

# Anisotropic P-wave traveltimes tomography implementing Thomsen's weak approximation in TOMO3D

Adrià Meléndez<sup>1</sup>, Clara Estela Jiménez<sup>1</sup>, Valentí Sallarès<sup>1</sup>, César R. Ranero<sup>2</sup>

<sup>1</sup>Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar (CSIC), Barcelona, 08003, Spain

5 <sup>2</sup>Barcelona Center for Subsurface Imaging, ICREA at Institut de Ciències del Mar (CSIC), Barcelona, 08003, Spain

Correspondence to: Adrià Meléndez ([melendez@icm.csic.es](mailto:melendez@icm.csic.es))

**Abstract.** We present the implementation of Thomsen's weak anisotropy approximation for VTI media within TOMO3D, our code for 2-D and 3-D joint refraction and reflection traveltimes tomographic inversion. In addition to the inversion of seismic P-wave velocity and reflector depth, the code can now retrieve models of the Thomsen's parameters  $\delta$  and  $\varepsilon$ . Here we  
10 test this new implementation following four different strategies on a canonical synthetic experiment [in ideal conditions with the purpose of estimating the maximum capabilities and potential weak points of our modeling tool and strategies](#). First, we study the sensitivity of traveltimes to the presence of a 25% anomaly in each of the parameters. Next, we invert for two combinations of parameters,  $(v, \delta, \varepsilon)$  and  $(v, \delta, v^\perp)$ , following two inversion strategies, simultaneous and sequential, and compare the results to study their performances and discuss their advantages and disadvantages. Simultaneous inversion is  
15 the preferred strategy and the parameter combination  $(v, \delta, \varepsilon)$  produces the best overall results. The only advantage of the parameter combination  $(v, \delta, v^\perp)$  is a better recovery of the magnitude of  $v$ . In each case we derive the fourth parameter from the equation relating  $\varepsilon$ ,  $v^\perp$  and  $v$ . Recovery of  $v$ ,  $\varepsilon$  and  $v^\perp$  is satisfactory whereas  $\delta$  proves to be impossible to recover even in the most favorable scenario. However, this does not hinder the recovery of the other parameters, and we show that it is still possible to obtain a rough approximation of  $\delta$  distribution in the medium by sampling a reasonable range of homogeneous  
20 initial  $\delta$  models and averaging the final  $\delta$  models that are satisfactory in terms of data fit.

## 1 Introduction

An isotropic velocity field is the rare exception in the Earth subsurface. Anisotropy is a multiscale phenomenon, and its causes are diverse. In the crust, it can be produced by the preferred orientation of mineral grains or their crystal axes (Schulte-Pelkum and Mahan, 2014; Almqvist and Mainprice, 2017), the alignment of cracks and fracture networks and the  
25 presence of fluids (Crampin, 1981; Maultzsch et al., 2003; Yousef and Angus, 2016), or the bedding of layers much thinner than the wavelength used to explore them (Backus, 1962; Johnston and Christensen, 1995; Sayers, 2005). In the mantle, anisotropy is related to the alignment of olivine crystals due to mantle flow (Nicolas and Christensen, 1987; Montagner et al., 2007), aligned melt inclusions (Holtzman et al., 2003; Kendall et al., 2005), large-scale deformation (Vinnik et al., 1992; Vauchez et al., 2000), and preexisting lithospheric fabric (Kendall et al., 2006), among others. Anisotropy has proven an

30 informative physical property in the understanding of the Earth's interior (Ismail and Mainprice, 1998; Long and Becker, 2010), most particularly in continental rifts (Eilon et al., 2016), mid-ocean ridges (Dunn et al., 2001), and subduction zones (Long and Silver, 2008).

The theory of anisotropic wave propagation has been described in several publications (Kraut, 1963; Babuska and Cara, 35 1991). Numerous formulations of varying complexity have been proposed to approximate anisotropy depending on its magnitude and the symmetry conditions of the medium (Nye, 1957). Twenty-one elastic stiffness parameters define the most general anisotropic medium with the lowest symmetry conditions, whereas for the highest symmetry not equivalent to isotropy, only five parameters are needed (Almqvist and Mainprice, 2017). Regarding the strength of anisotropy, in view of the overall success of isotropic methods in studying the Earth's subsurface, and of the experimental evidence and sample 40 measurements available, it is admitted that the anisotropy is generally weak (Thomsen, 1986). Specifically, anisotropy is considered weak when Thomsen's parameters are much smaller than 1, i.e. for a ~20% or smaller velocity variation with angle. Precisely Thomsen (1986) presented the formulation for the transverse isotropy symmetry on weakly anisotropic media, which is the reference that we follow in this work. Thomsen's parameters are by far the most common and convenient combinations of stiffness tensor elements used in seismic anisotropy modeling (Tsvankin, 1996; Thomsen and Anderson, 45 2015). Applications of simpler approximations exist, as is the case of elliptical symmetry (Song et al., 1998; Giroux and Gloaguen, 2012), as well as others that assume the most general anisotropic model (Zhou and Greenhalgh, 2008).

The objectives of this work are (1) presenting the anisotropic version of TOMO3D (Meléndez et al., 2015) for the study of VTI weakly-anisotropic media in terms of the Thomsen's parameters  $\delta$  and  $\epsilon$  using P-wave arrival times, and (2) comparing 50 several parametrizations and inversion strategies ~~to identify those providing the most accurate results under optimal and equal conditions for all parameters with the purpose of defining an upper limit to the code's capabilities, an ideal but generalizable estimation of the code's performance, as well as highlighting its potential weaknesses. The P-wave velocity,  $\delta$  and  $\epsilon$  models obtained from the modeling of field data would be useful and geologically informative by themselves, but they would also serve as initial models in anisotropic FWI.~~ Moreover, the development of this code is motivated by the need to 55 combine wide-angle and near-vertical traveltimes picks in field data applications, as we plan to do with the trench-parallel 2-D profile in Sallarès et al. (2013), which is affected by a ~15% anisotropy judging from the mismatch in the interplate boundary locations obtained separately from near-vertical and wide-angle data. The P-wave velocity,  $\delta$  and  $\epsilon$  models obtained from the modeling of field data would be useful and geologically informative by themselves, but they would also serve as initial models in anisotropic FWI.

60 In the following section, we describe the anisotropy formulation and the modifications implemented on our 3-D joint refraction and reflection traveltimes tomography code TOMO3D to incorporate the inversion of Thomsen's parameters. Next, in the third section, we present the synthetic tests performed and their results, including accuracy and sensitivity analyses and

synthetic inversions. In section four, these results are discussed and interpreted in terms of the ability of the code to retrieve  
65 both the velocity field and Thomsen's anisotropy parameters. Finally, in the last section we summarize the main conclusions  
of this work.

## 2 Modeling anisotropy

The first part of this section is a general overview of the treatment of anisotropy within the field of seismic inversion, while  
the second one describes the implementation of Thomsen's weak [VTI](#) anisotropy formulation in the 3-D joint refraction and  
70 reflection traveltime tomography code TOMO3D.

### 2.1 Anisotropy in seismic inversion methods

Anisotropy was first incorporated to seismic inversion methods in traveltime tomography with the development of the  
linearized perturbation theory (Cerveny, 1982; Cerveny and Jech, 1982; Jech and Psencik, 1989). Previously, the approach to  
deal with anisotropy was to approximately remove its estimated effect to then apply an isotropic method (e.g. McCann et al.,  
75 1989). Linearized perturbation theory was first implemented in anisotropic traveltime tomography by Chapman and Pratt  
(1992) and Pratt and Chapman (1992) assuming the weak anisotropy approximation, which allowed them to use isotropic ray  
tracing and approximate anisotropy effects as being caused by small perturbations of the isotropic system. The initial  
development of anisotropic ray tracing is attributed to Cerveny (1972). Methods for anisotropic ray tracing and traveltime  
computation depend on the symmetry assumptions made regarding the medium. The most common of those is rotational  
80 symmetry around a vertical pole. This formulation is known as vertical transverse isotropy (VTI), also polar anisotropy (e.g.  
Rüger and Alkhalifah, 1996; Alkhalifah, 2002), and it is the simplest geologically applicable case: it reproduces the  
symmetry exhibited by minerals in sedimentary rocks, and that produced by parallel cracks or fine layering. Furthermore, it  
significantly simplifies the mathematical formulae since anisotropy is defined by only five parameters, which contributes to  
a greater computational efficiency. The generalization of VTI to a tilted symmetry axis is the so called tilted transverse  
85 isotropy (TTI). Some authors argue that it is not possible to distinguish TTI from VTI in real experimental cases without a  
priori information (Bakulin et al., 2009). Assuming the most general anisotropic media has also become rather usual, in  
particular with the improvement of computational resources, allowing for a more detailed and complex reconstruction of the  
subsurface physical properties (e.g. Zhou and Greenhalgh, 2005) although successful field data applications are yet to be  
achieved to the best of our knowledge. Regarding the inversion process, the main difficulty arises from the trade-off between  
90 velocity heterogeneity and anisotropy (e.g. Bezada et al., 2014). To the best of our knowledge, Stewart (1988) was the first  
to propose an inversion algorithm, specifically for the recovery of Thomsen's parameters in a weakly-anisotropic VTI  
medium. Other authors have produced inversion algorithms for different formulations such as azimuthal anisotropy (e.g.  
Eberhart-Phillips and Henderson, 2004; Dunn et al., 2005) or a 3-D TTI medium (e.g. Zhou and Greenhalgh, 2008).  
Concerning FWI, anisotropy in active data is typically modeled following Thomsen's parameters and the VTI and/or TTI

95 approximation for the medium. The first anisotropic wave propagators appeared during the 80s and 90s (e.g. Helbig, 1983; Alkhalifah, 1998) and new improvements on this matter continue today (e.g. Fowler et al., 2010; Duveneck and Bakker, 2011). When performing anisotropic full waveform inversion (FWI), both 2-D and 3-D, some authors choose to invert only for the velocity field, fixing the initial anisotropy models throughout the inversion because it simplifies the process (e.g. Prioux et al., 2011; Warner et al., 2013). However, other works have explored the feasibility of multiparameter inversions, 100 that is, using different combinations of velocity and anisotropy parameters and of inversion strategies (e.g. Gholami et al., 2013a,b; Alkhalifah and Plessix, 2014).

## 2.2 Anisotropy in TOMO3D: Thomsen's weak [VTI](#) anisotropy formulation

We adapted TOMO3D (Meléndez et al., 2015) to perform anisotropic ray tracing and traveltime calculations, as well as inversion of Thomsen's parameters for P-wave data. In TOMO3D, the forward problem solver is parallelized to 105 simultaneously trace rays for multiple sources and receivers, and it uses an hybrid ray tracing algorithm that combines the graph or shortest path method (Moser ~~et al.~~, 1991) and the bending refinement method (Moser [et al.](#), 1992). The inverse problem is solved sequentially using the LSQR algorithm (Paige and Saunders, 1982). Velocity models are discretized as 3-D orthogonal and vertically-sheared grids that can account for topography and/or bathymetry. [Velocity values are assigned to the grid nodes, and the velocity field is built by trilinear interpolation within each cell.](#) Apart from first-arrival traveltimes, 110 the code allows for the inversion of reflection traveltimes to obtain the geometry of major geological boundaries associated to impedance contrasts that produce strong seismic energy reflections in the data recordings. Such reflecting interfaces are modeled as 2-D grids independent of the velocity grid. The code is also prepared to extract information from the water-layer multiple of refracted and reflected seismic phases (Meléndez et al., 2014~~3~~). A detailed description of the code can be found in Meléndez (2014).

115 Our anisotropy formulation is based on Thomsen (1986) and specifically in the following weakly-anisotropic velocity equation for the P-wave velocity:

$$v^a(v, \delta, \varepsilon, \theta) = v \cdot (1 + \delta \cdot \sin^2(\theta) \cdot \cos^2(\theta) + \varepsilon \cdot \sin^4(\theta)) \quad (1)$$

120 where  $v^a$  is the anisotropic velocity,  $v$  is the velocity along the symmetry axis ( $\alpha_0$  in Thomsen, 1986),  $\theta$  is the angle with respect to the symmetry axis, and  $\delta$  and  $\varepsilon$  are Thomsen's parameters controlling the anisotropic P-wave propagation. Studying the cases of  $\theta=0$  and  $\theta=\pi/2$  the meaning of  $\varepsilon$  becomes clear: it is the relative difference between the velocities along and across the symmetry axis, that we refer to as parallel and perpendicular velocities, respectively.

125  $v^a(v, \theta = 0) = v$

$$v^a \left( v, \theta = \frac{\pi}{2} \right) = v \cdot (1 + \varepsilon) \equiv v^\perp$$

$$\varepsilon = (v^\perp - v)/v \tag{2}$$

130 According to Thomsen (1986) ~~the meaning of  $\delta$  is far from intuitive but the author states that it is related-associated~~ to the near-vertical anisotropic response and ~~shows its meaning is far from intuitive. However, that it relates a mathematical relationship between  $\delta$ ,  $v$ , and the normal move-out velocity ( $V_{\text{NMO}}$ ) exists.~~  $V_{\text{NMO}}$  models are ~~a by-product of multichannel seismic reflection data processing, a mathematical construct that involves the assumptions of a stratified media with constant velocity layers and of small spread, i.e. near-vertical propagation built as part of the normal move out correction in seismic reflection data processing.~~ It does not seem wise to try estimating it by other means, less so if the data and modeling used do not necessarily fulfill the assumptions for the normal move-out correction that define  $V_{\text{NMO}}$ . Moreover, we want to combine traveltimes from as many types of seismic data sets as possible, notably from multichannel reflection (near-vertical propagation) and wide-angle (sub-horizontal propagation) experiments, in order to have the best polar coverage, and with that, the best recovery of  $v$  and  $\varepsilon$  (or  $v^\perp$ ). ~~At best, our traveltime tomographic method would be able to produce approximations of the actual  $V_{\text{NMO}}$  models. Thus, we do not consider  $V_{\text{NMO}}$  to be a useful parameter in describing the general anisotropic VTI media for our modeling method, and we did not implement parametrizations  $(v, V_{\text{NMO}}, \varepsilon)$  and  $(v, V_{\text{NMO}}, v^\perp)$ . We do think however that  $V_{\text{NMO}}$  models can help in the building of initial  $\delta$  models, just as the comparison between near-vertical propagation and sub-horizontal data can provide an initial estimation of  $\varepsilon$ . Furthermore, such approximations would only be meaningful, if ever, when derived from travel times of a seismic reflection data set, for which the normal move out correction and thus the  $V_{\text{NMO}}$  are defined. Of course, in such a case, actual  $V_{\text{NMO}}$  models would be obtained from the normal move out correction, and therefore  $\delta$  could be calculated provided that a  $v$  model is available, for instance from our traveltime tomography. Thus, we~~ In conclusion, we only considered Eq. (2), and we implemented two parametrizations of the medium:  $(v, \delta, \varepsilon)$  and  $(v, \delta, v^\perp)$ . From here on, for simplicity, we will refer to them as  $P[\varepsilon]$  and  $P[v^\perp]$  respectively.

150 The linearized inverse problem matrix equation including anisotropy parameters for a refraction-only case is as follows



$$t = \sum_{i=1}^N u_i^\alpha \cdot s_i \quad (5)$$

$$175 \quad \frac{\partial t}{\partial u_i^\alpha} = s_i \quad (6)$$

From Eq. (1) the first order partial derivatives of the anisotropic slowness with respect to the model parameters ( $u$ ,  $\delta$ ,  $\varepsilon$ , and  $v^\perp$ ) are as follows

$$180 \quad \frac{\partial u^\alpha}{\partial u} = \frac{1}{1 + \delta \cdot \sin^2(\theta) \cdot \cos^2(\theta) + \varepsilon \cdot \sin^4(\theta)} \quad (7)$$

$$\frac{\partial u^\alpha}{\partial \delta} = \frac{-u \cdot \sin^2(\theta) \cdot \cos^2(\theta)}{(1 + \delta \cdot \sin^2(\theta) \cdot \cos^2(\theta) + \varepsilon \cdot \sin^4(\theta))^2} \quad (8)$$

$$\frac{\partial u^\alpha}{\partial \varepsilon} = \frac{-u \cdot \sin^4(\theta)}{(1 + \delta \cdot \sin^2(\theta) \cdot \cos^2(\theta) + \varepsilon \cdot \sin^4(\theta))^2} \quad (9)$$

185

$$\frac{\partial u^\alpha}{\partial v^\perp} = \frac{-(u)^2 \cdot \sin^4(\theta)}{(1 + \delta \cdot \sin^2(\theta) \cdot \cos^2(\theta) + \varepsilon \cdot \sin^4(\theta))^2} \quad (10)$$

If the  $n$  th pick corresponds to a reflected ray, then its traveltme residual is related to changes in model parameters by the following equations

190

$$\Delta t_n^\pm = \Delta t_n^0 + \sum_{m=1}^4 r_m^z \cdot \frac{\partial t}{\partial z} \cdot \Delta z_m \quad (11)$$

with

$$\frac{\partial t}{\partial z} = \sum_{m=1}^4 r_m^z \cdot (u^\alpha(\theta_i) + u^\alpha(\theta_r)) \cdot \cos(\eta) \cdot \cos(\xi) \quad (12)$$

195

where the additional term corresponds to the depth kernel which has been modified to account for anisotropy from its isotropic version as derived by Bishop et al (1985).  $u^\alpha(\theta_i)$  and  $u^\alpha(\theta_r)$  are the anisotropic slownesses at the reflecting point on the interface for the incident and reflected rays respectively,  $\eta$  is the angle of the interface with respect to the horizontal, and  $\xi$  is the incidence angle with respect to the interface normal vector. For simplicity, in this work we focus the analysis on first arrival inversion so we do not use reflection picks.

### 200 3. Synthetic tests

We have performed a number of tests using canonical synthetic models made of an anomaly centered in a uniform background with two main objectives: (1) checking that the newly implemented anisotropic traveltimes tomography method works properly, and (2) providing a quantitative measure of the potential recovery of anisotropy based on P-wave traveltimes alone. ~~To test this we first calibrated the code by comparing the synthetic data that it generates to analytically calculated data.~~ ~~Second~~First, we run a sensitivity test to assess the effect that a variation in each model parameter has in the synthetic traveltimes, ~~and we calibrated the code by comparing the synthetic data that it generates to analytically calculated data.~~ ~~Finally~~Next, we performed a number of synthetic inversion tests considering both possible parametrizations and inversion strategies. ~~These tests are conducted under ideal and equal conditions for all parameters with the purpose of obtaining an upper-limit but widely-applicable estimation of the code's performance, and detecting any potential weak points.~~

210

The models in all these tests are cubes of 5-km-long edge. The background model of all four parameters is set to a constant value, i.e.  $v$ ,  $\delta$ ,  $\varepsilon$  and  $v^\perp$  background models are homogeneous. Note that the z axis positive direction points downwards. Grid spacing is 0.125 km for the four parameters in all three dimensions so that differences in model discretization do not influence the test results. The volume of the anomaly is determined by the  $3\sigma$  region of a 3-D Gaussian function centered in the cube setting  $3\sigma = 0.5$  km. The values of  ~~$v$ ,  $\delta$ ,  $\varepsilon$  and  $v^\perp$  the grid nodes~~ within this volume are homogeneously increased resulting in a discretized representation of a spherical anomalous body.

