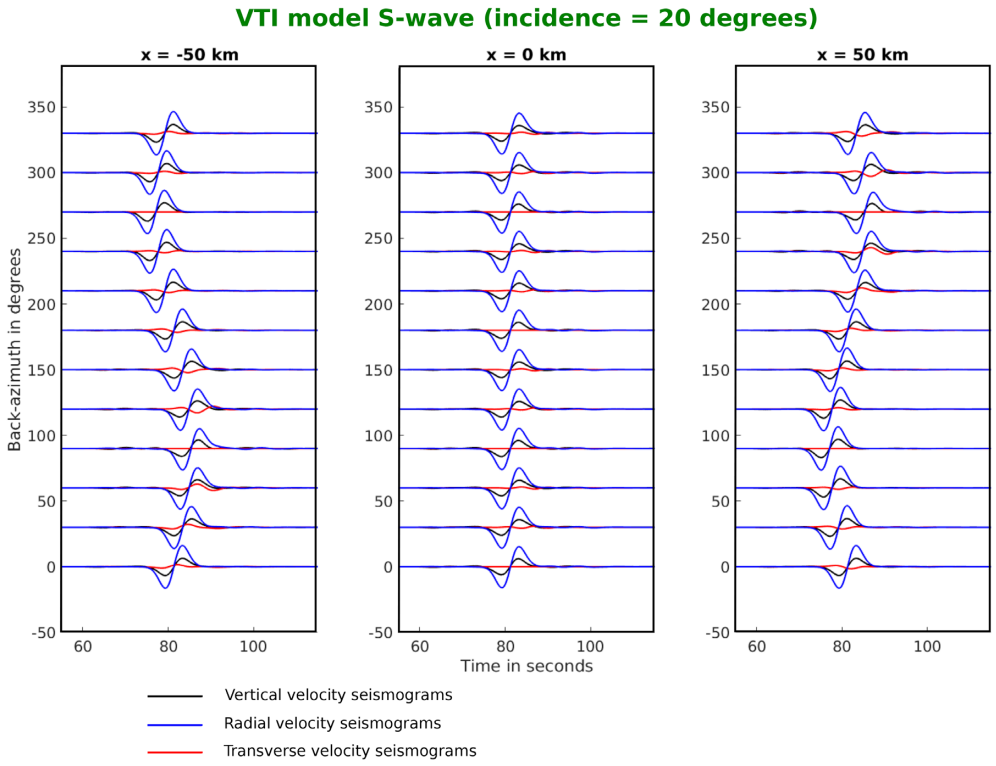

Figure A6. Synthetic waveforms computed in the VTI inclusion model for a S wave with 20° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

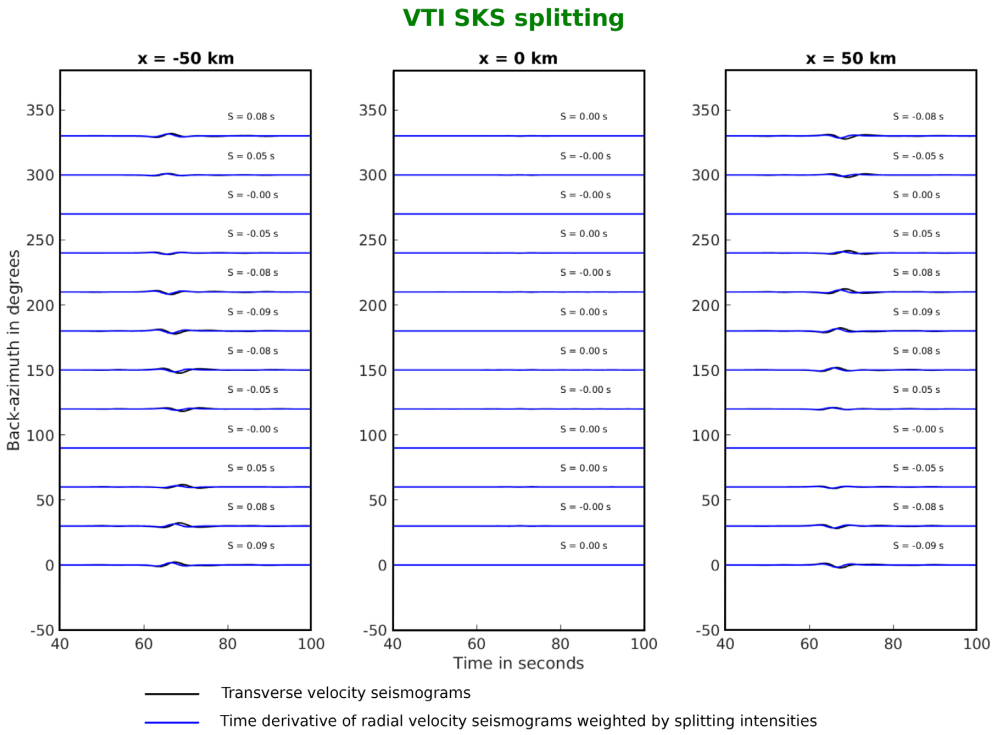

Figure A7. SKS splitting for the VTI experiment. The three panels present the azimuthal variations of the transverse component (black) and derivative of radial component weighted by the splitting intensity (blue) waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively. Splitting intensity S is given on top of each waveform.