215

#### 3.1 Accuracy

~~Alternately for each of the four parameters, the said anomaly was added at the center of the cube representing a 25% increase on the background value, while the models for the rest of parameters remained homogeneous. Table 1 summarizes the background and anomaly values for all parameters.~~

220

~~The acquisition configuration consists of 482 diametrically opposed source-receiver pairs. Each receiver records exclusively the first arrival traveltimes from its corresponding source for a total of 482 traveltimes (Fig. 1). Sources and receivers are located at 2.5 km of the center of the model, at the locus defined by the surface of the sphere inscribed in the cube, and placed at the crossing points of 32 meridians with 15 parallels, and at each pole.~~

225

~~In these four cases, we compared the synthetic traveltimes obtained with our code to the analytic solution. The comparison along two selected meridians at 0 rad and  $\pi/4$  rad azimuths is displayed in Fig. 2. Table 2 contains the mean traveltimes misfits relative to the analytic traveltimes in percentage and their respective mean deviations, for each of these four tests and along the two selected meridians, as well as the overall values.~~

230

### 3.12 Sensitivity

We define the sensitivity of a parameter as the difference between the first arrival traveltimes with and without the anomaly in that parameter. Figure 2 shows synthetic and analytic sensitivities for two selected meridians. We express sensitivity both as normalized and as relative traveltime difference. In its normalized form, traveltime differences are divided by the greatest of these differences among all parameters, i.e. normalized to 1, whereas in its relative form, these differences are given with respect to the traveltime without the anomaly and multiplied by 100. Note that the analytic response does not change between meridians, i.e. given the symmetry of the models and of the anisotropic formulation, sensitivity is independent of the azimuth angle.

The acquisition configuration consists of 482 diametrically-opposed source–receiver pairs. Each receiver records exclusively the first arrival traveltime from its corresponding source for a total of 482 traveltimes (Fig. 1). Sources and receivers are located at 2.5-km of the center of the model, at the locus defined by the surface of the sphere inscribed in the cube, and placed at the crossing points of 32 meridians with 15 parallels, and at each pole.

According to the definition of sensitivity, alternately for  $v$ ,  $\delta$ , and  $\varepsilon$ , the said anomaly was added at the center of the cube representing a 25% increase on the background value, while the models for the rest of parameters in the parametrization remained homogeneous. Table 1 summarizes the background and anomaly values for all parameters in each sensitivity test, along with their equivalence in the alternative parametrization. Since  $v$  is related to  $\varepsilon$  and  $v^\perp$  through Eq. (2) its sensitivity pattern changes depending on the parametrization used (Fig. 2). Indeed, a 25% increase in  $v$  with respect to equivalent background models in  $P[\varepsilon]$  and  $P[v^\perp]$  (Table 1) yields different sensitivities because of the different parameters involved ( $\varepsilon$  or  $v^\perp$ ) in the representation of the medium. Contrarily,  $\delta$  sensitivity pattern is independent of the parametrization used.  $\varepsilon$  and  $v^\perp$  sensitivities are only defined in their respective parametrizations. However, a 25% increase in  $v^\perp$  yields an equivalent  $\varepsilon$  of 0.45 which is greater than the  $\sim 0.2$  limit for weak anisotropy approximation. Thus, instead of establishing the comparison with  $v^\perp$  sensitivity based on a proportional anomaly increment and measuring its effect on traveltimes, we based it on an equal traveltime change, i.e. the same change in traveltime requires a change of 25% in  $\varepsilon$  but only a  $\sim 3.4\%$  change in  $v^\perp$  (Table 1), indicating that data is  $\sim 7$  times more sensitive to  $v^\perp$  than to  $\varepsilon$  changes. For the calculation of sensitivities, we used the same data, that is, the same models and acquisition geometry, as for the accuracy test. For instance, when computing the sensitivity of velocity, we add the 25% spherical anomaly to the velocity background model, while the models for the other parameters remain homogeneous.

~~We define sensitivity as the difference between the first arrival traveltimes with and without the anomaly. Figure 3 shows synthetic and analytic sensitivities for the two selected meridians and the four parameters. Note that the analytic response~~

does not change between meridians, i.e. given the symmetry of the models and of the anisotropic formulation, sensitivity is independent of the azimuth angle.

265

Comparison of Fig. 3a,b with the values in Table 2 and Fig. 2 indicates that the forward calculation of traveltimes is accurate enough with respect to the traveltime residuals expected for the selected anomalies. Sensitivity is at least 5 times and up to two orders of magnitude greater than traveltime accuracy depending on the parameters, with the exception of angles for which sensitivity tends to zero.

270

$v$  sensitivity in  $P[\varepsilon]$  is the highest for all angles, 4.5% to 5% (Fig. 23), and it can be shown that expressed in its relative form, the analytic solution is a constant 5% (see mathematical proof in supplementary material). In its normalized form it follows the same sinusoidal pattern as in  $P[v^\perp]$  (Fig. 2e), and both have equal maxima in the directions parallel to the symmetry axis. In both cases minima are found: in the directions perpendicular to the symmetry axis, but in  $P[v^\perp]$  they reach down to 0 whereas in  $P[\varepsilon]$  the value is  $\sim 0.85$ .  $v^\perp$  sensitivity is virtually identical to  $v$  sensitivity, but as we move away from this direction sensitivity decreases to 0% at the direction of the symmetry axis, as expected from Eq. (1).  $\varepsilon$  sensitivity logically goes to 0% in the directions parallel to the symmetry axis and has its maxima ( $\sim 0.8\%$  or  $\sim 0.15$ ) for the polar angles perpendicular to it.  $v^\perp$  sensitivity would follow the same angular dependence as  $v^\perp$  but with maxima its magnitude of the order of magnitude is as much as 5 times smaller than that of  $v^\perp$  sensitivities  $\sim 0.8\%$  at its maxima. Finally,  $\delta$  sensitivity is, at its maxima, more than one order of magnitude smaller than  $v$  sensitivity, that is around 0.25% or 0.05. These sensitivity results indicate that we can generally expect similar a better recoveries for  $v$  and  $v^\perp$ , better than for  $\varepsilon$  and  $v^\perp$ , and that retrieving  $\delta$  might prove complicated. Keep in mind that these sensitivities for  $v$ ,  $\delta$ , and  $\varepsilon$  are produced by an anomaly that represents a 25% increase with respect to the background value, and that an anomaly in  $v^\perp$  would produce the same sensitivity pattern as  $\varepsilon$  with only a  $\sim 3.4\%$  increment.

285

The differences in synthetic sensitivities between meridians arise from the discretization of the model space in a Cartesian system of coordinates. Such approximation inevitably defines privileged directions for ray tracing, and consequently produces differences in synthetic traveltimes. The mismatch between synthetic and analytic sensitivities (Fig. 32) occurs because the discretization used cannot represent the surface of a perfect sphere. These effects are most notable in the  $v$  sensitivity, in  $P[\varepsilon]$  and to a lesser extent in  $P[v^\perp]$ , precisely because these are the most sensitive parameters, and thus the errors in the representation of a sphere and the existence of privileged directions have a much larger influence on the calculated traveltimes. Figure S1 shows how refining the  $v$  model generates a much more accurate sensitivity pattern, and reduces the relative traveltime misfit error (Fig. S1a), and generates a more accurate sensitivity pattern (Fig. S1b). In a real case study one can always refine the grid spacing of a particular parameter to achieve better accuracy, but here we wish to test the performance of the code in the modeling and recovery of each parameter under the same conditions, i.e. equivalent anomalies and identical model discretization.

295

### 3.2 Accuracy

For the four simulations in the sensitivity analysis, we compared the synthetic traveltimes obtained with our code to the analytic solution to quantify the accuracy of the code's performance. The comparison along the two selected meridians at 0 rad and  $\pi/4$  rad azimuths is displayed in Fig. 3 expressed as relative traveltime error, i.e. the difference between synthetic and analytic traveltimes relative to the latter in percentage. Table 2 contains the mean of the relative traveltime errors and their respective mean deviations for these four tests and along the two selected meridians, as well as the overall values for each of them.

Comparison of Fig. 2a,b with the values in Table 2 and Fig. 3 indicates that the forward calculation of traveltimes is accurate enough with respect to the traveltime residuals expected for the selected anomalies. Sensitivity is at least 5 times and up to two orders of magnitude greater than traveltime accuracy errors depending on the parameters, with the exception of angles for which sensitivity tends to zero. Figure S2 illustrates that the code is able to reproduce nearly identical accuracies using both parametrizations. Furthermore, given that we are using the same synthetic traveltimes that we used for the sensitivity analysis, this also implies that the sensitivity patterns obtained with the alternative parametrization would be virtually equal to those in Fig. 2.

### **3.3. Inversion results**

For the inversion tests, we considered a synthetic medium defined by the anomaly models of all four parameters. Here we refer to these models as target models, and the goal of the inversion is to retrieve the heterogeneity in each of them. These tests are conducted for the two parametrizations of the anisotropic medium described in section 2: P[ $\varepsilon$ ] and P[ $v^\perp$ ]. Note that, in order to perform the inversion tests on equivalent cases for both parametrizations, the heterogeneity in  $v^\perp$  is calculated with Eq. (2) considering the 25% anomalies in  $v$  and  $\varepsilon$ , which yields a  $\sim 29.3\%$  anomaly in  $v^\perp$  (Table 3). If not indicated otherwise, we use background models as initial models. Finally, we study the potential recovery of  $\delta$  because inverting this particular parameter proves notoriously difficult due to its low sensitivity (Fig. 23).

~~In this case, The synthetic data set the acquisition geometry~~ is made of 114 sources each recorded at with 113 receivers for a total 12,882 first arrival traveltimes. For the acquisition geometry, again sources and receivers are located at the surface defined by the sphere inscribed in the cube (Fig. 1). ~~The 114 positions at the surface of this sphere are shared by sources and receivers, and each receiver records all sources, except for the one source located at its same position. Again, sources and receivers are located at the surface defined by the sphere inscribed in the cube, but now all receivers record all sources, except for the one coinciding in location.~~

For both parametrizations, we compared two inversion strategies: simultaneously inverting for all parameters and a two-step sequential inversion. First, in [Figs 4 and 5](#)[Figs 5 and 6](#) we show the best results for the simultaneous inversion strategy. For each parametrization we derived the fourth parameter applying Eq. (2).

Table 4 shows several statistical measures to quantify the quality of these inversion results. As a measure of data fit improvement, we provide the RMS of traveltime residuals for the first and last iterations. As a measure of model recovery or fit, for each parameter, we calculate the mean relative [misfits-difference](#) for the background area between the inverted model and either the target or the initial one, as they are identical in this area, as well as for the anomaly area comparing the inverted model to both the target and the initial models. In the case of a perfect recovery, mean relative [misfit-difference](#) for the background area would be 0%, whereas for the anomaly area, it would be 0% when using the target model as a reference, and 25% (~29.3% for  $v^\perp$ ) when comparing to the initial model.

In an attempt to improve the recovery of  $\delta$ , we repeated these two tests for different values of the smoothing constraints, but it proved impossible. Correlation lengths tested for all four parameters include 0.25 km (twice the grid spacing), 0.5 km, and 1 km. The weights of the smoothing submatrices for each parameter,  $\lambda$  in Eq. (3), were varied between 1 and 100, with intermediate values of 2, 5, 10, 20, 30, and 60. For successful inversions, very similar results for the other parameters were obtained regardless of the final  $\delta$  model. In other words, the low sensitivity of  $\delta$  makes it [extremely hard, if not impossible](#), to recover [it-this parameter](#) from traveltime data, but for this same reason it has little or no influence on the recovery of  $v$  and  $\varepsilon$  or  $v^\perp$ .

The two-step sequential inversion strategy was also tested for both parametrizations  $P[\varepsilon]$  and  $P[v^\perp]$ . For the first step, we tested two options: (a) inverting for  $v$  while fixing  $\delta$  and  $\varepsilon$  or  $v^\perp$  and (b) fixing only  $\delta$ . In the second step we used the inverted models from step 1 as initial models and tested three options: (c) inverting for all three parameters, (d) fixing only  $\delta$ , when following option (a) in step 1, and (e) fixing  $v$  and/or  $\varepsilon$  or  $v^\perp$ , when following option (b) in step 1. [Figure 4 shows a flowchart describing all the sequential inversion combinations tested](#). Again smoothing constraints were varied for similar correlation lengths and submatrix weights as detailed for the simultaneous inversion case.

[As indicated in Fig. 4, in](#) the case of  $P[\varepsilon]$ , the best result ([Fig. 6](#)[Fig. 7](#)) was obtained inverting for  $v$  and  $\varepsilon$  while fixing  $\delta$  in the first step, and fixing only  $\varepsilon$  in the second step, [whereas, As](#) for  $P[v^\perp]$ , the best combination for the two-step inversion ([Fig. 7](#)[Fig. 8](#)) was fixing only  $\delta$  in step 1, and inverting for the three parameters in step 2. Tables 5 and 6 summarize the statistical quantification of data and model fit for each parametrization.

### 3.4 Modelling $\delta$

360 Observing that good results for  $v$  and  $\varepsilon$  or  $v^\perp$  are achieved regardless of the result in  $\delta$ , and knowing that the sensitivity of  $\delta$  is notably smaller than that of the other parameters, we explored a strategy to have an estimate of this parameter. First, as a reference, we considered an unrealistically optimal scenario in which the real  $v$  and  $\varepsilon$  or  $v^\perp$  models are known to us. [Figs 8 and 9](#) [Figs 9 and 10](#) show the resulting  $\delta$  achieved by repeating inversions in [Figs 4 and 5](#) [Figs 5 and 6](#) with  $v$  and  $\varepsilon$  or  $v^\perp$  target models as initial models. Table 7 summarizes the traveltimes residuals RMS and the mean relative [misfits-difference](#) for each parameter in these inversions. Again, these two tests were repeated for ranges of smoothing constraints in all four parameters, as described for the cases in [Figs 4 and 5](#) [Figs 5 and 6](#). Table 7 and [Figs 8 and 9](#) [Figs 9 and 10](#) correspond to the best results obtained, which indicate that the recovery of  $\delta$  is, at best, extremely complicated due to the limited sensitivity of traveltimes data to changes in this parameter.

370 Next, we decided to try neglecting  $\delta$  in Eq. (1), and we repeated a number of inversions, such as the ones displayed in [Figs 4 and 5](#) [Figs 5 and 6](#), following the equation

$$v^a(v, \varepsilon, \theta) = v \cdot (1 + \varepsilon \cdot \sin^4(\theta)) \quad (131)$$

375 The purpose of these tests was checking whether it was possible to invert  $v$  and  $\varepsilon$  or  $v^\perp$  with data generated following Eq. (1) using the approximation in Eq. (131) given that the influence of  $\delta$  on the results for other parameters is rather small, that  $\delta$  cannot be accurately retrieved from traveltimes alone, and that it has the smallest sensitivity. To do so, a homogeneous model of  $\delta=0$  was fixed throughout the inversions. These tests were unsuccessful, with noticeably poorer results than when considering a dependence on  $\delta$  (Table 8). However, they were useful in proving that even if a detailed  $\delta$  model is not necessary to successfully retrieve the other parameters, at least a rough approximation of the  $\delta$  field is needed to recover the other parameters, e.g. the background  $\delta$  model that we used as initial model in inversions displayed in [Figs 4 and 5](#) [Figs 5 and 6](#).

385 Finally, we tested whether it would be possible to obtain at least this rough approximation of  $\delta$  in the medium, valid in the sense that it allows for the successful recovery of the rest of parameters, [using any a priori information available such as compilations of anisotropy measurements \(e.g. Thomsen, 1986; Almqvist and Mainprice, 2017\)](#). Once again we repeated inversions from [Figs 4 and 5](#) [Figs 5 and 6](#) (initial  $\delta = 0.16$ ), now using different homogeneous initial models for  $\delta$  within a range of possible values from 0.1 to 0.24. Table 9 contains the initial and final RMS of traveltimes residuals, as well as  $\delta$  mean values for the inverted model along with the corresponding mean deviations. It is straightforward to note that, for a central subrange of the tested initial  $\delta$  values, final RMS values are an order of magnitude smaller, a few tenths of millisecond compared to the few milliseconds outside this subrange. Specifically, very similar results to those in [Figs 4 and](#)

[5Figs 5 and 6](#) in terms of traveltimes residuals RMS are produced by initial  $\delta$  values between 0.13 and 0.22 for  $P[\varepsilon]$ , and between 0.12 and 0.22 for  $P[v^\perp]$ . The narrowing of the initial  $\delta$  distribution to a smaller subrange of mean  $\delta$  values for the inverted models is indicative of a good general convergence trend.

395

The rough estimate of the  $\delta$  field could be built, for instance, as the average of the mean  $\delta$  values for the inverted models in the central subranges defined by the change in magnitude of the final RMS of traveltimes residuals. One final inversion could be run using a homogeneous initial  $\delta$  model with this average value, with the additional option of fixing it and inverting only for the other two parameters. As mentioned in subsection 2.2, potentially more detailed initial  $\delta$  models could be obtained from the normal move-out correction of near-vertical reflection seismic data.

400

#### 4. Discussion

We have tested two parametrizations of the VTI anisotropic media,  $P[\varepsilon]$  ( $v$ ,  $\delta$ ,  $\varepsilon$ ) and  $P[v^\perp]$  ( $v$ ,  $\delta$ ,  $v^\perp$ ), and two inversion strategies, ~~the~~ simultaneously inverting for all three parameters and a two-step sequential process fixing some of the parameters in each step. We consider three criteria for evaluating and comparing the quality of the inversion results obtained following the four possible combinations of strategies and parametrizations: visual inspection of the results, as well as traveltimes data and model fits. For both inversion strategies and both parametrizations, the recovery of the parameters is positively correlated with their respective sensitivities; the more sensitive parameters are systematically better recovered.

405

For the simultaneous inversion, both parametrizations were able to produce acceptable final results ([Figs 4 and 5](#)[Figs 5 and 6](#)). According to our tests,  $P[\varepsilon]$  provides the best outcome, specifically because data and model fits (Table 4) are better for this option but, more importantly, because of the difference in the quality of the recovery of  $\varepsilon$ . ~~Also, the recovery of  $\delta$ , even though it is far from acceptable, is significantly better in the case of  $P[\varepsilon]$ .~~ However, visual comparison of the recovery of  $v$  as well as the  $v$  model fit in Table 4 indicates that  $P[v^\perp]$  yields a slightly better result regarding the magnitude of the anomaly for this parameter, as is particularly evident at ~~the its center of the anomaly.~~ In  $P[\varepsilon]$ , given the disparity in sensitivities between  $v$  and  $\varepsilon$ , the former might be prone to accounting for data misfit that actually corresponds to the latter, whilst in  $P[v^\perp]$  this disparity is less acute. The boundaries of the anomalies for both velocities are still better retrieved with  $P[\varepsilon]$ . Also, the recovery of  $\delta$ , even though it is far from acceptable, is significantly better in the case of  $P[\varepsilon]$ . This might be explained because the disparity of  $\delta$  sensitivity is greater with respect to the other two parameters in  $P[v^\perp]$  than in  $P[\varepsilon]$ .

410

415

420

In general, sequential inversion is a more complex process that requires more human intervention and fine tuning in each step. In addition, fixing some of the parameters in the first step may result in the inverted parameters artificially accounting for part of the data misfit that is actually related to the fixed ones. This can more easily lead convergence into a local minimum, and it might be impossible to correct this tendency in the second step. For this inversion strategy, it is also  $P[\varepsilon]$

425 that produces the best results (~~Figs 6 and 7~~[Figs 7 and 8](#)). Data fit and the model fit of  $\varepsilon$  are slightly better than for the simultaneous inversion of this parametrization ([Tables 4 and 5](#)), whereas model fits and the visual aspect of both velocities are almost identical to those obtained by simultaneously inverting all three parameters ([Figs 5 and 7](#)~~Tables 4 and 5~~). Visually, it is difficult to decide whether the recovery of  $\varepsilon$  is better or not than for the simultaneous inversion ([Figs 4-5 and 6-7](#)). As for  $\delta$ , recovery is unsuccessful and artifacts appear in the background area of the model but, according to both its model fit and its visual aspect, it is notably better than for the simultaneous inversion. As in the case of simultaneous inversion, the only advantage of using  $P[v^\perp]$  instead of  $P[\varepsilon]$  is that it yields a better recovery of the anomaly magnitude of  $v$  ([Tables 4 and 6](#)). The results for both velocities are virtually identical to those obtained by simultaneous inversion of this parametrization ([Figs 5-6 and 7-8](#)).  $\delta$  and  $\varepsilon$  are not properly retrieved, but the results are significantly better than for the simultaneous inversion of this same parametrization.

435 ~~For both inversion strategies, results are always better for  $P[\varepsilon]$  than for  $P[v^\perp]$ . This behavior can be attributed to crosstalk between  $v$  and  $v^\perp$  as indicated by their sensitivities (Fig. 3). Indeed, for a moderately wide range of angles, roughly within  $\pi/2 \pm \pi/4$  and  $3\pi/2 \pm \pi/4$ , both sensitivity patterns are of the same order of magnitude which can result in a substantial trade-off for data within those ranges. Contrarily, the sensitivity patterns of the parameters used in  $P[\varepsilon]$  do not interfere as much with each other given their differences in magnitude.~~

440  $\delta$  has been shown to be by far the most complicated parameter to retrieve because of the low sensitivity of traveltimes to its variation (Fig. [3-2](#)). Even when excellent  $v$  and  $\varepsilon$  or  $v^\perp$  models are available, i.e. the target models for these parameters in our synthetic tests, the recovery of  $\delta$  is limited at best (~~Figs 8 and 9~~[Figs 9 and 10](#) and Table 7). However, and for the same reason, poor recoveries of  $\delta$  do not affect the recovery of the other two parameters, meaning that a detailed  $\delta$  model is not necessary to satisfactorily retrieve  $v$  and  $\varepsilon$  or  $v^\perp$  ([Figs 4-5-9-10](#)). Still, our inversion tests also proved that neglecting  $\delta$  in Eqs (1) and (3) is not an option, the accuracy in the recovery of the other parameters resulting severely affected (Table 8). Thus, even if a detailed inversion of  $\delta$  is, at the very least, hard to achieve, and it is not needed for a successful result in the other parameters, some sort of simple, even homogeneous, initial  $\delta$  model with a value or values about the average  $\delta$  in the medium, is necessary for a good recovery of the other parameters.

450 We showed that given some a priori information on the range of possible  $\delta$  values in the medium, it should be possible to create the necessary initial  $\delta$  model. In order to illustrate this, we chose a range of  $\delta$  values for the initial model and rerun the inversions in ~~Figs 4 and 5~~[Figs 5 and 6](#). The results indicate that for any homogeneous initial  $\delta$  model in a certain subrange close to the actual average  $\delta$  value of the medium, the results for  $v$  and  $\varepsilon$  or  $v^\perp$  are satisfactory and virtually identical to those of the original inversions (Table 9). This subrange is easily defined by looking at the final RMS of traveltimes residuals, which experiences a notorious change of ~~an one~~[an one](#) order of magnitude. Any model within this subrange works similarly well as initial  $\delta$  model. Alternatively, a possibly more robust selection of the constant value for a homogeneous initial  $\delta$  model might

be the mean (or also the median or the mode) of the mean  $\delta$  values for the inverted models in this subrange. It is worth noting that, whereas for the purpose of this work we used the same discretization for all parameters, in a real case study it would probably be recommendable to use a coarser discretization for  $\delta$  than for the other parameters, and in general a finer discretization for the more sensitive parameters (Fig. S1). [Indeed, a heterogeneity of a given spatial scale and relative variation will produce a greater effect on data for a parameter of greater sensitivity. Thus, for a parameter of greater sensitivity, it will be easier for the code to identify smaller heterogeneities both in scale and variation, which will require a finer grid.](#)

## 465 5 Conclusions

We have successfully implemented and tested a new anisotropic traveltimes tomography code. For this implementation we had to modify both the forward problem and the inversion algorithms of the TOMO3D code (Meléndez et al., 2015). The forward problem was adapted to compute the velocities observed by rays considering Eq. (1) for the weak [VTI](#) anisotropy formulation in Thomsen (1986). The inversion solver was extended to include the  $\delta$ ,  $\varepsilon$  and  $\nu^\perp$  kernels in the linearized forward problem matrix equation, as well as smoothing and damping matrices for these parameters defined following the same scheme as for velocity in the isotropic code (Eqs 3-102).

Regarding the synthetic tests, ~~after checking the proper performance of the code by comparing with analytic solutions (Fig. 2),~~ we [first](#) determined the sensitivity of traveltimes data to changes in each of the parameters defining anisotropy in the medium ~~of interest~~ ( $\nu$ ,  $\delta$ ,  $\varepsilon$ ,  $\nu^\perp$ ) (Fig. 32), [and we checked the proper performance of the code by comparing the synthetic traveltimes produced with their respective analytic solutions \(Fig. 3\).](#) Next, we performed [canonical](#) inversion tests to compare two possible media parametrization,  $P[\varepsilon]$  and  $P[\nu^\perp]$ , and two possible inversion strategies, simultaneous and sequential. According to our tests, both parametrizations have their strengths:  $P[\varepsilon]$  produces the best overall result in the sense that all parameters are acceptably recovered, [with the exception of  \$\delta\$ ,](#) and trade-off between parameters is lower, but  $P[\nu^\perp]$  yields the best result for the magnitude of the anomaly in  $\nu$ . Regarding the inversion strategy, simultaneous inversion is more straightforward and involves less human intervention, and given that both strategies yield similar results, it would be our first choice. Sequential inversion is always a more complex process that can be shown to work in a synthetic case because the target models are available, but in field data applications the complexity would most likely be unmanageable. [These tests were conducted under ideal conditions, and thus the conclusions provided by their results are an upper-limit but generalizable estimation of the strengths and weaknesses in the code's performance.](#)

An acceptable recovery of  $\delta$  turned out to be impossible due to the small sensitivity of traveltimes to this parameter, but we verified that it cannot simply be neglected in the equations. Whereas the recovery of the other parameters is not significantly affected by that of  $\delta$ , a rough estimate of the average  $\delta$  value in the medium is necessary and sufficient to generate a

490 | homogeneous initial model that allows for satisfactory inversion results in the [se](#) other parameters. We also proved that it is possible to obtain it, provided that some a priori knowledge on  $\delta$  values in the medium is available to define a range of plausible values, such as field or laboratory measurements.

### **Code availability**

495 | The anisotropic version of TOMO3D will be made available [only](#) for academic purposes ~~only~~ on our group website. Currently a copy of the code can be obtained by sending an e-mail to the corresponding author.

### **Author contribution**

500 | The formulation of the overarching research goals of this work is a product of discussion among the four co-authors. AM was in charge of software development and data curation, analysis, visualization, and validation. AM also prepared the manuscript. The methodology for the synthetic tests was designed by AM, CEJ, and VS. CR was responsible for [funding the acquisition of financial support](#).

### **Competing interests**

The authors declare that they have no conflict of interest.

### **Acknowledgements**

505 | AM and CLJ are funded by Respol through the SOUND collaboration project with CSIC, and the work in this paper was conducted at the Grup de Recerca de la Generalitat de Catalunya 2009SGR146: Barcelona Center for Subsurface Imaging (B-CSI). We thank all our fellows at the B-CSI for their contribution to this work. We also wish to thank Dr. Marko Riedel for his constructive comments at the 18th edition of the biennial International Symposium on Deep Seismic Profiling of the Continents and their Margins (SEISMIX 2018). [The comments of two anonymous referees have also contributed significantly to the improvement of our work.](#)

### 510 **References**

Alkhalifah, T.: Acoustic approximations for processing in transversely isotropic media, *Geophysics*, 63, 623-631, <https://doi.org/10.1190/1.1444361>, 1998.

Alkhalifah, T.: Travel Time computation with the linearized eikonal equation for anisotropic media, *Geophys. Prospect.*, 50, 373-382, <https://doi.org/10.1046/j.1365-2478.2002.00322.x>, 2002.

- 515 Alkhalifah, T. and Plessix, R. É.: A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study, *Geophysics*, 79, R91-R101, <https://doi.org/10.1190/geo2013-0366.1>, 2014.
- Almqvist B. S. G. and Mainprice, D.: Seismic properties and anisotropy of the continental crust: Predictions based on mineral texture and rock microstructure, *Rev. Geophys.*, 55, 367-433, <https://doi.org/10.1002/2016RG000552>, 2017.
- Babuska, V. and Cara, M.: *Seismic Anisotropy in the Earth*, 1st edition, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.
- 520 Backus G. E.: Long-wave elastic anisotropy produced by horizontal layering. *J. Geophys. Res.*, 67, 4427-4440, <https://doi.org/10.1029/JZ067i011p04427>, 1962.
- Bakulin, A., Woodward, M., Osypov, K., Nichols, D., and Zdraveva, O.: Can we distinguish TTI and VTI media?, in: SEG Technical Program Expanded Abstracts 2009 of the 79th SEG Annual Meeting, Houston, USA, 25-30 October 2009, 226-
- 525 230, <https://doi.org/10.1190/1.3255313>, 2009.
- Bamford, D. and Crampin, S.: Seismic anisotropy – the state of the art, *Geophys. J. R. astr. Soc.*, 49, 1-8, <https://doi.org/10.1111/j.1365-246X.1977.tb03697.x>, 1977.
- Bezada, M. J., Faccenda, M., Toomey, D. R., and Humphreys, E.: Why ignoring anisotropy when imaging subduction zones could be a bad idea, in: AGU Fall Meeting Abstracts of the 47th AGU Fall Meeting, San Francisco, USA, 15-19 December
- 530 2014, S41D-04, 2014.
- Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., Spindler, D. A., and Wyld, H. W.: Tomographic determination of velocity and depth in laterally varying media, *Geophysics*, 50, 903-923, <https://doi.org/10.1190/1.1441970>, 1985.
- Crampin, S.: A review of wave motion in anisotropic and cracked elastic-media, *Wave Motion*, 3, 343-391, [https://doi.org/10.1016/0165-2125\(81\)90026-3](https://doi.org/10.1016/0165-2125(81)90026-3), 1981.
- 535 Cervený, V.: Seismic rays and ray intensities in inhomogeneous anisotropic media, *Geophys. J. R. Astr. Soc.*, 29, 1-13, <https://doi.org/10.1111/j.1365-246X.1972.tb06147.x>, 1972.
- Cervený, V.: Direct and inverse kinematic problems for inhomogeneous anisotropic media – linearized approach, *Contr. Geophys. Inst. Slov. Aca. Sci.*, 13, 127-133, 1982.
- 540 Cervený, V. and Jech, J.: Linearized solutions of kinematic problems of seismic body waves in inhomogeneous slightly anisotropic media, *J. Geophys.*, 51, 96-104, 1982.
- Chapman C. H. and Pratt, R. G.: Travelttime tomography in anisotropic media – I. Theory, *Geophys. J. Int.*, 109, 1-19, <https://doi.org/10.1111/j.1365-246X.1992.tb00075.x>, 1992.
- Dunn, R. A., Toomey, D. R., Detrick, R. S., and Wilcock, W. S. D.: Continuous Mantle Melt Supply Beneath an
- 545 Overlapping Spreading Center on the East Pacific Rise, *Science*, 291, 1955-1958, <https://doi.org/10.1126/science.1057683>, 2001.

- Dunn, R. A., Lekic, V., Detrick, R. S., and Toomey, D. R.: Three-dimensional seismic structure of the Mid-Atlantic Ridge (35°N): Evidence for focused melt supply and lower crustal dike injection, *J. Geophys. Res.*, 110, <https://doi.org/10.1029/2004JB003473>, 2005.
- 550 Duveneck, E. and Bakker, P. M.: Stable P-wave modeling for reverse-time migration in tilted TI media, *Geophysics*, 76, S65–S75, <https://doi.org/10.1190/1.3533964>, 2011.
- Eberhart-Phillips, D. and Henderson, M.: Including anisotropy in 3-D velocity inversion and application to Marlborough, New Zealand, *Geophys. J. Int.*, 156, 237–254, <https://doi.org/10.1111/j.1365-246X.2003.02044.x>, 2004.
- Eilon, Z., Abers, G. A., and Gaherty, J. B.: A joint inversion for shear velocity and anisotropy: the Woodlark Rift, Papua  
 555 New Guinea, *Geophys. J. Int.*, 206(1), 807–824, <https://doi.org/10.1093/gji/ggw177>, 2016.
- Fowler, P.J., Du, X., and Fletcher, R.P.: Coupled equations for reverse-time migration in transversely isotropic media, *Geophysics*, 75, S11–S22, <https://doi.org/10.1190/1.3294572>, 2010.
- Giroux, B. and Gloaguen, E.: Geostatistical traveltimes tomography in elliptically anisotropic media, *Geophys. Prospect.*, 60(6), 1133–1149, <https://doi.org/10.1111/j.1365-2478.2011.01047.x>, 2012.
- 560 Gholami, Y., Brossier, R., Operto, S., Ribodetti, A., and Virieux, J.: Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: Sensitivity and trade-off analysis, *Geophysics*, 78, R81–R105, <https://doi.org/10.1190/geo2012-0204.1>, 2013.
- Gholami, Y., Brossier, R., Operto, S., Prioux, V., Ribodetti, A., and Virieux, J.: Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 2: Synthetic and real data case studies from Valhall.  
 565 *Geophysics*, 78, R107–R124, <https://doi.org/10.1190/geo2012-0203.1>, 2013.
- Helbig, K.: Elliptical anisotropy—Its significance and meaning, *Geophysics*, 48, 825–832, <https://doi.org/10.1190/1.1441514>, 1983.
- Holtzman, B. K., Kohlstedt, D. L., Zimmerman, M. E., Heidelbach, F., Hiraga, T., and Hustoft, J.: Melt segregation and strain partitioning: implications for seismic anisotropy and mantle flow, *Science*, 301, 1227–1230,  
 570 <https://doi.org/10.1126/science.1087132>, 2003.
- Ismaïl, W. B. and Mainprice, D.: An olivine fabric database: an overview of upper mantle fabrics and seismic anisotropy, *Tectonophysics*, 296, 145–157, [https://doi.org/10.1016/S0040-1951\(98\)00141-3](https://doi.org/10.1016/S0040-1951(98)00141-3), 1998.
- Jech, J. and Psencik, I.: First-order perturbation method for anisotropic media, *Geophys. J. Int.*, 99, 369–376, <https://doi.org/10.1111/j.1365-246X.1989.tb01694.x>, 1989.
- 575 Johnston, J. E. and Christensen, N. I.: Seismic anisotropy of shales, *J. Geophys. Res.*, 100, 5991–6003, <https://doi.org/10.1029/95JB00031>, 1995.
- Kendall, J. M., Stuart, G. W., Ebinger, C. J., Bastow, I. D., and Keir, D.: Magma-assisted rifting in Ethiopia, *Nature*, 433, 146–148, <https://doi.org/10.1038/nature03161>, 2005.

- Kendall, J. M., Pilidou, S., Keir, D., Bastow, I. D., Stuart, G. W., and Ayele, A.: Mantle upwellings, melt migration and the  
580 rifting of Africa: insights from seismic anisotropy, *Geol. Soc. Lond. Spec. Publ.*, 259(1), 55–72,  
<https://doi.org/10.1144/GSL.SP.2006.259.01.06>, 2006.
- Kraut, E. A.: Advances in the Theory of Anisotropic Elastic Wave Propagation, *Rev. Geophys.*, 1(3), 401-448,  
<https://doi.org/10.1029/RG001i003p00401>, 1963.
- Long, M. D. and Silver, P. G.: The Subduction Zone Flow Field from Seismic Anisotropy: A Global View, *Science*, 319,  
585 315-318, <https://doi.org/10.1126/science.1150809>, 2008.
- Long, M. D. and Becker, T. W.: Mantle dynamics and seismic anisotropy, *Earth Planet. Sci. Lett.*, 297, 341-354,  
<https://doi.org/10.1016/j.epsl.2010.06.036>, 2010.
- Maultzsch, S., Chapman, M., Liu, E., and Li, X. Y.: Modelling frequency-dependent seismic anisotropy in fluid-saturated  
rock with aligned fractures: implication of fracture size estimation from anisotropic measurements, *Geophys. Prospect.*,  
590 51(5), 381-392, <https://doi.org/10.1046/j.1365-2478.2003.00386.x>, 2003.
- McCann, C., Assefa, S., Sothcott, J., McCann, D. M., and Jackson, P. D.: In-situ borehole measurements of compressional  
and shear wave attenuation in Oxford clay, *Sci. Drilling*, 1, 11-20, 1989.
- Meléndez, A.: Development of a New Parallel Code for 3-D Joint Refraction and Reflection Travel-Time Tomography of  
Wide-Angle Seismic Data – Synthetic and Real Data Applications to the Study of Subduction Zones, PhD thesis, Universitat  
595 de Barcelona, Barcelona, Spain, pp. 171, <http://hdl.handle.net/2445/65200>, 2014.  
[Meléndez, A., Sallarès, V., Ranero, C. R. and Kormann, J.: Origin of water layer multiple phases with anomalously high  
amplitude in near-seafloor wide-angle seismic recordings, \*Geophys. J. Int.\*, 196, 243-252, <https://doi.org/10.1093/gji/ggt391>,  
2014.](https://doi.org/10.1093/gji/ggt391)
- Meléndez, A., Korenaga, J., Sallarès, V., Miniussi, A., and Ranero, C. R.: TOMO3D: 3-D joint refraction and reflection  
600 traveltimes tomography parallel code for active-source seismic data – synthetic test, *Geophys. J. Int.*, 203, 158-174,  
<https://doi.org/10.1093/gji/ggv292>, 2015.
- Montagner, J. P., Marty, B., and Stutzmann, E.: Mantle upwellings and convective instabilities revealed by seismic  
tomography and helium isotope geochemistry beneath eastern Africa, *Geophys. Res. Lett.*, 34, L21303,  
<https://doi.org/10.1029/2007GL031098>, 2007.
- 605 [Moser, T. J.: Shortest path calculation of seismic rays, \*Geophysics\*, 56, 59-67, <https://doi.org/10.1190/1.1442958>, 1991.](https://doi.org/10.1190/1.1442958)  
[Moser, T. J., Nolet, G., Snieder, R.: Ray bending revisited, \*B. Seismol. Soc. Am.\*, 82, 259-288, 1992.](https://doi.org/10.1190/1.1442958)
- Nicolas, A. and Christensen, N. I.: Formation of anisotropy in upper mantle peridotites-A review. Composition, structure and  
dynamics of the lithosphere-asthenosphere system, 16, 111-123, <https://doi.org/10.1029/GD016p0111>, 1987.
- Nye, J. F.: Physical Properties of Crystals: their representation by tensors and matrices, 1st edition, Oxford University Press,  
610 Oxford, England, 1957.
- Pratt, R. G. and Chapman C. H.: Traveltime tomography in anisotropic media – II. Application, *Geophys. J. Int.*, 109, 20-37,  
<https://doi.org/10.1111/j.1365-246X.1992.tb00076.x>, 1992.

- Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O. I., and Kommedal, J. H.: On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophysical Journal International*, 187, 1495-1515, 615 <https://doi.org/10.1111/j.1365-246X.2011.05209.x>, 2011.
- Rüger, A. and Alkhalifah, T.: Efficient two-dimensional anisotropic ray tracing, in: *Seismic Anisotropy*, edited by: Erling, F., Rune, M. H., and Jaswant, S. R., Soc. of Expl. Geophys., 556-600, <https://doi.org/10.1190/1.9781560802693.ch18>, 1996.
- Sallarès, V., Meléndez, A., Prada, M., Ranero, C. R., McIntosh, K., and Grevenmeyer, I.: Overriding plate structure of the Nicaragua convergent margin: Relationship to the seismogenic zone of the 1992 tsunami earthquake, *Geochem. Geophys. Geosyst.*, 14, 3436–3461, <https://doi.org/10.1002/ggge.20214>, 2013. 620
- Sayers, C. M.: Seismic anisotropy of shales. *Geophys. Prospect.*, 53, 667-676, <https://doi.org/10.1111/j.1365-2478.2005.00495.x>, 2005.
- Schulte-Pelkum, V. and Mahan, K. H.: A method for mapping crustal deformation and anisotropy with receiver functions and first results from USArray. *Earth Planet. Sci. Lett.*, 402, 221-233, <https://doi.org/10.1016/j.epsl.2014.01.050>, 2014.
- 625 Song, L., Zhang, S., Liu, H., Chun, S., and Song, Z.: A formalism for acoustical traveltimes tomography in heterogeneous anisotropic media, in: *Review of Progress in Quantitative Nondestructive Evaluation*, edited by: Thompson, D. O. and Chimenti, D. E., Springer, Boston, USA, 1529-1536, [https://doi.org/10.1007/978-1-4615-5339-7\\_198](https://doi.org/10.1007/978-1-4615-5339-7_198), 1998.
- Stewart, R. R.: An algebraic reconstruction technique for weakly anisotropic velocity, *Geophysics*, 53, 1613-1615, <https://doi.org/10.1190/1.1442445>, 1988.
- 630 Thomsen, L.: Weak elastic anisotropy, *Geophysics*, 51, 1954-1966, <https://doi.org/10.1190/1.1442051>, 1986.
- Thomsen, L. and Anderson, D. L.: Weak elastic anisotropy in global seismology. *Geol. Soc. Spec. Pap.*, 514, 39-50, [https://doi.org/10.1130/2015.2514\(04\)](https://doi.org/10.1130/2015.2514(04)), 2015.
- Tsvankin, I.: P- wave signatures and notation for transversely isotropic media: An overview, *Geophysics*, 61(2), 467-483, <https://doi.org/10.1190/1.1443974>, 1996.
- 635 Vauchez, A., Tommasi, A., Barruol, G., and Maumus, J.: Upper mantle deformation and seismic anisotropy in continental rifts, *Phys. Chem. Earth A: Solid Earth Geod.*, 25(2), 111–117, [https://doi.org/10.1016/S1464-1895\(00\)00019-3](https://doi.org/10.1016/S1464-1895(00)00019-3), 2000.
- Vinnik, L. P., Makeyeva, L. I., Milev, A., and Usenko, A. Y.: Global patterns of azimuthal anisotropy and deformations in the continental mantle. *Geophys. J. Int.*, 111(3), 433-447, <https://doi.org/10.1111/j.1365-246X.1992.tb02102.x>, 1992.
- Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C., Conroy, 640 G., and Bertrand, A.: Anisotropic 3D full-waveform inversion. *Geophysics*, 78(2), R59-R80, <https://doi.org/10.1190/geo2012-0338.1>, 2013.
- Yousef B. M. and Angus, D. A.: When do fractured media become seismically anisotropic? Some implications on quantifying fracture properties, *Earth Planet. Sci. Lett.*, 444, 150-159, <https://doi.org/10.1016/j.epsl.2016.03.040>, 2016.
- Zhou, B. and Greenhalgh, S.: 'Shortest path' ray tracing for most general 2D/3D anisotropic media, *J. Geophys. Eng.*, 2, 54- 645 63, <https://doi.org/10.1088/1742-2132/2/1/008>, 2005.

**Tables**

<u>Sensitivity test for</u>	<u>Model area</u>	<u>P[<math>\epsilon</math>]</u>			<u>P[<math>v^\perp</math>]</u>		
		<u><math>v</math></u>	<u><math>\delta</math></u>	<u><math>\epsilon</math></u>	<u><math>v</math></u>	<u><math>\delta</math></u>	<u><math>v^\perp</math></u>
<u><math>v</math> in P[<math>\epsilon</math>]</u>	<u>Anomaly</u>	<u>2.5</u>	<u>0.16</u>	<u>0.16</u>	<u>2.5</u>	<u>0.16</u>	<u>2.9</u>
	<u>Background</u>	<u>2</u>	<u>0.16</u>	<u>0.16</u>	<u>2</u>	<u>0.16</u>	<u>2.32</u>
<u><math>v</math> in P[<math>v^\perp</math>]</u>	<u>Anomaly</u>	<u>2.5</u>	<u>0.16</u>	<u>-0.072</u>	<u>2.5</u>	<u>0.16</u>	<u>2.32</u>
	<u>Background</u>	<u>2</u>	<u>0.16</u>	<u>0.16</u>	<u>2</u>	<u>0.16</u>	<u>2.32</u>
<u><math>\delta</math></u>	<u>Anomaly</u>	<u>2</u>	<u>0.2</u>	<u>0.16</u>	<u>2</u>	<u>0.2</u>	<u>2.32</u>
	<u>Background</u>	<u>2</u>	<u>0.16</u>	<u>0.16</u>	<u>2</u>	<u>0.16</u>	<u>2.32</u>
<u><math>\epsilon</math> (and <math>v^\perp</math>)</u>	<u>Anomaly</u>	<u>2</u>	<u>0.16</u>	<u>0.2</u>	<u>2</u>	<u>0.16</u>	<u>2.4</u>
	<u>Background</u>	<u>2</u>	<u>0.16</u>	<u>0.16</u>	<u>2</u>	<u>0.16</u>	<u>2.32</u>

650 **Table 12:** Values for anomaly and background areas of the models used in each of the four sensitivity tests (white cells), and their equivalences in the alternative parametrization (gray cells) (Fig. 2). Accuracy tests in Fig. 3 are conducted for these same four cases (white cells). A comparison with the accuracies achieved using the alternative parametrizations can be seen in Fig. S2.

<u>Anomaly in</u>	<u><math>v</math> in P[<math>\epsilon</math>]</u>		<u><math>v</math> in P[<math>v^\perp</math>]</u>		<u><math>\delta</math></u>		<u><math>\epsilon</math></u>	
<u>Azimuth</u>	<u>0 rad</u>	<u><math>\pi/4</math> rad</u>	<u>0 rad</u>	<u><math>\pi/4</math> rad</u>	<u>0 rad</u>	<u><math>\pi/4</math> rad</u>	<u>0 rad</u>	<u><math>\pi/4</math> rad</u>
<u>Mean relative error <math>\pm</math> mean deviation (%)</u>	<u><math>0.8 \pm 0.1</math></u>	<u><math>0.6 \pm 0.1</math></u>	<u><math>0.6 \pm 0.2</math></u>	<u><math>0.5 \pm 0.2</math></u>	<u><math>0.01 \pm 0.01</math></u>	<u><math>0.01 \pm 0.01</math></u>	<u><math>0.04 \pm 0.05</math></u>	<u><math>0.03 \pm 0.04</math></u>
	<u>Overall values</u>							
	<u><math>0.7 \pm 0.1</math></u>		<u><math>0.5 \pm 0.2</math></u>		<u><math>0.017 \pm 0.009</math></u>		<u><math>0.04 \pm 0.03</math></u>	

655 **Table 2:** Mean relative traveltimes misfit errors in percentage and their mean deviations for the two selected meridians in Fig. 32, and for the entire set of 482 source – receiver pairs. Compared to Fig. 23a,b, these average traveltimes misfit error values indicate that the code is sufficiently accurate to model the traveltimes residuals arising from the inclusion of the selected anomalies.

	$v$ (km/s)	$\delta$	$\varepsilon$	$v^\perp$ (km/s)
Background value	2	0.16	0.16	2.32
Anomaly value	2.5	0.2	0.2	3

**Table 3:** For inversion tests, background and anomaly values of all four parameters for the initial/background model and for the anomaly in the target model. The model is a cube of edge 5 km. The anomaly is a discretized sphere of 1 km in diameter at the center of the cube.

	Residuals RMS (ms)	Mean relative <u>misfits-differences</u> (%)											
		$v$			$\delta$			$\varepsilon$			$v^\perp$		
		BG	AI	AT	BG	AI	AT	BG	AI	AT	BG	AI	AT
P[ $\varepsilon$ ]	30 – 0.4	0.5	21.0	3.3	4.8	22.3	15.2	1.6	11.1	11.2	0.5	22.8	5.0
P[ $v^\perp$ ]	30 – 0.5	0.8	25.9	1.9	5.0	20.8	29.1	5.8	26.7	41.0	0.6	21.3	6.2

660 **Table 4:** Quantification of the quality of recovery in terms of data and model fit for the simultaneous inversion of both  
665 parametrizations (Figs 4 and 5/Figs 5 and 6). Initial and final RMS values for traveltime residuals to quantify data fit. For each  
parameter, as a measure of model fit, we computed the mean relative misfits-differences between the background areas of the  
inverted and target models (BG), the anomaly areas of the inverted and initial models (AI), and the anomaly areas of the inverted  
and target models (AT). Since the initial model is equal to the target model in the background area, it is not necessary to calculate  
the misfit-difference between inverted and initial models for this area. The ideal misfit-difference value for BG is 0%. AI and AT  
indicate the resemblance between true and recovered anomalies, and their ideal values are 0% and 25% (~29.3% in the case of  $v^\perp$ )  
respectively. Recovery is consistent with sensitivity (Fig. 23):  $v$  and  $v^\perp$  are well retrieved,  $\varepsilon$  is only partially recovered with P[ $\varepsilon$ ],  
whilst inversion of  $\delta$  is unsuccessful in both cases.

P[ $\varepsilon$ ]	Residuals RMS (ms)	Mean relative <u>misfits-differences</u> (%)											
		$v$			$\delta$			$\varepsilon$			$v^\perp$		
		BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>
Step 1	30 – 0.5	0.5	21.4	2.9	-	-	-	1.3	8.2	13.5	0.5	22.8	5.1
Step 2	0.5 – 0.3	0.5	21.4	3.0	0.9	5.3	19.4	-	-	-	0.5	22.8	5.1

670 **Table 5:** Same as Table 4 but for the two-step sequential inversion of P[ $\varepsilon$ ] (Fig. 6/Fig. 7). Recoveries in terms of model fit are  
virtually identical to those for the simultaneous inversion of this parametrization, with  $\varepsilon$  recovery being just slightly better. Final  
data fit is also better than the one achieved by simultaneous inversion of P[ $\varepsilon$ ].

	Residuals RMS (ms)	Mean relative <u>misfits-differences</u> (%)											
		$\nu$			$\delta$			$\varepsilon$			$\nu^\perp$		
		BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>
Step 1	30 – 0.7	0.7	26.7	2.7	-	-	-	5.4	33.0	45.7	0.5	21.0	6.4
Step 2	0.7 – 0.5	0.7	26.2	2.5	1.8	17.4	14.9	5.7	29.8	41.8	0.5	21.5	6.1

675 **Table 6:** Same as Table 5 but for the two-step sequential inversion of  $P[\nu^\perp]$  (Fig-7Fig. 8). Model fits are similar to those achieved by the simultaneous inversion of this parametrization:  $\nu$  and  $\nu^\perp$  are well retrieved, while recovery of  $\delta$  and  $\varepsilon$  is unsuccessful. Final data fit is identical to the one obtained by simultaneously inverting for  $P[\nu^\perp]$ .

	Residuals RMS (ms)	Mean relative <u>misfits-differences</u> (%)											
		$\nu$			$\delta$			$\varepsilon$			$\nu^\perp$		
		BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>
$P[\varepsilon]$	19 – 0.1	0.1	25.2	0.3	0.8	15.4	7.7	0.3	24.4	0.5	0.1	29.4	0.3
$P[\nu^\perp]$	21 – 0.4	0.2	24.2	1.1	1.4	30.6	9.9	1.2	31.9	7.4	0.1	29.7	0.5

680 **Table 7:** Same as Table 4 but for the simultaneous inversion of  $P[\varepsilon]$  and  $P[\nu^\perp]$  using  $\nu$  and  $\varepsilon$  or  $\nu^\perp$  target models as initial models (Figs 8 and 9Figs 9 and 10). Model fits for  $\nu$  and  $\varepsilon$  or  $\nu^\perp$  are close to perfect as expected, but even so the recovery of  $\delta$  is partial at most. Data misfit is smaller than for the original inversions in Figs 4 and 5Figs 5 and 6.  $P[\varepsilon]$  yields a better result according to all indicators, and as for all previous tests, the recovery of  $\nu^\perp$  from  $P[\varepsilon]$  is notably better than that of  $\varepsilon$  from  $P[\nu^\perp]$ .

	Residuals RMS (ms)	Mean relative <u>misfits-differences</u> (%)											
		$\nu$			$\delta$			$\varepsilon$			$\nu^\perp$		
		BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>	BG	<u>AI</u>	<u>AT</u>
$P[\varepsilon]$	61 – 0.8	3.3	22.0	3.1	-	-	-	14.6	28.6	5.2	2.6	26.8	3.2
$P[\nu^\perp]$	61 – 0.8	3.5	21.1	3.3	-	-	-	24.7	40.7	15.2	2.0	27.8	2.4

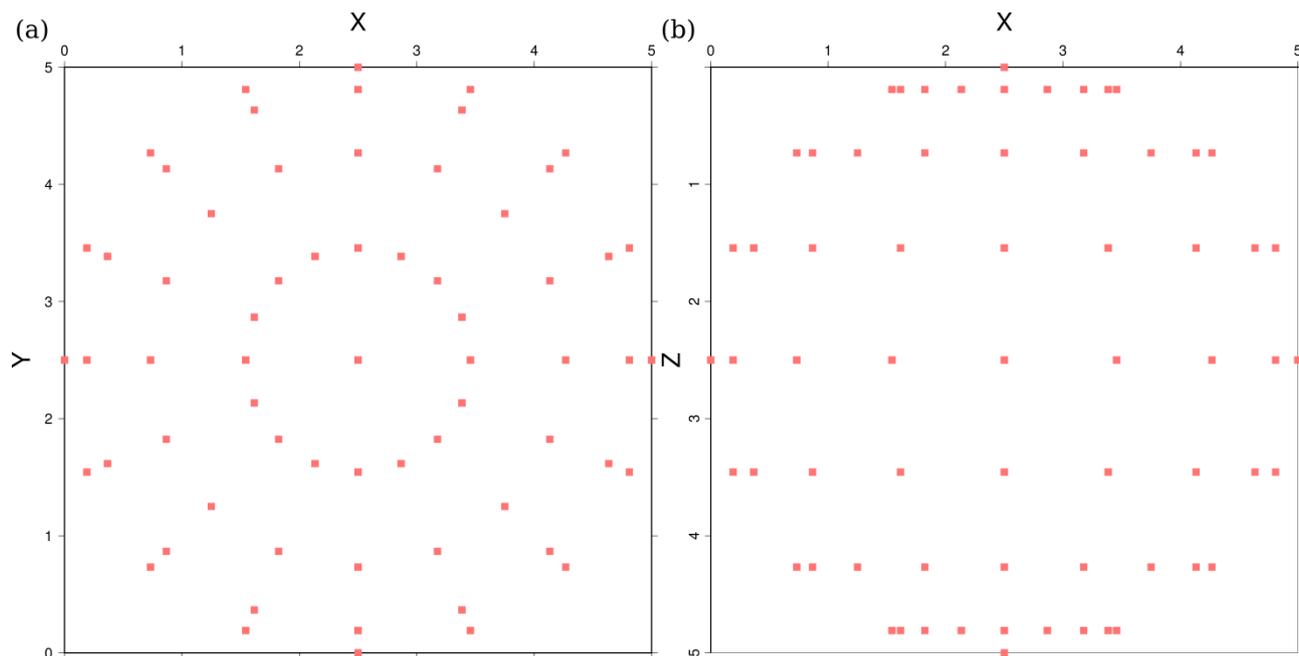
**Table 8:** Same as Table 4 but for the simultaneous inversion of  $P[\varepsilon]$  and  $P[\nu^\perp]$  following Eq. (131), i.e. neglecting  $\delta$  in Eqs (1) and (3). Model fits-differences for  $\nu$ ,  $\varepsilon$ , and  $\nu^\perp$ , as well as data misfit are all significantly worse than for any of the previous tests.