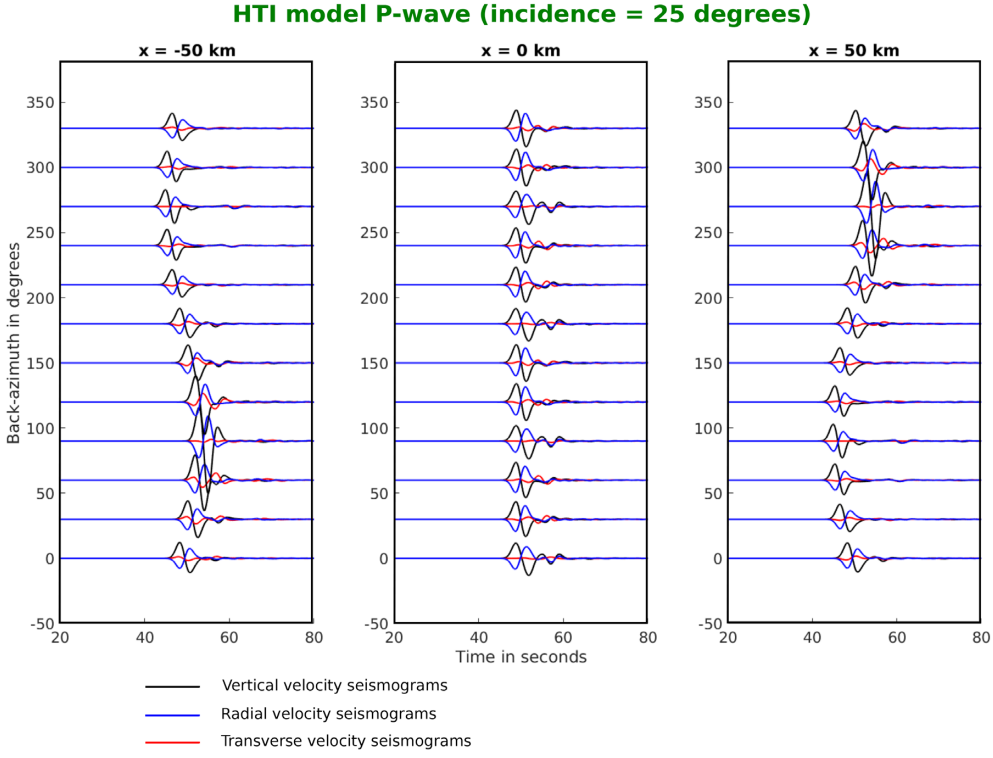

Figure A8. Synthetic waveforms computed in the HTI inclusion model for a P wave with 15° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively.

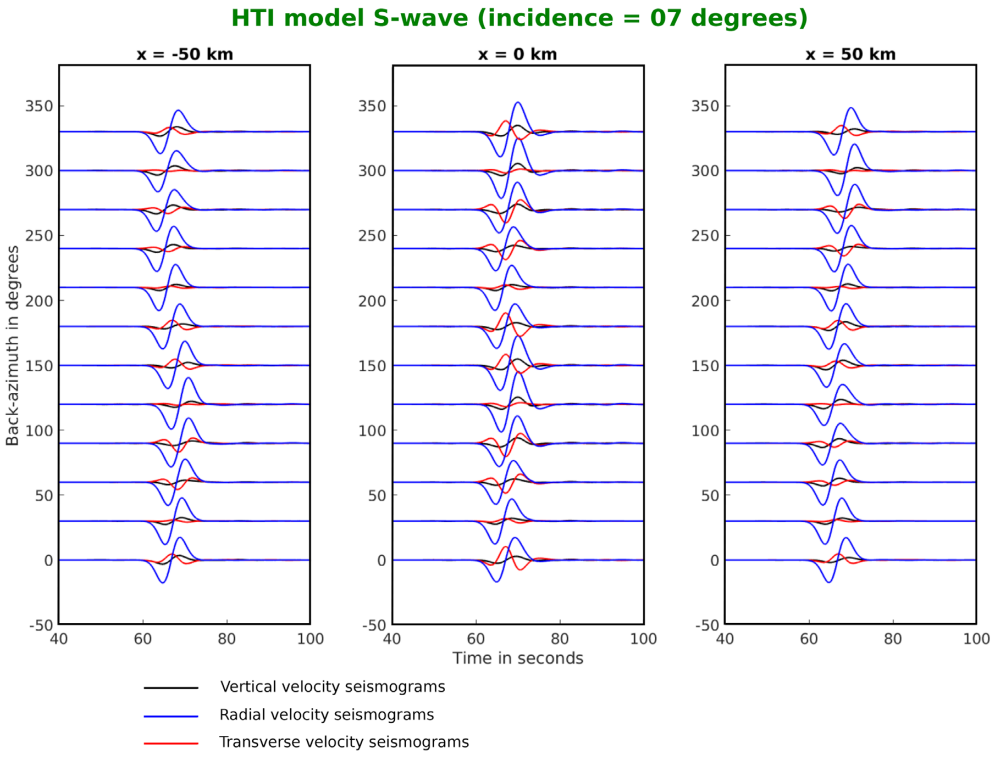

Figure A9. Synthetic waveforms computed from the HTI inclusion model for a P wave with 25° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