$\delta$ value for initial model	P[ $\varepsilon$ ]		P[ $\nu^{\perp}$ ]	
	Initial and final traveltime residuals RMS (ms)	Mean $\pm$ mean deviation of inverted $\delta$ model	Initial and final traveltime residuals RMS (ms)	Mean $\pm$ mean deviation of inverted $\delta$ model
0.1	39 – 11	0.13 $\pm$ 0.05	39 – 5	0.13 $\pm$ 0.04
0.11	37 – 8	0.13 $\pm$ 0.04	37 – 4	0.14 $\pm$ 0.03
0.12	35 – 4	0.15 $\pm$ 0.03	35 – 0.5	0.14 $\pm$ 0.02
0.13	34 – 0.5	0.15 $\pm$ 0.01	34 – 0.5	0.14 $\pm$ 0.02
0.14	32 – 0.4	0.15 $\pm$ 0.01	32 – 0.5	0.15 $\pm$ 0.01
0.15	31 – 0.4	0.16 $\pm$ 0.01	31 – 0.5	0.16 $\pm$ 0.01
0.16	30 – 0.4	0.160 $\pm$ 0.008	30 – 0.5	0.160 $\pm$ 0.008
0.17	28 – 0.4	0.16 $\pm$ 0.01	29 – 0.6	0.17 $\pm$ 0.01
0.18	27 – 0.4	0.17 $\pm$ 0.01	28 – 0.5	0.17 $\pm$ 0.01
0.19	27 – 0.5	0.17 $\pm$ 0.01	27 – 0.5	0.18 $\pm$ 0.01
0.2	27 – 0.5	0.18 $\pm$ 0.02	27 – 0.5	0.18 $\pm$ 0.02
0.21	27 – 0.5	0.18 $\pm$ 0.02	27 – 0.5	0.19 $\pm$ 0.02
0.22	27 – 0.5	0.19 $\pm$ 0.02	27 – 0.5	0.19 $\pm$ 0.02
0.23	27 – 5	0.20 $\pm$ 0.04	27 – 2	0.20 $\pm$ 0.03
0.24	28 – 11	0.22 $\pm$ 0.06	28 – 5	0.20 $\pm$ 0.04

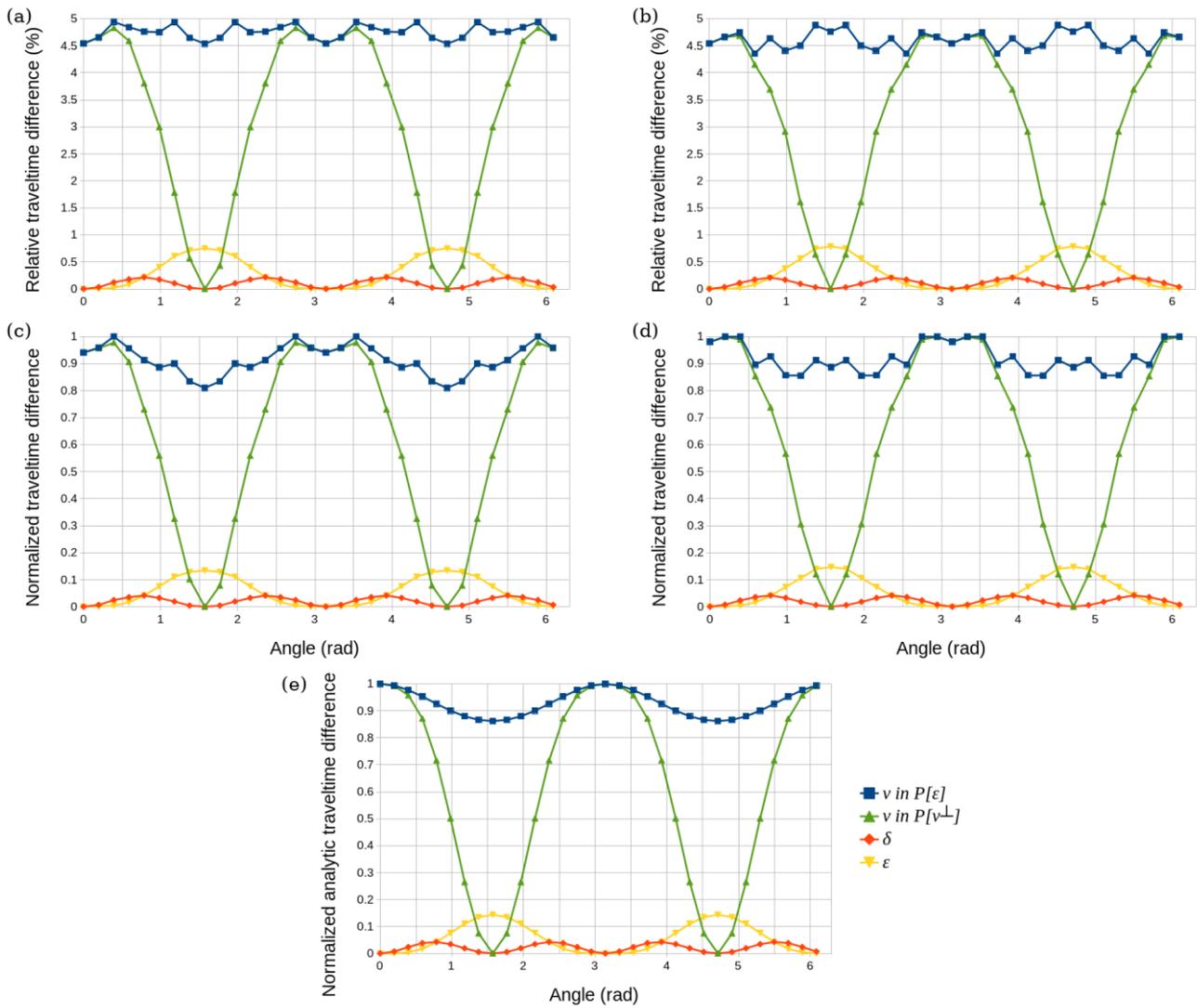
**Table 9: Results of the procedure to approximate an initial  $\delta$  model. The RMS of the final traveltime residuals shows a clear change in order of magnitude in the subranges (0.13,0.22) and (0.12,0.22) depending on the parameterization. Results for initial  $\delta = 0.16$  correspond to the examples in [Figs 4 and 5](#) [Figs 5 and 6](#).**

685

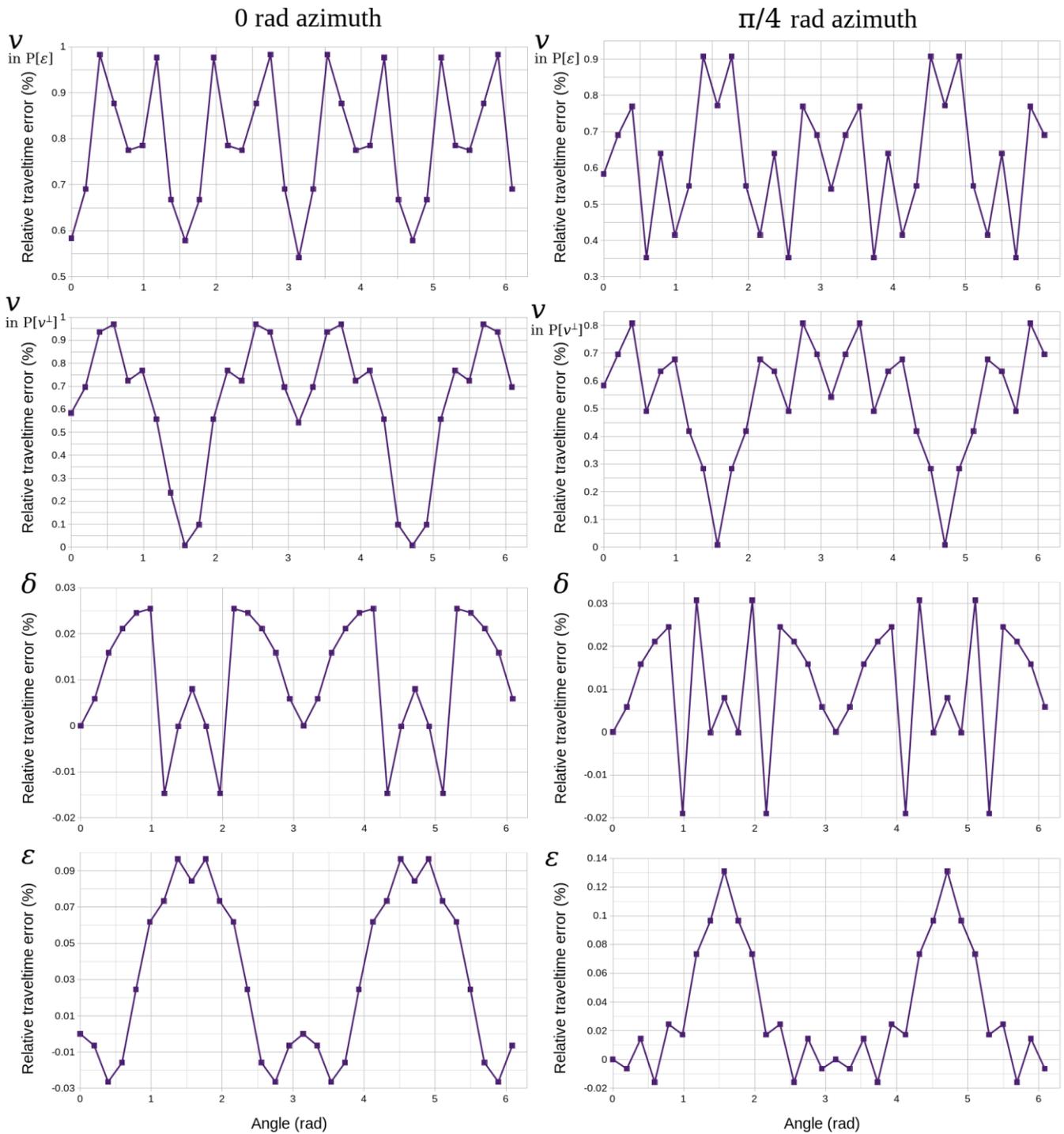
## Figures



690 **Figure 1: (a) Horizontal and (b) vertical views of the acquisition geometry for the inversion tests. 114 sources and 114 receivers (red boxes) are located at 2.5 km from the center of the model, at the locus defined by the surface of the sphere inscribed in the cube, and placed at the crossing points of 16 meridians with 7 parallels, and at each pole. Thus, in the inversion tests we used 12,882 traveltimes from 114 sources each recorded at 113 receivers, i.e. all receivers record all sources, except for the one coinciding in location. The acquisition geometry for the accuracy and sensitivity tests is similar, only in this case for sources and**  
695 **receivers at the crossing points of 32 meridians and 15 parallels, plus one of each at the two poles, and using just 482 traveltimes from diametrically-opposed source–receiver pairs arranged, i.e. each receiver exclusively records the first arrival travelttime from its paired source.**



700 | **Figure 23:** Sensitivities for the meridians at azimuths 0 rad (left) and  $\pi/4$  rad (right) as a function of the polar angle (origin in the downward vertical axis). (a) and (b) Synthetic relative sensitivities in percentage, (c) and (d) synthetic normalized sensitivities, and (e) normalized analytic sensitivities. Sensitivity values displayed correspond to all source – receiver pairs along the selected meridians of the acquisition configuration.



705 | **Figure 32: For both selected meridians, 0 rad and  $\pi/4$  rad azimuths, Relative traveltime misfits errors in percentage with respect to the analytic value for each of the four simulations used in the sensitivity analysis possible anomalies separately, and for each of**

~~the two selected meridians, 0 rad (left) and  $\pi/4$  rad (right) azimuths.~~ Polar angle origin is in the downward vertical axis. Mean values and deviations are shown in Table 2.

## Two-step sequential inversion strategy

### Step 1

(a) Inverting for  $v$  while fixing  $\delta$  and  $\varepsilon$  (or  $v^\perp$ ).

(b) Inverting for  $v$  and  $\varepsilon$  (or  $v^\perp$ ) while fixing  $\delta$ .

### Step 2

(c) Inverting for  $v$  and  $\varepsilon$  (or  $v^\perp$ ) while fixing  $\delta$ .

(d) Inverting for all three parameters.

(e)

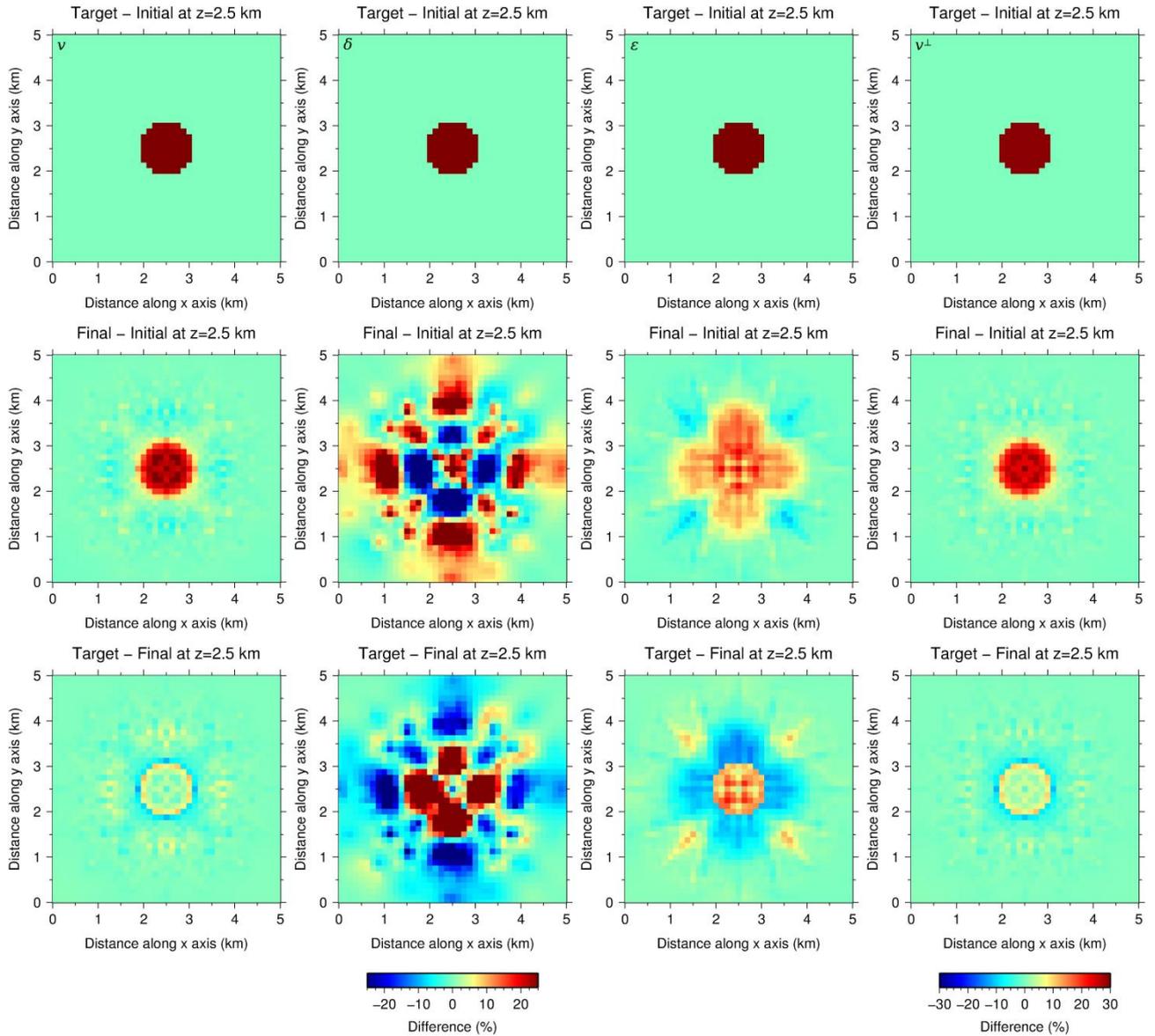
Inverting  $\delta$  while fixing  $v$  and  $\varepsilon$  (or  $v^\perp$ ).

Inverting  $\delta$  and  $\varepsilon$  (or  $v^\perp$ ) while fixing  $v$ .

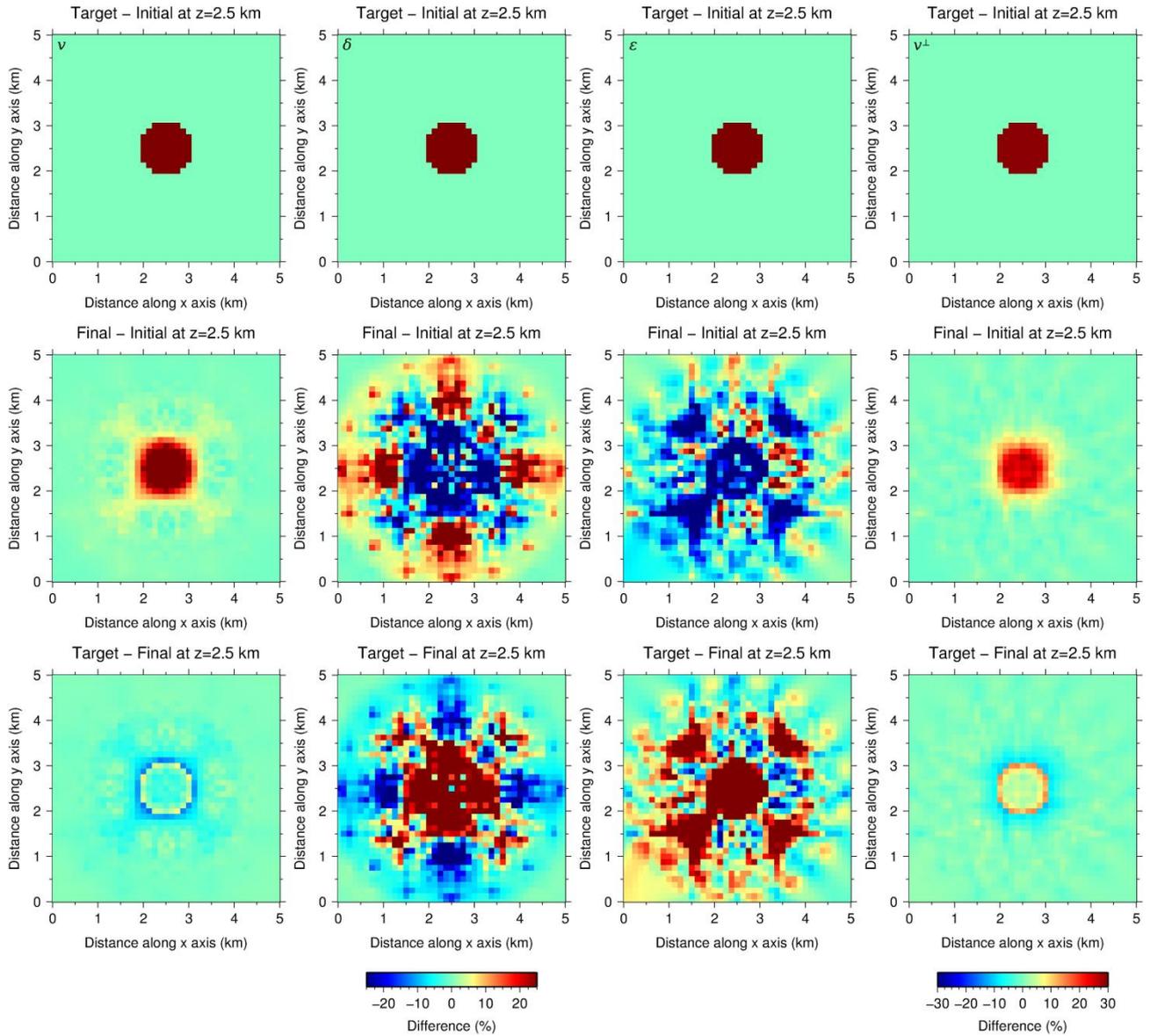
Inverting  $\delta$  and  $v$  while fixing  $\varepsilon$  (or  $v^\perp$ ).

710

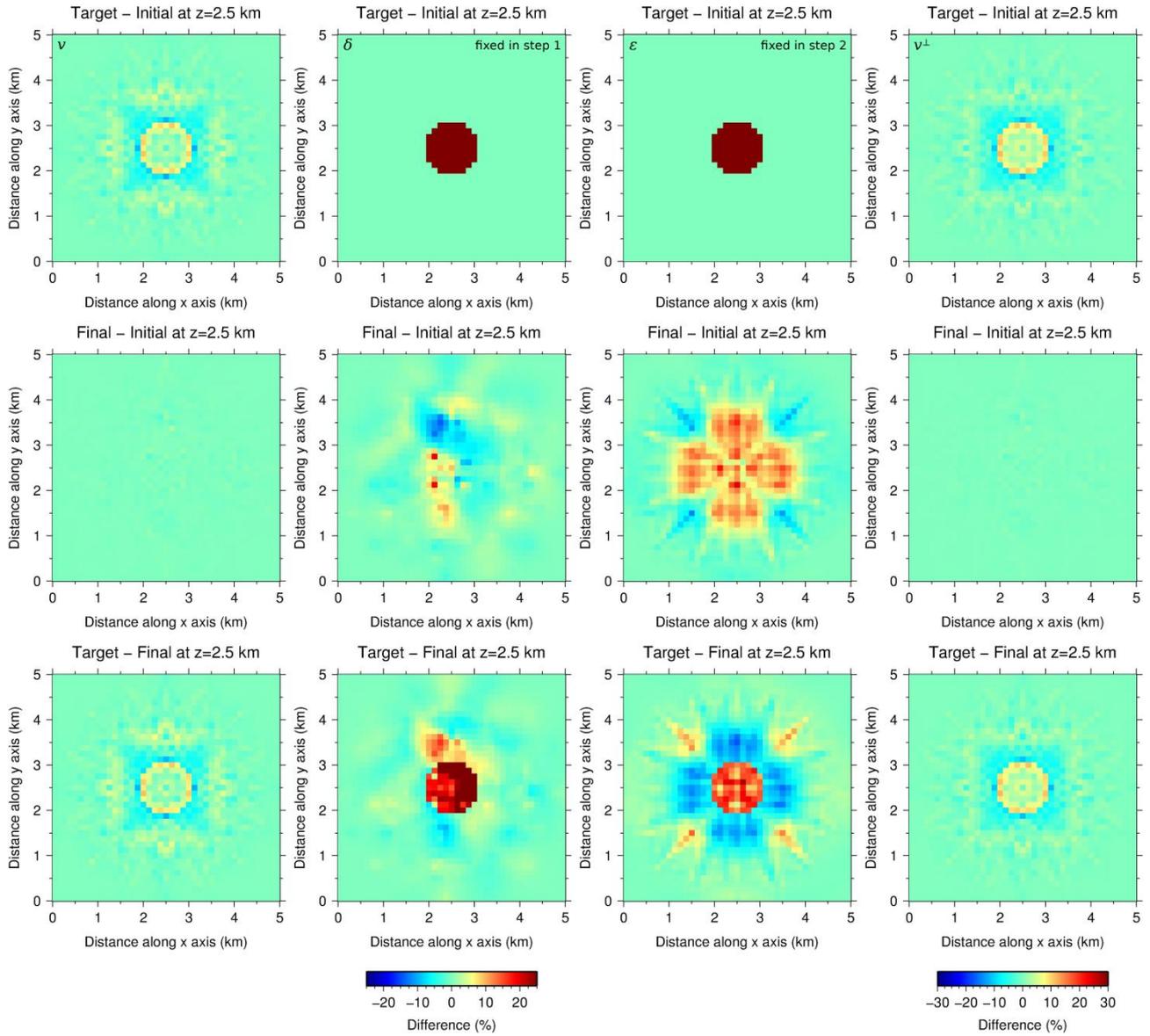
[Figure 4: Flowchart describing the two steps for all the sequential inversion options tested. The best options for the sequential inversion strategy in  \$P\[\varepsilon\]\$  \(Fig. 7\) and in  \$P\[v^\perp\]\$  \(Fig. 8\) are marked in green and red respectively.](#)



715 | **Figure 54:** Simultaneous inversion with  $P[\varepsilon]$ . Horizontal slices of the relative differences between target and initial (first row), final  
720 | and initial (second row), and target and final (third row) models at 2.5-km depth for the four parameters.  $v^\perp$  is derived from Eq. (2). The range of the color scale for  $v^\perp$  is wider than for the rest of parameters because the heterogeneity is calculated considering the 25% anomalies in  $v$  and  $\varepsilon$ , which yields a  $\sim 29.3\%$  anomaly in  $v^\perp$ . First and second row would be identical if the inversion were perfect, whereas the third row would display a homogeneous value of 0%. The quality of the recovery of each parameter is in correlation with their sensitivities (Fig. 23). Recovery of  $v$  is satisfactory with anomaly values close to the target and well-defined anomaly boundaries.  $\varepsilon$  recovery is partial, the anomaly is centered but its magnitude and shape are not as accurate as in the case of both velocities; even so it allows for a successful recovery of  $v^\perp$  through Eq. (2), both in anomaly magnitude and shape. As for  $\delta$ , recovery is unsuccessful.

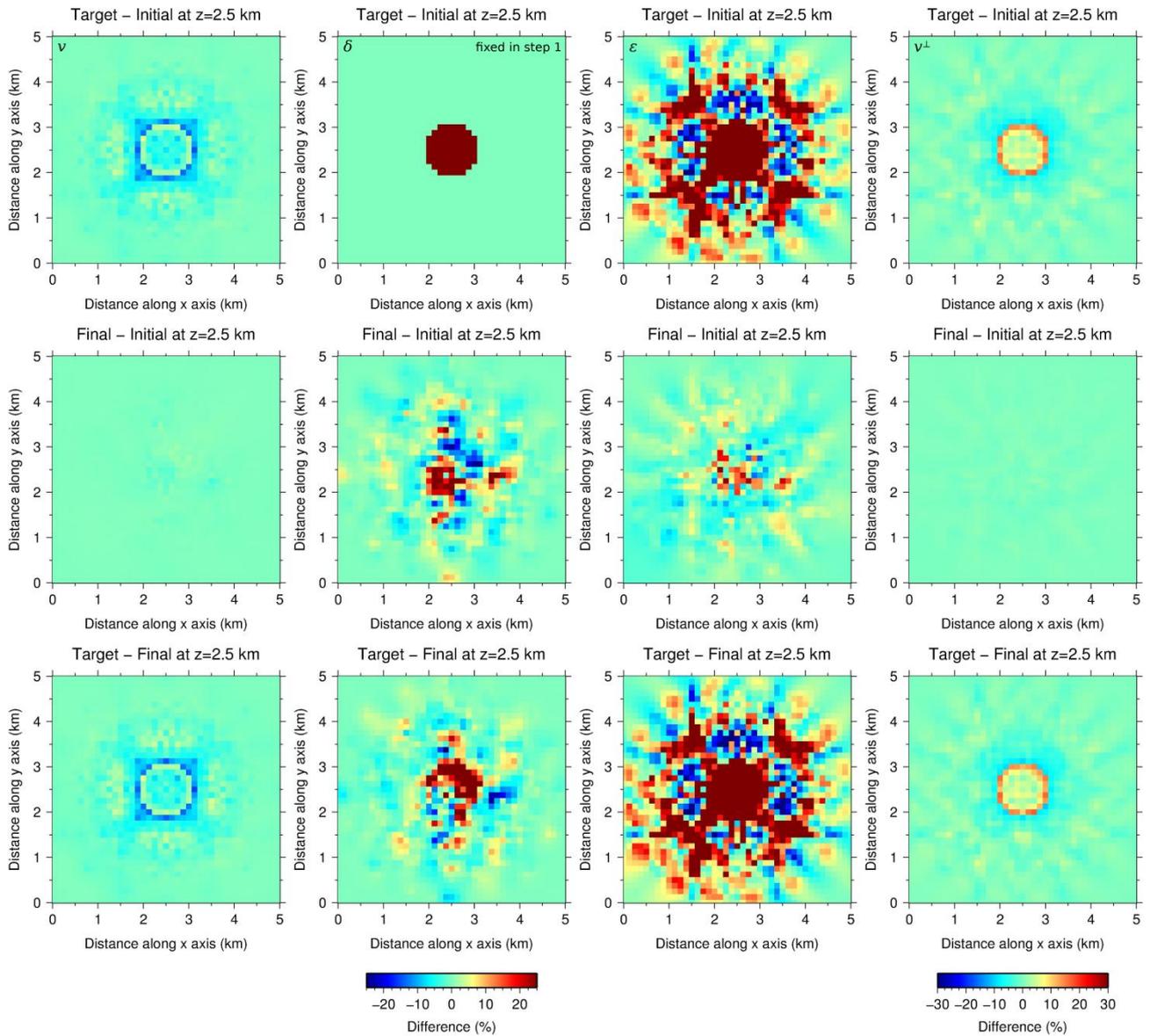


725 | **Figure 65:** Same as [Fig. 4](#)[Fig. 5](#) but with  $P[v^\perp]$ .  $\varepsilon$  is derived from Eq. (2). The quality of the recovery is in correlation with sensitivity ([Fig. 32](#)). Both velocities are satisfactorily recovered. The magnitude of the anomaly in  $v$  is better recovered than for  $P[\varepsilon]$ , whereas the opposite occurs for  $P[v^\perp]$ . Anomaly boundaries for both velocities are not as well determined as for  $P[\varepsilon]$ .  $\varepsilon$  and  $\delta$  are not recovered.



730

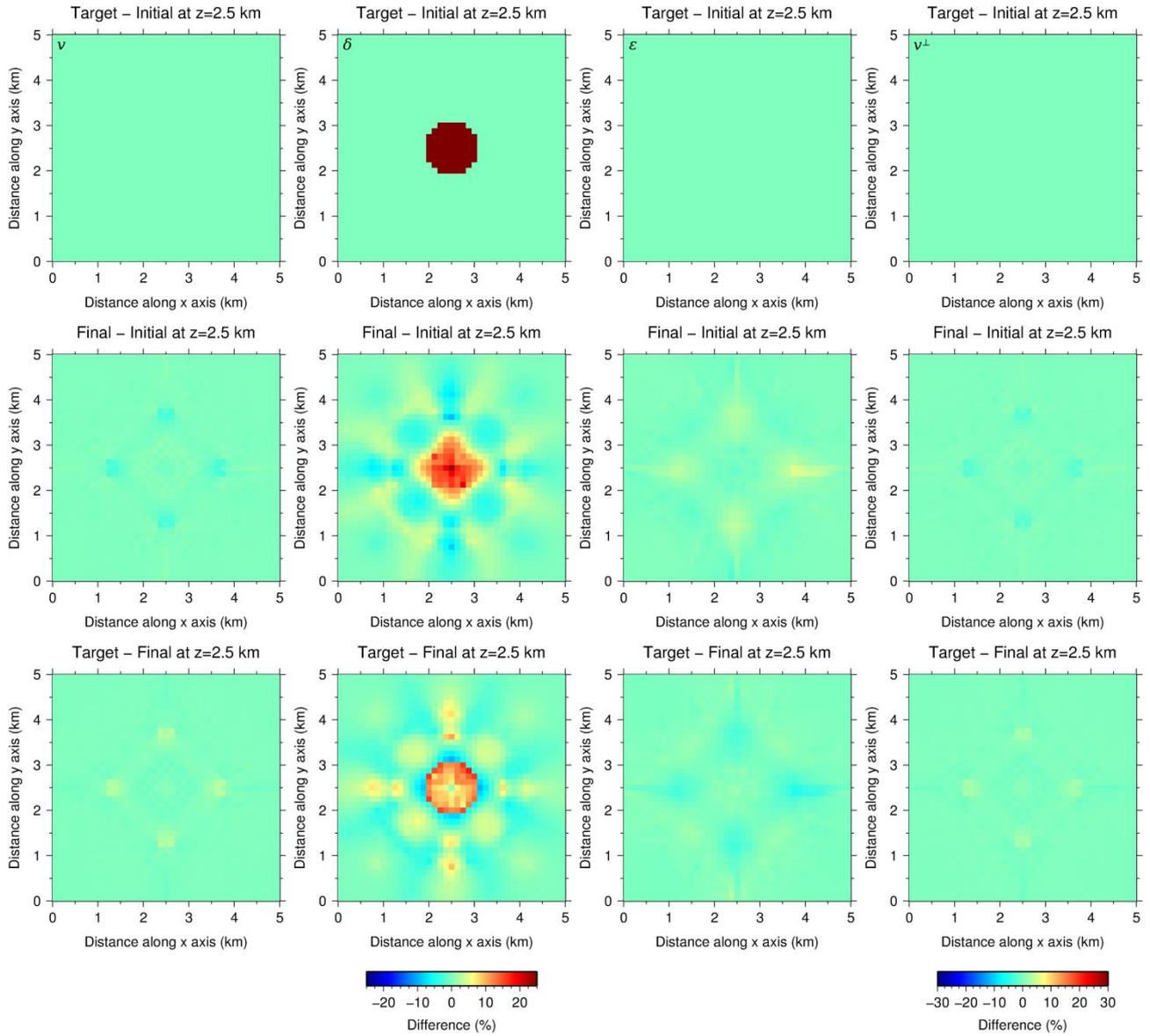
Figure 76: Same as Fig. 4 Fig. 5 but for the two-step sequential inversion strategy with  $P[\epsilon]$ . In the first step only  $\delta$  was fixed. In the second step, only  $\epsilon$  was fixed. Final models from step 1 were used as initial model for step 2.  $v$  (1st column) is well recovered in step 1 (top panel), and it is barely modified by the second step (middle and bottom panels).  $\delta$  is fixed to the initial homogeneous model in step 1, and its recovery is unsuccessful in step 2. Recovery of  $\epsilon$  is limited compared to  $v$ , but significantly better than that of  $\delta$ . Nonetheless, it proves to be good enough to provide a satisfactory recovery of  $v^\perp$  using Eq. (2).



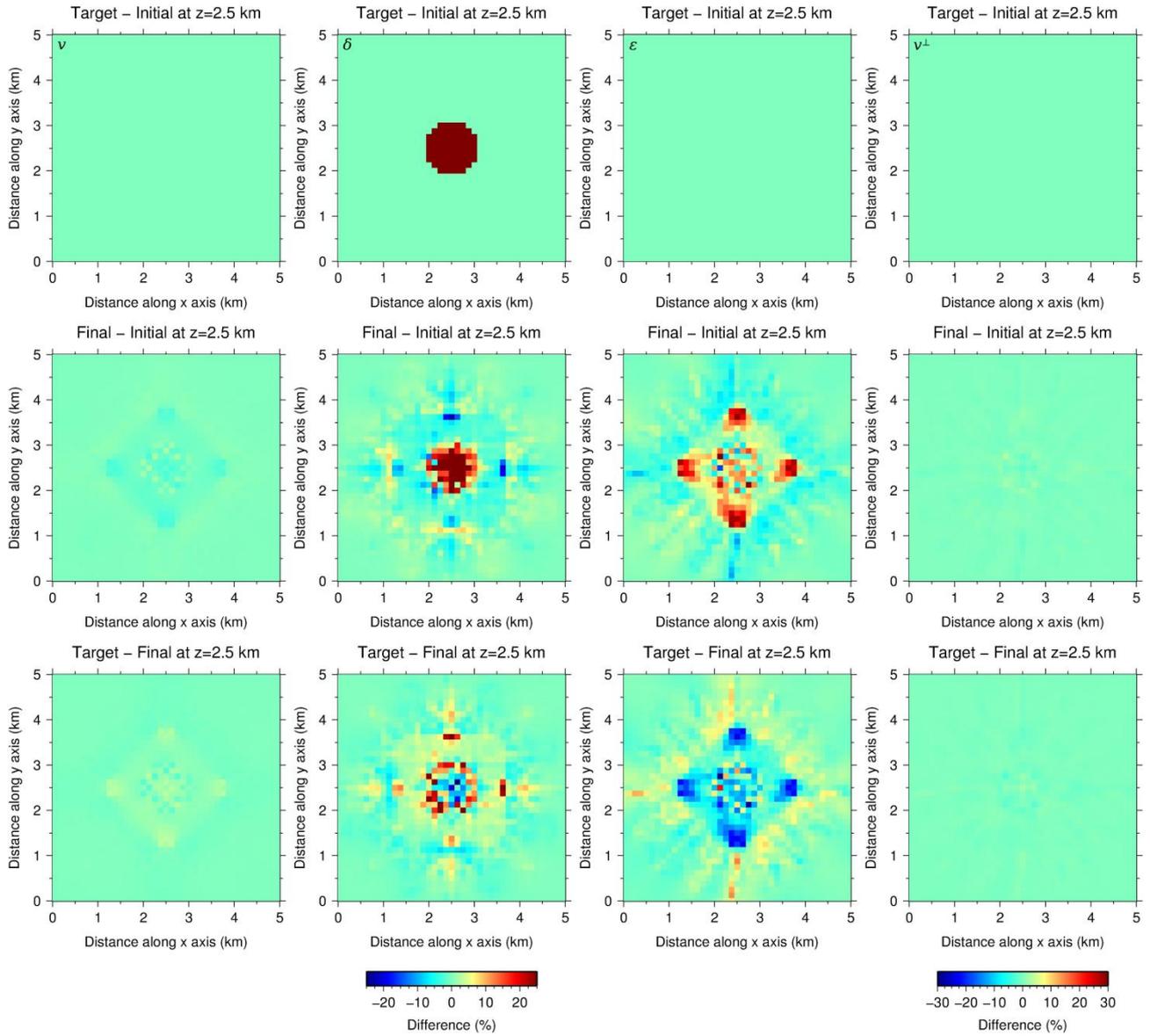
735

Figure 87: Same as Fig. 6 but for the two-step sequential inversion strategy with  $P[v^\perp]$ . In the first step only  $\delta$  is fixed, whereas in the second one all parameters are inverted.  $v$  (1st column) and  $v^\perp$  (4th column) are well recovered in step 1 (top panel), and they are barely modified by the second step (middle and bottom panels).  $\delta$  and  $\epsilon$  are not properly recovered; in both cases some sort of irregular perturbations approximately centered in the cube are retrieved but bearing no resemblance to the target anomalies.

740



745 | **Figure 98:** Same as [Fig. 4](#)[Fig. 5](#) but using target models as initial models for all parameters in  $P[\epsilon]$  but  $\delta$ . This test was conducted to study the recovery of  $\delta$  under unrealistically optimal circumstances, and even with these perfect initial conditions, recovery is, at best, extremely complicated due to the small sensitivity ([Fig. 32](#)); magnitude and shape are only partially recovered. In the first row, differences for  $v$  and  $\epsilon$  are 0% since we use target models as initial ones. For this same reason, the second and third rows show that differences between target and inverted for these three parameters are hardly observable, indicating that inversion is not modifying  $v$  and  $\epsilon$  even though they are not fixed. Consequently, the resulting recovery of  $v^\perp$  through [Eq. \(2\)](#) is almost perfect as well.