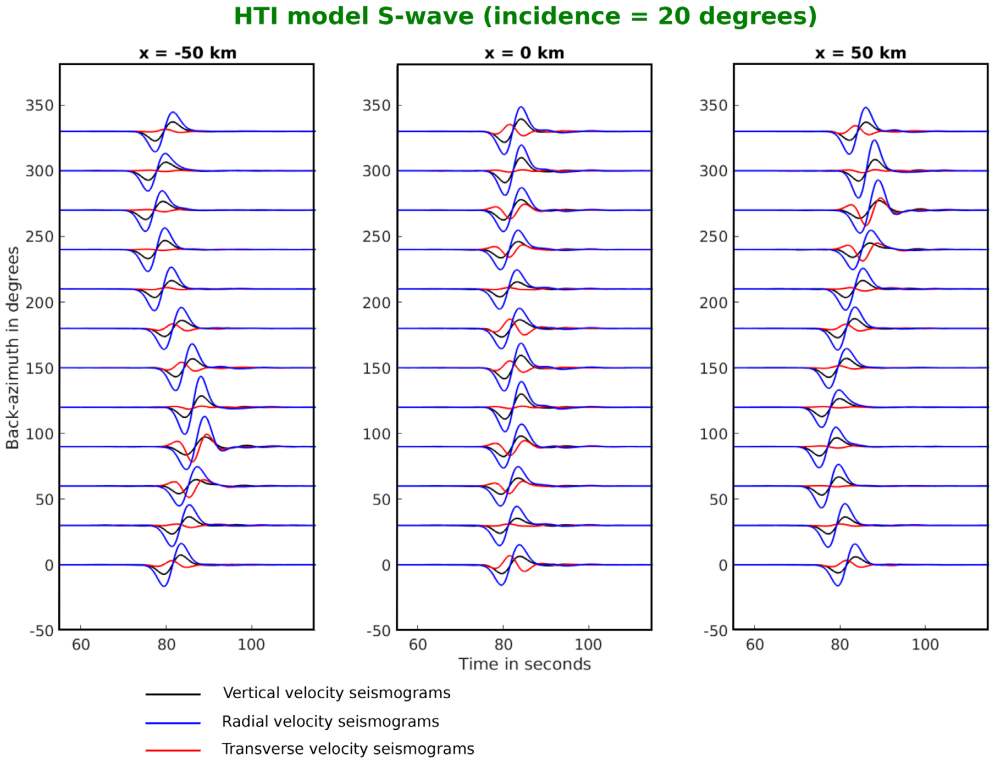

Figure A10. Synthetic waveforms computed in the HTI inclusion model for a S wave with 7° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

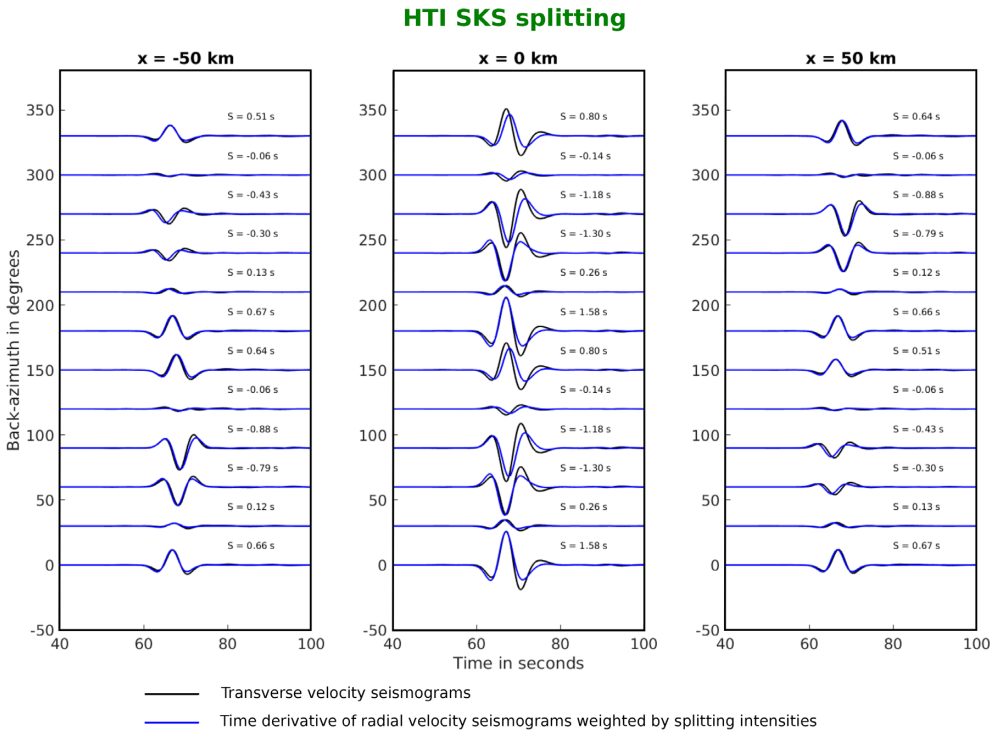

Figure A11. Synthetic waveforms computed in the HTI inclusion model for a S wave with 20° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