750

Figure 109: Same as Fig-8Fig. 9 but using target models as initial models for all parameters in  $P[v^\perp]$  except for  $\delta$ . The magnitude of the  $\delta$  anomaly is better recovered than for  $P[\epsilon]$ , but the shape is not as well retrieved, and artifacts appear in the background area. Differences between inverted and target  $v$  and  $v^\perp$  models are still hardly observable but not as much as for  $v$  and  $\epsilon$  in the case of  $P[\epsilon]$ , and thus the recovery of  $\epsilon$  through Eq. (2) is also not as good.

755

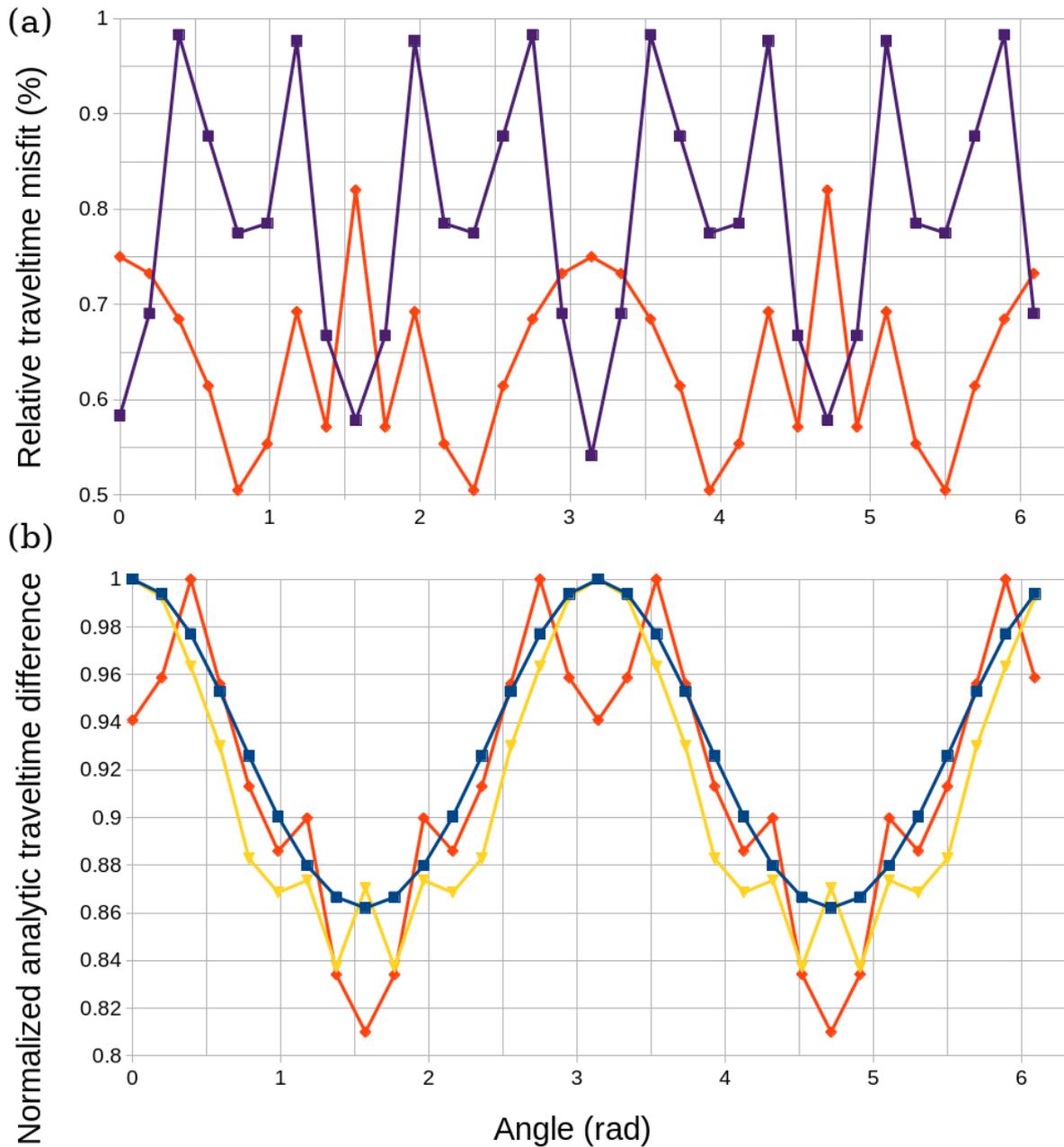
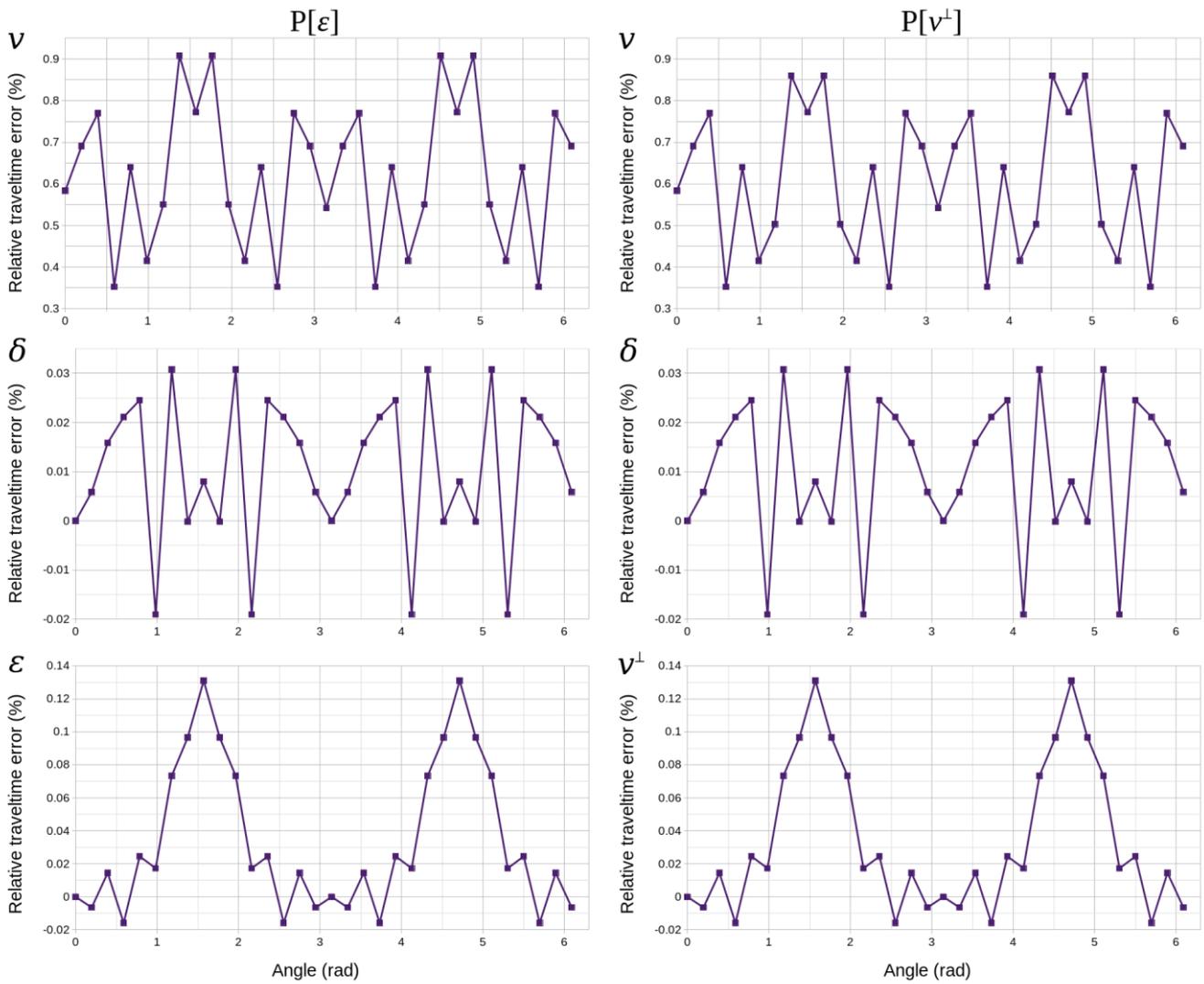


Figure S1: Improvement of (a) accuracy and (b) sensitivity pattern of  $\nu$  in  $P[\varepsilon]$  for the 0 rad meridian with refinement of the model grid by a factor of 2. In (a) the purple line marks the accuracy for the grid used in the synthetic tests in this paper as a reference for comparison with the orange line indicating the accuracy achieved using a refined grid. The refined average relative traveltime error is  $0.64\% \pm 0.08\%$  compared to the  $0.8\% \pm 0.1\%$ . In (b) the comparison is established among the blue line marking the analytic sensitivity as a reference, and the sensitivities for the grid that we used (orange line) and the refined grid (yellow line). The refined grid reduces the relative traveltime error and produces a better fit of the analytic sensitivity.



765 **Figure S2: For the  $\pi/4$  rad meridian, relative traveltme errors in percentage with respect to the analytic value for three pairs of equivalent simulations used in the sensitivity tests. Polar angle origin is in the downward vertical axis. Both parametrizations produce almost identical accuracies.**

*Mathematical proof for the shapes of  $v$  sensitivity in both parametrizations*

770 Regarding the shapes of relative and normalized sensitivities for  $v$ , in the following we provide mathematical proof showing that the former is constant whereas the latter is sinusoidal.

We use  $\Delta t$  to refer to the difference between the travel times measured with and without the 25% anomaly.  $x_A$  is the thickness of the anomaly whereas  $x_B$  is the total ray path length from source to receiver.  $v_A$  and  $v_B$  are the anisotropic velocities in the anomaly and in the background respectively.

$$\Delta t = \Delta \left( \frac{x}{v} \right) = \frac{x_B}{v_B} - \left( \frac{x_A}{v_A} + \frac{x_B - x_A}{v_B} \right) = \frac{x_A(v_A - v_B)}{v_A v_B}$$

775 According to the definition we have just given, the sensitivity expressed as relative difference is

$$S_R = \frac{\frac{x_A(v_A - v_B)}{v_A v_B}}{t_R}$$

where  $t_R$  is the travel time measured without the anomaly

$$t_R = x_B / v_B$$

With that  $S_R$  becomes

$$S_R = \frac{x_A}{x_B} \left( 1 - \frac{v_B}{v_A} \right)$$

In the case of a 25% anomaly in the  $v$  parameter in  $P[\varepsilon]$ ,  $v_A$  and  $v_B$  are as follows

$$v_A = v_{PA}(1 + \delta \sin^2 \theta \cos^2 \theta + \varepsilon \sin^4 \theta)$$

$$v_B = v_{PB}(1 + \delta \sin^2 \theta \cos^2 \theta + \varepsilon \sin^4 \theta)$$

We use  $v_{PA}$  and  $v_{PB}$  to refer to the axis-parallel velocity parameter  $v$ . For a 25% anomaly in  $v$  their proportion is

$$v_{PA} = 1.25 v_{PB}$$

780 The values for  $x_A$  and  $x_B$  are

$$x_A = 1 \text{ km and } x_B = \text{ km}$$

Thus, the final expression for  $S_R$  for a 25% anomaly in  $v$  in  $P[\varepsilon]$  is a constant 0.05 (or 5%).

$$S_R = \frac{1}{5} \left( 1 - \frac{1}{1.25} \right) = 0.05$$

Regarding the sensitivity expressed as normalized difference, the general expression for an anomaly in any of the four parameters is

$$S_N = \frac{\frac{x_A(v_A - v_B)}{v_A v_B}}{\Delta t_{MAX}} = \frac{x_A}{\Delta t_{MAX}} \left( \frac{1}{v_B} - \frac{1}{v_A} \right)$$

785 where  $\Delta t_{MAX}$  is the maximum travel time difference among all four parameters. This expression will contain some combination of sinusoidal functions for all four possible anomalies. The sensitivity (normalized and relative) pattern for  $v$  in  $P[v^\perp]$  is different than for  $P[\varepsilon]$  although it follows the same sinusoidal pattern and it has equal maxima. We can see that in the case of  $P[v^\perp]$  both  $S_R$  and  $S_N$  will display a sinusoidal shape. Now  $v_A$  and  $v_B$  are

$$v_A = v_{PA} \left( 1 + \delta \sin^2 \theta \cos^2 \theta + \left( \frac{v^\perp}{v_{PA}} - 1 \right) \sin^4 \theta \right)$$

$$v_B = v_{PB} \left( 1 + \delta \sin^2 \theta \cos^2 \theta + \left( \frac{v^\perp}{v_{PB}} - 1 \right) \sin^4 \theta \right)$$

and the sine and cosine functions do not cancel out in  $S_R$ , as they did in the case of an anomaly in  $v$  in  $P[\varepsilon]$ , nor in  $S_N$ . For the  
 790 latter it is trivial to see that the sinusoidal dependencies are identical to the case in  $P[\varepsilon]$ .

795 *Comments in file se-2019-44-RC1.pdf*

3- Are substantial conclusions reached?

Partly. Tests were done with synthetic data, which allow evaluating quantitatively the performance of the inversion. However, I think that the conclusions are not fully supported by the presented work, this for three reasons. First, the authors  
800 did not study the influence of noise on the robustness of the results. Inevitably, noise is present in field data (picking accuracy, timing accuracy, statics, etc) and at least the effect of some gaussian noise should be investigated with the synthetic data. Second, the data acquisition geometry could never be achieved in reality. For the presented tests, an anomalous spheric body is surrounded by sources (Tx) and receivers (Rx) in the whole space. This geometry supposes first that the location of the anomalous body is known, and second that access is possible underground almost everywhere around  
805 the body. At best, surface and a few borehole Tx & Rx are typically available for typical surveys. For such geometries, the forward operator does not allow uniform resolution, such as illustrated in the results of the paper. So I think that the capacity to resolve the anomaly is over optimistic. Third, the initial model is quite close to the true model (the initial model is equal to the background of the true model). I don't think that it is realistic to know that well the properties of the background, especially when the background itself is anisotropic.

810 (\*) We are aware that the synthetic experiment that we designed to perform our tests is not realistic. In fact, our goal is to use this canonical benchmark to assess accuracy, sensitivity, and performance of the four inversion strategies under ideal and equal conditions for all anisotropic parameters, avoiding the specificities and biases of a synthetic experiment simulating a particular field case study. These tests on a canonical benchmark provide conclusions that are generally informative of the code's performance. In other words, here we are not interested in the performance of the code in a more or less specific  
815 geological and experimental context, but in obtaining an upper limit to the code's capabilities, an ideal but generalizable estimation for the code's performance.

We have modified the manuscript to clarify the purpose of our testing approach and the reasons behind it. Specifically, we have edited the Abstract (lines 10 and 11), the third paragraph in the Introduction, the first paragraph to section 3 Synthetic  
820 tests, and the second paragraph in the Conclusions.

Nonetheless, in an upcoming paper presenting an application of the code to an anisotropic field data case, we will first evaluate the code's performance on a realistic synthetic experiment simulating this particular field study. This simulation should yield an estimation of the potential quality of the results as well as information on the best modeling strategy to  
825 approach this specific case, and it will include noise, an initial model obtained from isotropic tomography, and will replicate the same exact acquisition geometry used in the field. By synthetically reproducing a more or less specific type of seismic

experiment under some realistic circumstances, e.g. a certain noise level, we will get an idea of what to expect in that particular case with the selected noise level, receiver and source densities and distributions, geological features and anomalies, a priori information available in the initial models, etc. Therefore, this sort of realistic synthetic testing and the conclusions that can be drawn from it become relevant and meaningful as preliminary work linked to a particular field data application.

We believe that adding these other type of tests here would result in an exceedingly long manuscript covering too many aspects. Also, we think that realistic synthetic tests are better presented along with the field data that they are simulating. Thus, we prefer to separate our work into two publications, this first one of technical and methodological content, and a second one focusing on a field data application. Figure R1 corresponds to this anisotropic field case. In Sallarès et al. [2013] we obtained an isotropic  $V_p$  model from a refraction and wide-angle reflection seismic (WAS) data set (sub-horizontal propagation) and compared it to the image obtained from multichannel seismic (MCS) reflection data (near-vertical propagation). The top plot shows the isotropic  $V_p$  model, with the vertical coordinate converted from depth to two-way time (TWT), superimposed onto the MCS image. The white circles inside the model delineate the geometry of inter-plate reflection imaged by MCS data. The thick red line corresponds to the TWT-converted inter-plate boundary obtained from WAS data. The mismatch between the two locations of the inter-plate boundary is most likely due to some degree of seismic anisotropy between near-vertical and sub-horizontal propagations. In the bottom plot we increase the  $V_p$  values by 15%, and with this the MCS and WAS locations of the inter-plate boundary now display a good match. Thus, this 15% increase is a good initial estimate of  $\epsilon$ , and it indicates, as a general trend, that near-vertical propagation is ~15% faster than sub-horizontal propagation in this area of the subsurface ( $\epsilon \approx 0.15$ ). For an estimate of  $\delta$  we will use the  $V_{NMO}$  model from the normal move-out correction of MCS data processing and the isotropic  $V_p$  model.

4- Are the scientific methods and assumptions valid and clearly outlined?

Some assumptions are not realistic (see point above).

See answer to point 3.

5- Are the results sufficient to support the interpretations and conclusions?

See point 3.

See answer to point 3.

10- Is the overall presentation well structured and clear?

Overall yes, but the section describing the inversion strategies is somewhat hard to follow (especially the details of the sequential inversion). I think that a figure with flowcharts could help.

We added a figure depicting the flowchart for the two-step sequential inversion (new Figure 4).

12- Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Overall yes, but I think that the anisotropy parameters epsilon & delta should be formally defined in section 2. Some parameters in equation 4 are not defined.

865 (#) The definitions of  $\delta$  and  $\varepsilon$  are available in Thomsen [1986] which is the main reference for our work and for weak VTI anisotropy in general. We chose to focus on our formulation, and we only reproduced the formula for anisotropic  $V_p$  because it is at the core of our anisotropic modeling tool. However if the Editor deems it necessary to include these definitions in our manuscript we will do so. Also we corrected the definitions for Equation 4.

870 13- Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

Perhaps the terms for the reflected rays could be removed from the equations (the case of reflected rays is not treated in the paper).

We removed Equations 11 and 12 as well as the terms related to reflections and interface depth in Equation 3. They are not used in this paper, and we will discuss them in the upcoming article presenting an application to anisotropic field data.

875

14- Are the number and quality of references appropriate?

Yes, but some references cited in the text are not in the list at end of the paper.

We have added the missing citations to the reference list. Lines 596 to 598, 605, and 606.

880 *Comments in se-2019-44-RC1-supplement.pdf*

Section 2 Modelling anisotropy

This section should contain the mathematical definition of epsilon and delta.

See answer marked with (#).