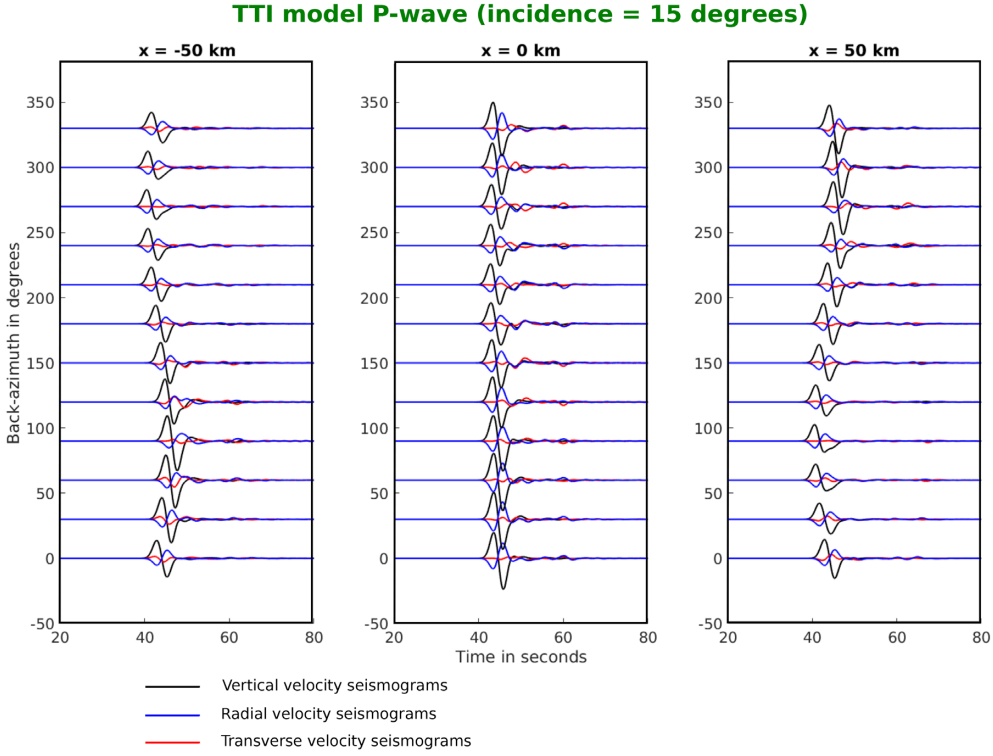

Figure A12. SKS splitting for the HTI experiment. The three panels present the azimuthal variations of the transverse (black) and derivative of radial weighted by the splitting intensity (blue) waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively. Splitting intensity S is given on top of each waveform.

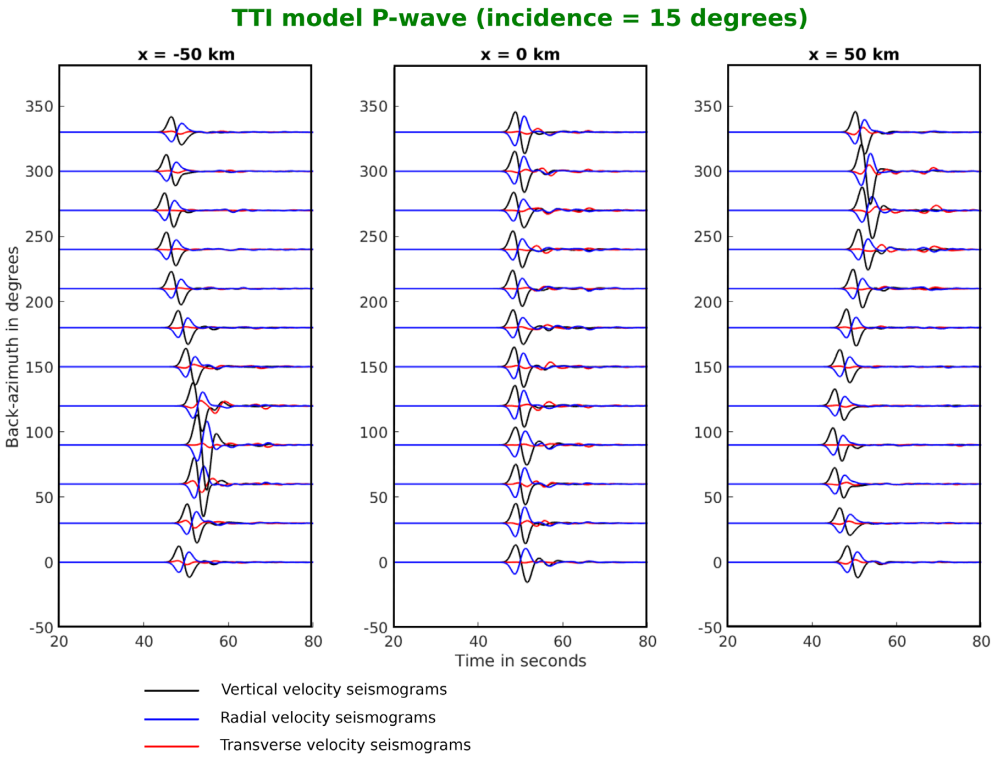

Figure A13. Synthetic waveforms computed in the TTI inclusion model for a P wave with 15° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

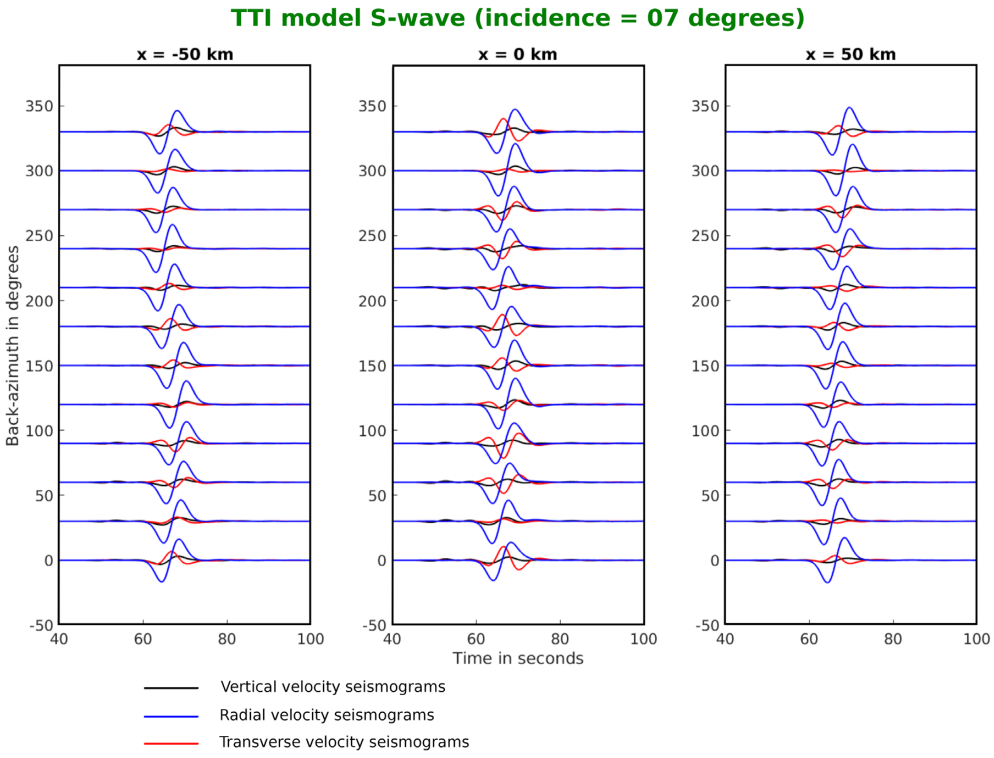

Figure A14. Synthetic waveforms computed in the TTI inclusion model for a P wave with 25° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

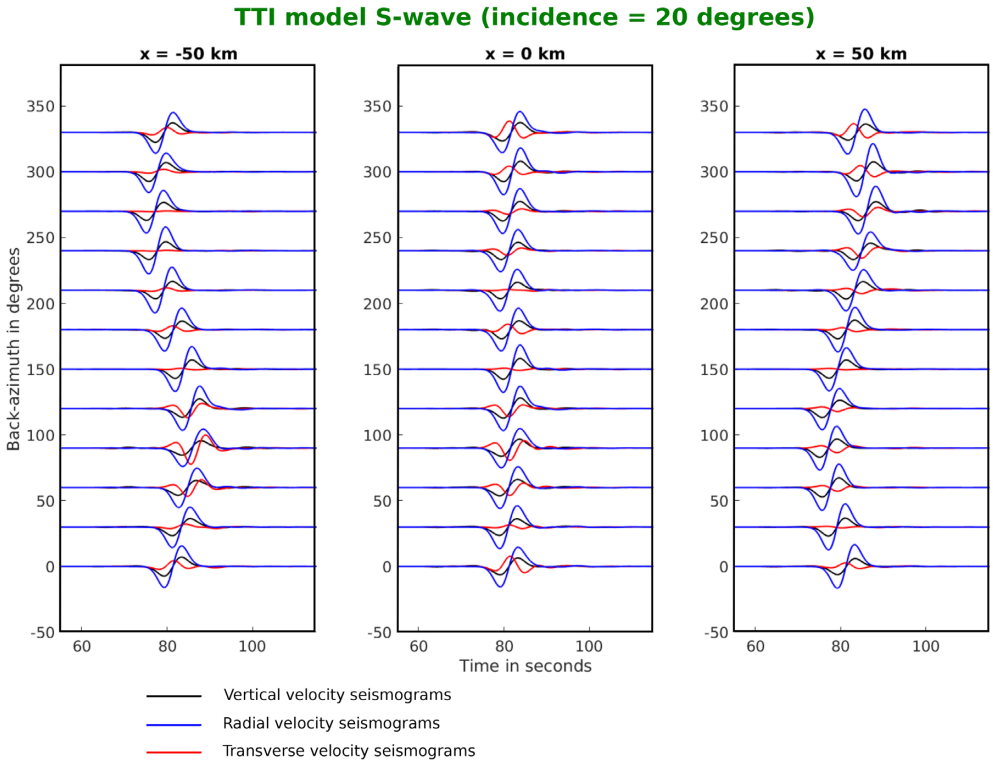

Figure A15. Synthetic waveforms computed in the TTI inclusion model for a *S* wave with 7° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

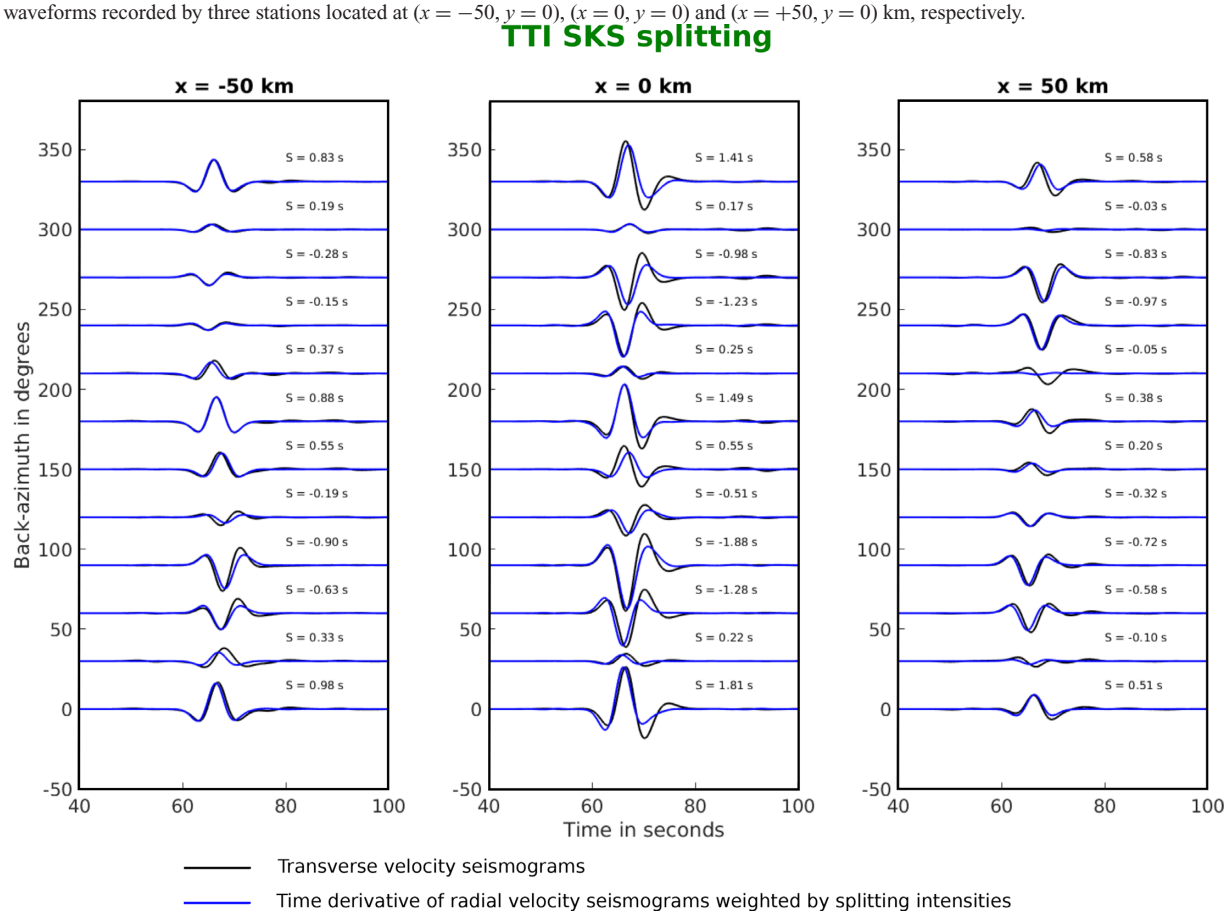

Figure A16. Synthetic waveforms computed in the TTI inclusion model for a *S* wave with 20° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

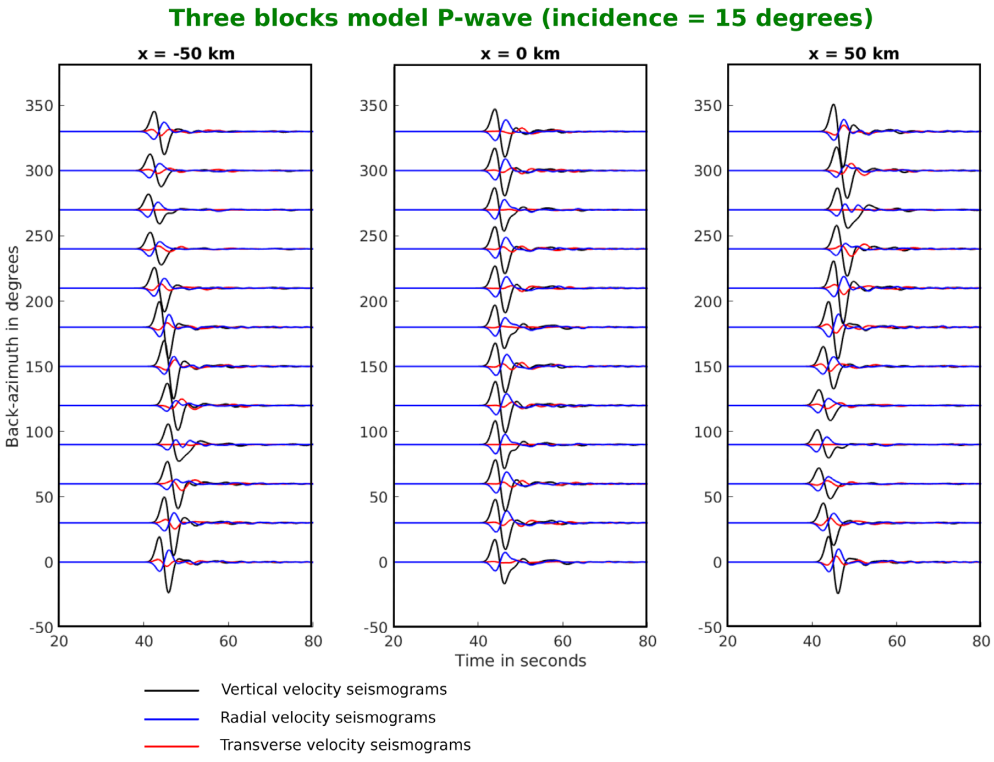

Figure A17. SKS splitting for the TTI experiment. The three panels present the azimuthal variations of the transverse (black) and derivative of radial weighted by the splitting intensity (blue) waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively. Splitting intensity S is given on top of each waveform.

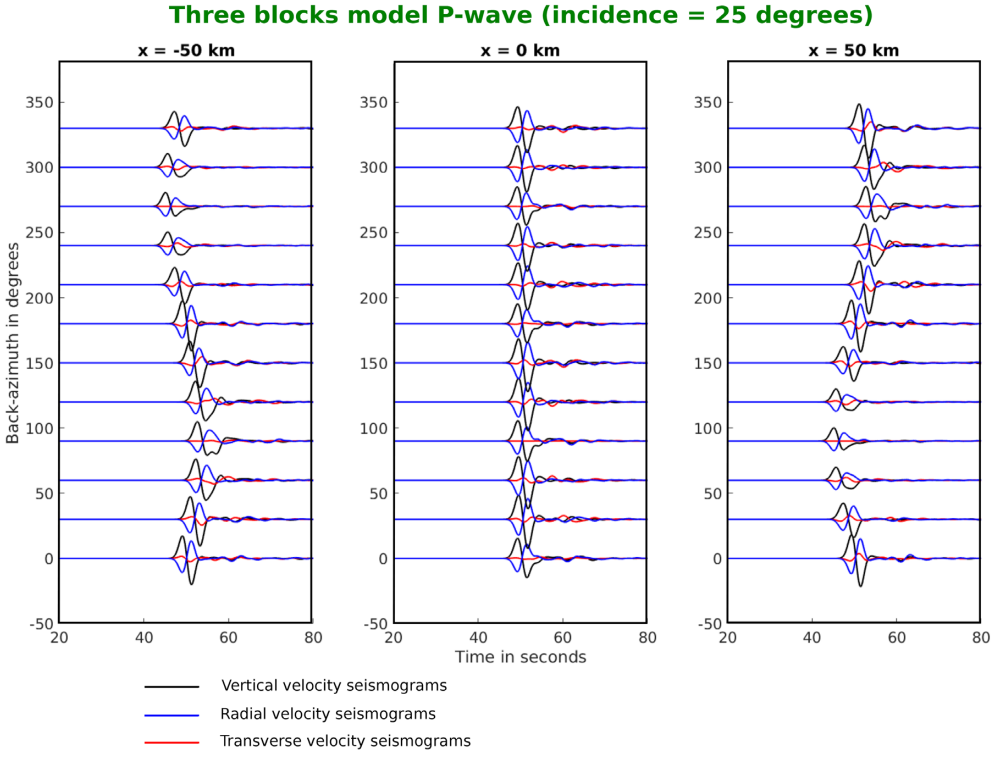

Figure A18. Synthetic waveforms computed in the three-block model for a P wave with 15° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively.

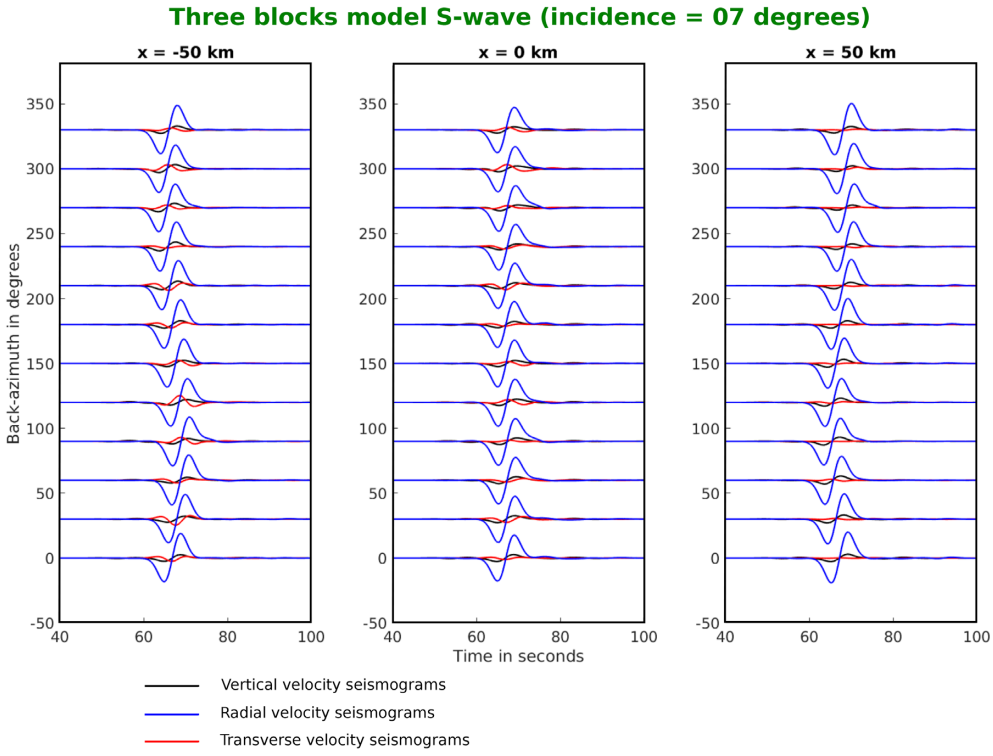

Figure A19. Synthetic waveforms computed in the three-block model for a *P* wave with 25° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively.

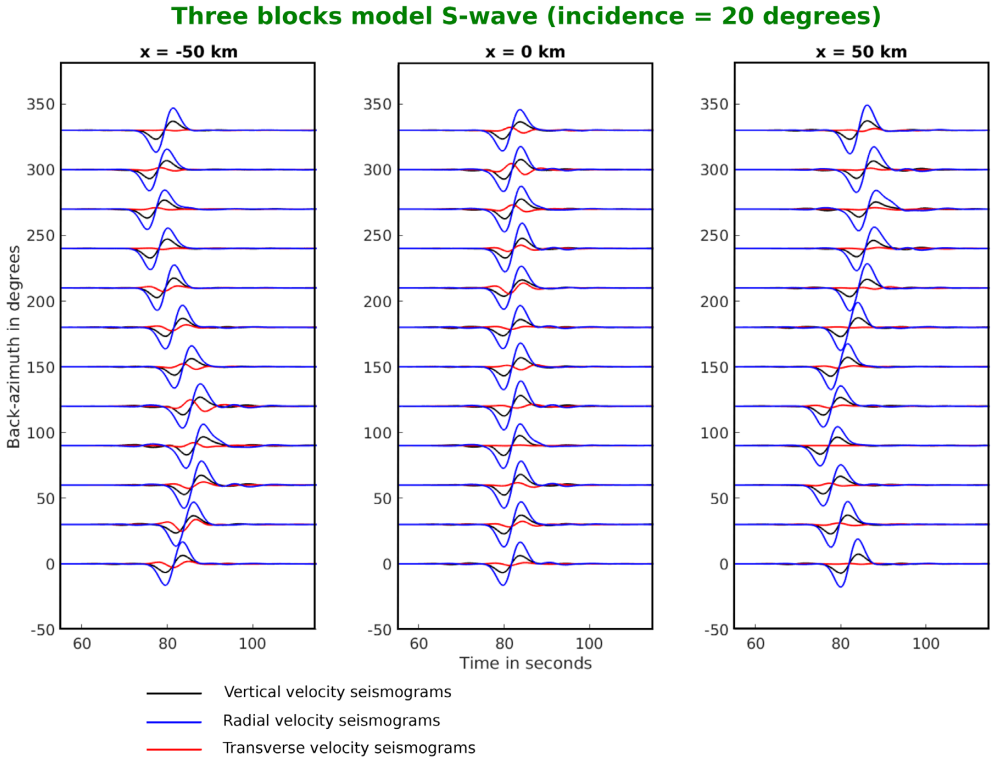

Figure A20. Synthetic waveforms computed in the three-block model for a *S* wave with 7° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at ($x = -50$, $y = 0$), ($x = 0$, $y = 0$) and ($x = +50$, $y = 0$) km, respectively.

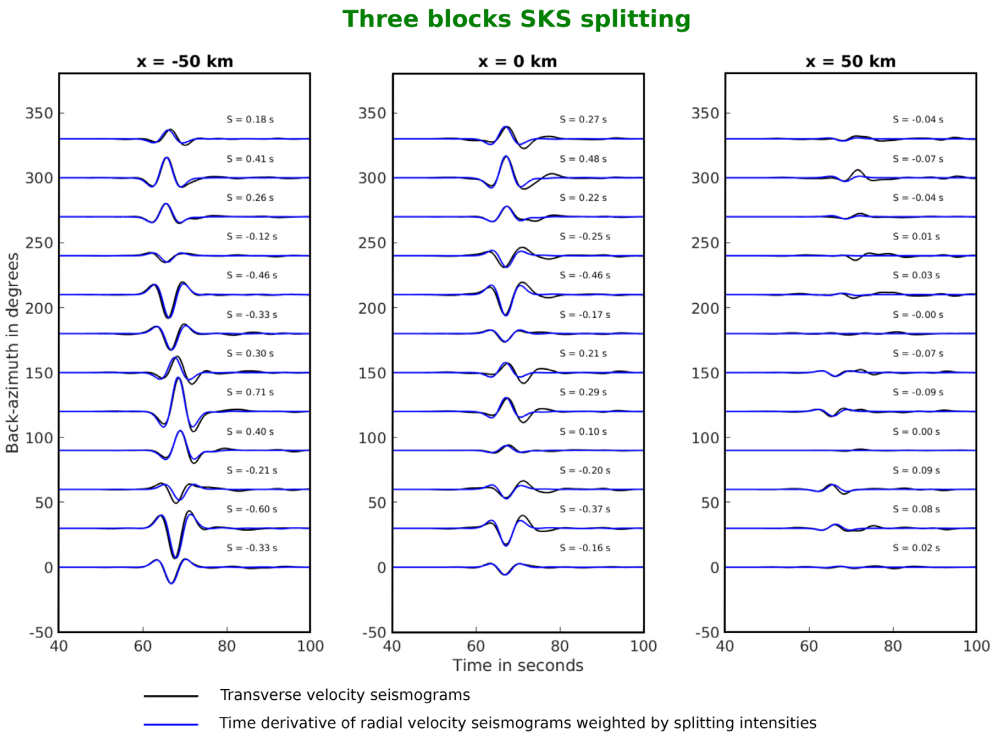

Figure A21. Synthetic waveforms computed in the three-block model for a *S* wave with 20° incidence angle. The three panels present the azimuthal variations of the waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively.

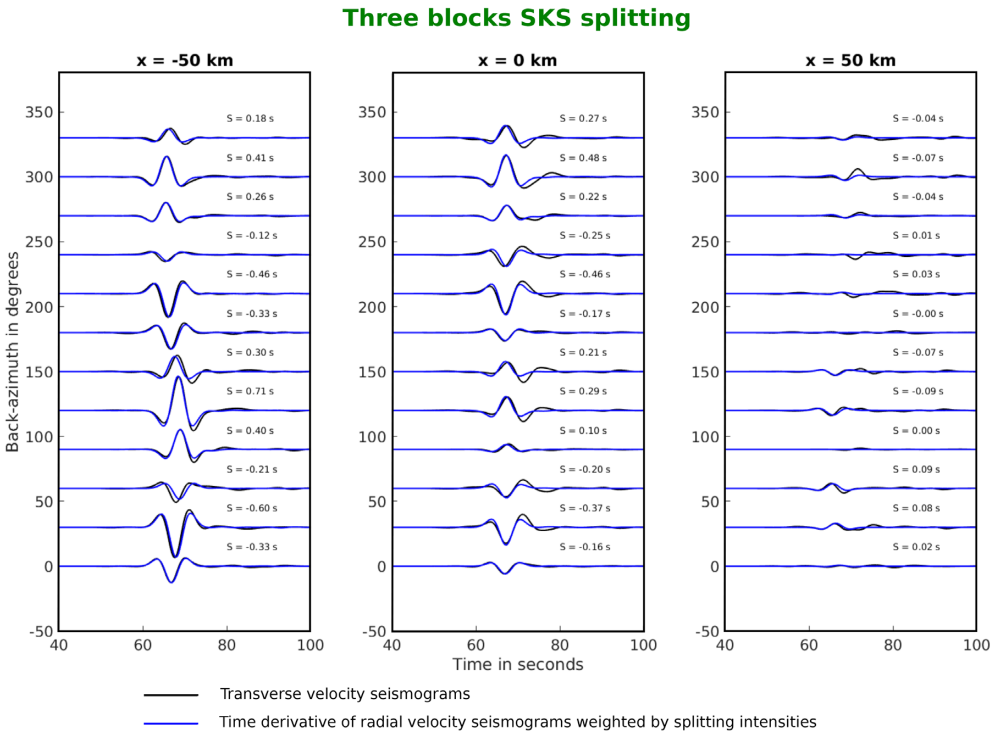

Figure A22. SKS splitting for the three-block experiment. The three panels present the azimuthal variations of the transverse (black) and derivative of radial weighted by the splitting intensity (blue) waveforms recorded by three stations located at $(x = -50, y = 0)$, $(x = 0, y = 0)$ and $(x = +50, y = 0)$ km, respectively. Splitting intensity *S* is given on top of each waveform.