885

Subsection 2.2 Anisotropy in TOMO3D: Thomsen's weak anisotropy approximation

Three cited publications missing in the reference list.

We added the three missing citations to the reference section. See answer to point 14.

890 Do you assign the properties within the cells or at the grid nodes? e.g. is the number of unknown parameters a function of the number of cells or the number of nodes?

Properties are assigned to the grid nodes, so that the number of unknowns is a function of the number of nodes. We have modified the text to clarify this point (lines 108 and 109).

895 Equation 3

Delta z not defined.

We removed the parameters related to reflected rays and reflector depth from Equation 3, which includes  $\Delta z$ , since they are not used in this paper. They will be discussed in the field data application paper.

900 Equation 4

$r_m$  not defined.

The definition of  $r_m$  was already in the text (line 165) although the subscript m was missing. However, we have corrected lines 162 to 167 to clarify the description of Equation 4.

905 Section 3 Synthetic tests, line 195

You mean the values of the v, epsilon and delta parameters at grid nodes?

Yes. We have edited this phrase to clarify its meaning (line 215).

Subsection 3.1 Accuracy, lines 198 and 199

910 The body will not eventually contain anomalous v, epsilon and delta?

Yes, and that is what we consider when performing the inversion tests. However, regarding accuracy and sensitivity we wish to evaluate each parameter separately so that we can compare the code's performances for each of them. We have edited and reordered sections 3.1 and 3.2 to stress this point, and also as a result of a comment by the other referee.

915 Lines 203 to 205

This\* is totally unrealistic, how do you expect that your conclusions will be valid for real life scenarios? [\*referring to the acquisition geometry]

As mentioned in the answer marked with (\*), our goal is to obtain an upper limit to the code's performance, in this case regarding the accuracy for each parameter. The conclusions are valid in the sense that they inform us on the limits of our modeling tool. We do not pretend in any way that the canonical benchmark model or acquisition used simulate any particular field study.

920 Line 209

When computed with respect to the analytic solution, this is a relative error. A misfit is generally the difference with an observation.

We replaced "misfits" by "errors" here and elsewhere in the manuscript when referring to the relative difference between calculated and analytic traveltimes.

Subsection 3.2 Sensitivity, line 232

930 Would be interesting to see if the relationship is linear with respect to changes in epsilon and delta.

Figure R2 shows the analytic  $\delta$  sensitivities for anomalies of 10%, 15%, 20%, and 25%. The linear relationship between sensitivity and anomaly increment is particularly clear observing the evolution of the maxima. Given the form of Equation 1, one can infer that  $\epsilon$  sensitivity will have an analogous behavior.

935 Subsection 3.3 Inversion results, line 250

How can one expect to know the anisotropy parameters of the background? Moreover, you did not study the effect of noise on the robustness of the inversion.

(^) There exist compilations of values for Thomsen's anisotropy parameters (e.g. Thomsen [1986] for sedimentary rocks and related materials) or for the components of the elastic modulus tensor (e.g. Almqvist and Mainprice [2017] for continental crust rocks) from which Thomsen's parameters can be derived. Also, we can extract estimates of  $\delta$  from multichannel reflection data processing, specifically using the mathematical relationship between  $v$ ,  $\delta$ , and  $V_{NMO}$ , and an isotropic estimate of  $v$ . An estimate for  $\epsilon$  can be obtained by comparing velocity models derived from multichannel and wide-angle seismic data since they are representative of near-vertical and sub-horizontal propagations respectively, and  $\epsilon$  represents the relative difference between them (Fig. R1). We will detail and conduct these procedures to estimate  $\delta$  and  $\epsilon$  for initial model building  
945 in field data applications in the upcoming paper.

As discussed in answer (\*), we assume that we know this background value perfectly in order to evaluate the code's performance in retrieving an anomalous body within this background in what is an ideal situation providing an upper-limit but generalizable estimation of the code's capabilities and indicating potential weaknesses. For identical reasons we do not  
950 include noise in our data.

Lines 254 and 255

Not clear [referring to the description of the acquisition geometry].

We edited lines 321 to 326 to clarify the acquisition geometry.

955

Lines 277 to 279

This\* is quite confusing, use a flowchart for each [\*referring to the description of the sequential inversion].

We added a figure with a flowchart reproducing the two-step sequential inversion strategy (new Figure 4).

960 Section 4 Discussion, line 368

The problem is actually to obtain this a priori information.

As mentioned in answer (^), it is possible to obtain reasonable estimates of the anisotropy parameters.

Lines 376 to 378

965 What is the basis for this statement? [referring to “it would probably be recommendable to use a coarser discretization for  $\delta$  than for the other parameters, and in general a finer discretization for the more sensitive parameters (Fig. S1)”].

The heterogeneities that can be resolved for a particular parameter depend on its sensitivity. The more sensitive a parameter is, the smaller the spatial scale and the relative variation of the resolvable heterogeneities. An heterogeneity of a given scale and variation will produce a greater effect on data for a parameter of greater sensitivity. Thus, for a parameter of greater  
970 sensitivity, it will be easier for the code to identify smaller heterogeneities both in scale and variation, which will require a finer grid. We have modified the manuscript to clarify this point (lines 461 to 464).

Section Author contribution

Data acquisition? There is no data in that paper.

975 “Funding acquisition” means “acquisition of the financial support for the project leading to this publication” as defined in section The CRediT Roles at <https://www.casrai.org/credit.html>. We have replaced “funding acquisition” by “acquisition of financial support”.

Table 4

980 The opposite [referring to the ideal values for AI and AT in “AI and AT indicate the resemblance between true and recovered anomalies, and their ideal values are 0% and 25% (~29.3% in the case of  $v \perp$  ) respectively”].

Corrected.

Table 5 (to Table 8)

985 Keep consistent notation [referring to the systematic use of either AI or Ai, and AT or At].

Corrected.

990 *Comments in file se-2019-44-RC2.pdf*

#### General comments

The authors present a modification to TOMO3D seismic tomographic inversion code which allows to include anisotropy parameters (assuming VTI symmetry) in the inversion. Such code is a valuable and interesting contribution, as most of the existing tomographic codes assume an isotropic medium. The manuscript presents necessary evaluation of reliability of the code. Formally, the structure of the paper is correct, it contains all necessary parts and is generally well written, although some parts (details below) are hard to follow. The main goal of the manuscript is to evaluate three characteristics of the code/method: the accuracy of the forward code, sensitivity of the method for anomalies of model parameters and resolving ability of the inversion. In my opinion, the papers gives satisfactory answer for first two questions are basically satisfactory, while the third question is studied only partially (details below). An interesting outcome of the inversion tests is the discussion about poor sensitivity of the travel time data on the  $\delta$  parameter, even in case of 'ideal' measurements geometry (spherical geometry and uniform angular spacing of sources/receivers locations).

#### Specific comments

From the parametrization used (Thomsen parameters) it implicitly follows that the VTI symmetry of the medium is assumed (also, the inversion does not solve for orientation of the symmetry axis, so I guess it is assumed to be vertical) but it is not clearly stated in the text, and statements like 'anisotropic tomography' (also in the title!) suggest at the first glance that the code can be used for media with more general and more complicated symmetries. If I'm right about this, I would recommend to state it explicitly that the code can model VTI media only. This would make clear that it cannot be used e.g. for modeling of azimuthal anisotropy (which is usually due to HTI or TTI medium).

The first line of the Abstract states that with this article "We present the implementation of Thomsen's weak anisotropy approximation for VTI media within TOMO3D". In the third paragraph of the Introduction we describe the two main objectives of this work pointing out that the we present "the anisotropic version of TOMO3D for the study of VTI weakly-anisotropic media". We believe that with these two statements it is clear that the anisotropic version of TOMO3D assumes VTI symmetry for the media, but we also added "VTI" to the title of subsection 2.2 that now reads "Anisotropy in TOMO3D: Thomsen's weak VTI anisotropy formulation", in line 69 in the introduction for section 2 Modelling anisotropy, and in the first paragraph of the Conclusions.

Regarding the title, we consider that mentioning Thomsen's weak approximation is sufficient since it implies that VTI symmetry is used, as the referee corroborates. However, we would not oppose changing it to "Anisotropic P-wave travelttime tomography implementing Thomsen's weak VTI approximation in TOMO3D".

L120: The feasibility of determination of the  $\delta$  parameter is widely discussed in the text, but the definition of the parameter itself is not given. I think it should be added to the text.

1025 The definition of  $\delta$  is available in Thomsen [1986] which is the main reference for our work. We chose to focus on our formulation, and we only reproduced the formula for anisotropic  $V_p$  because it is key to our anisotropic modeling tool. However if the Editor deems it necessary to include this definition in our manuscript we will do so.

L128-132: I understand the relationship between anisotropy parameters and  $V_{nmo}$ , but the comment and conclusion about the  $V_{nmo}$  in L128-132 is unclear for me.

1030 However, a mathematical relationship between  $\delta$ ,  $v$ , and the normal move-out velocity ( $V_{NMO}$ ) exists.  $V_{NMO}$  models are built as part of the normal move-out correction in seismic reflection data processing. At best, our travel time tomographic method would be able to produce approximations of the actual  $V_{NMO}$  models. Furthermore, such approximations would only be meaningful, if ever, when derived from travel times of a seismic reflection data set, for which the normal move-out correction and thus the  $V_{NMO}$  are defined. Of course, in such a case, actual  $V_{NMO}$  models would be obtained from the normal move-out correction, and therefore  $\delta$  could be calculated provided that a  $v$  model is available, for instance from our travel time tomography. Thus, we only consider Eq. (2), and we implemented two parametrizations of the medium:  $(v, \delta, \epsilon)$  and  $(v, \delta, v_{\perp})$ .

1040 The conclusion is that we do not consider  $V_{NMO}$  to be a useful parameter in describing the anisotropic VTI media so that we do not implement parametrizations  $(v, V_{NMO}, \epsilon)$  and  $(v, V_{NMO}, v_{\perp})$ . Unlike axis-parallel and axis-perpendicular propagation velocities, which define the physics of the medium,  $V_{NMO}$  is a by-product of multichannel seismic reflection data processing, a mathematical construct that involves the assumptions of a stratified media with constant velocity layers and of small spread, i.e. near-vertical propagation. The best possible way to obtain a  $V_{NMO}$  model is obviously as this by-product of the processing of the sort of data that it is defined for. It does not seem clever to try and obtain it by other means, less so if the data and modeling used to do so do not fulfill the necessary assumptions of the processing for which  $V_{NMO}$  is defined.  $\delta$  is needed to describe the weak VTI anisotropy as proposed by Thomsen [1986], whereas we do not need  $V_{NMO}$ .

1050 We admit that the original text did not convey these ideas in a clear way, and so we have modified the text to clarify our points (paragraph starting at line 131).

-Accuracy and sensitivity tests:

Fig. 3: what is the meaning of the ‘normalized difference’? And why ‘relative difference’ for  $v$  is constant for all angles, while ‘normalized difference’ is sinusoidal?

1055 (^) Normalized difference is the difference between the travel times measured with and without anomaly divided by the greatest of these differences among all four parameters. Relative difference is the difference between the travel times

measured with and without anomaly divided by the travel time without anomaly (and multiplied by 100). These are the relative changes produced in travel times by an anomaly of 25% relative change in each particular parameter and as a function of the polar angle. We realize that these two manners of presenting sensitivity are not properly described in the text. We have modified it to include their definitions (first paragraph of new section 3.1 Sensitivity).

1060

Regarding the shapes of relative and normalized sensitivities for  $v$ , in the following we provide mathematical proof showing that the former is constant whereas the latter is sinusoidal (also in the supplementary material).

1065

We use  $\Delta t$  to refer to the difference between the travel times measured with and without the 25% anomaly.  $x_A$  is the thickness of the anomaly whereas  $x_B$  is the total ray path length from source to receiver.  $v_A$  and  $v_B$  are the anisotropic velocities in the anomaly and in the background respectively.

$$\Delta t = \Delta \left( \frac{x}{v} \right) = \frac{x_B}{v_B} - \left( \frac{x_A}{v_A} + \frac{x_B - x_A}{v_B} \right) = \frac{x_A(v_A - v_B)}{v_A v_B}$$

According to the definition we have just given, the sensitivity expressed as relative difference is

$$S_R = \frac{\frac{x_A(v_A - v_B)}{v_A v_B}}{t_R}$$

where  $t_R$  is the travel time measured without the anomaly

$$t_R = x_B / v_B$$

With that  $S_R$  becomes

$$S_R = \frac{x_A}{x_B} \left( 1 - \frac{v_B}{v_A} \right)$$

1070

In the case of a 25% anomaly in the  $v$  parameter in P[ $\varepsilon$ ],  $v_A$  and  $v_B$  are as follows

$$v_A = v_{PA}(1 + \delta \sin^2 \theta \cos^2 \theta + \varepsilon \sin^4 \theta)$$

$$v_B = v_{PB}(1 + \delta \sin^2 \theta \cos^2 \theta + \varepsilon \sin^4 \theta)$$

We use  $v_{PA}$  and  $v_{PB}$  to refer to the axis-parallel velocity parameter  $v$ . For a 25% anomaly in  $v$  their proportion is

$$v_{PA} = 1.25 v_{PB}$$

The values for  $x_A$  and  $x_B$  are

$$x_A = 1 \text{ km and } x_B = 1 \text{ km}$$

Thus, the final expression for  $S_R$  for a 25% anomaly in  $v$  in P[ $\varepsilon$ ] is a constant 0.05 (or 5%).

$$S_R = \frac{1}{5} \left( 1 - \frac{1}{1.25} \right) = 0.05$$

1075

Regarding the sensitivity expressed as normalized difference, the general expression for an anomaly in any of the four parameters is

$$S_N = \frac{x_A(v_A - v_B)}{v_A v_B \Delta t_{MAX}} = \frac{x_A}{\Delta t_{MAX}} \left( \frac{1}{v_B} - \frac{1}{v_A} \right)$$

where  $\Delta t_{MAX}$  is the maximum travel time difference among all four parameters. This expression will contain some combination of sinusoidal functions for all four possible anomalies.

1080 L198: The authors wrote “..for each of the four parameters. . .25% anomaly was added, while . . . the rest of parameters was homogeneous” – This seems to be contradictory (or it is explained in a misleading way). Parameters  $v$ ,  $vT$  (I use  $vT$  instead of the symbol used in the manuscript for  $v$  in perpendicular direction) and  $\epsilon$  are interrelate by Eq. 2, so changing  $vT$  and keeping  $v$  constant must change also  $\epsilon$  – in result, two parameters are changed. I know that this is not a problem for inversion, as parameters  $vT$  and  $\epsilon$  are used alternatively, not in the same time, but such description introduces confusion in  
1085 interpretation of the sensitivity test results. Also, even if anomalies of every parameter are the same (25%), we can hardly compare cases of  $\epsilon$  and  $vT$  anomalies: - 25%  $\epsilon$  anomaly results in value of 0.2 (compared to background 0.16), while: - for  $vT$ , the same anomaly (25%) with respect to background 2.32 results in  $vT=2.9$ . This, connected with  $v$  in the anomaly being 2.0 (undisturbed, so equal to background), results in  $\epsilon=0.45$  –an anomaly few times larger than in previous case. (This is probably the reason why in Fig. 3 sensitivity for  $\epsilon$  reaches  $\sim 0.7\%$ , while sensitivity of  $vT$  is few times larger and reaches  
1090 4.5%).

There are various points to be addressed in this comment. We most likely did not explain the procedure with sufficient detail. Moreover, we wish to note that this has been a much useful comment that has made us revisit our sensitivity analysis and realize that it needed some corrections. We have modified and reordered sections 3.1 and 3.2 to correctly address these points and to present the new elements in the sensitivity analysis regarding  $v$ ,  $\epsilon$ , and  $v^\perp$ . More specifically, we created new  
1095 Figures 2 and 3, and included a new Table 1 describing the various model combinations used in both the accuracy and sensitivity analyses.

First of all, just as in the inversion, the forward problem is solved for either one or the other of the two possible parametrizations ( $P[\epsilon]$  and  $P[v^\perp]$ ), and equation 2 is respected in both cases. When computing the sensitivity for  $\epsilon$  we are using  $P[\epsilon]$  and thus the model is defined by a  $v$  model (background), a  $\delta$  model (background), and an  $\epsilon$  model (background with the intruded anomaly). Analogously, when calculating the sensitivity for  $v^\perp$  we are using  $P[v^\perp]$  with a  $v$  model (background), a  $\delta$  model (background) and a  $v^\perp$  model (background with intruded anomaly). However, as the referee observed, this combination implies  $\epsilon=0.45$  in  $P[\epsilon]$ . This value is greater than the  $\sim 0.2$  limit assumed for the weak anisotropy approximation [Thomsen, 1986]. Therefore, since it is not possible to establish the comparison between  $\epsilon$  and  $v^\perp$  based on an  
1105 equal anomaly increase of 25% and measuring their respective effects on travel times, we now decided to compare the

sensitivities of these two parameters based on an equal travel time change, i.e. the same change in travel time requires a change of 25% in  $\varepsilon$  but only a ~3.4% change in  $v^\perp$ .

1110 Second, about the comparison of the sensitivities of  $v$ ,  $\delta$ , and  $\varepsilon$ , precisely because we are considering the same relative increase to generate the anomalies, we can most definitely compare the sensitivities for these parameters. We are expressly working with relative changes so as to be able to compare sensitivities for parameters that have significantly different magnitude ranges. We are computing the relative change in travel times produced by proportional perturbations of each of these parameters. In other words, we see how the same proportion of change in each parameter is translated into different relative data changes, i.e. different sensitivities of the data to equal variations in each parameter.

1115

Third, in the case of  $v^\perp$ , as we just said, we now establish the comparison via the relative increase needed in  $\varepsilon$  and  $v^\perp$  to achieve the same travel time change. Indeed, in terms of the travel time change produced and following equation 2, for a given  $v$ , a certain relative change in  $v^\perp$  is equivalent to a greater relative change in  $\varepsilon$ . This is simply another way of saying that sensitivity to  $v^\perp$  is greater than to  $\varepsilon$ , i.e. to generate the same change in the travel time data, the change in  $\varepsilon$  must be of a  
 1120 greater proportion than in  $v^\perp$ . This is in agreement with the fact that in the inversion tests  $v^\perp$  is notably better recovered than  $\varepsilon$ .

Finally,  $v$  is the only parameter in equation 2 that is used in both parametrizations, and so we have to evaluate its sensitivity in both parametrizations. This involves using the background models for  $\delta$  and for  $\varepsilon$  or  $v^\perp$  depending on the case. We made  
 1125 sure that the P[ $v^\perp$ ] case does not violate the weak anisotropy assumption of  $\varepsilon$  being smaller or equal to ~0.2 in absolute value, that is, an absolute relative difference between  $v$  and  $v^\perp$  of up to approximately 20%. The sensitivity (normalized and relative) pattern for  $v$  in P[ $v^\perp$ ] is different than for P[ $\varepsilon$ ] although it follows the same sinusoidal pattern and it has equal maxima. Following the mathematical proof in the answer marked with (^) we can see that in the case of P[ $v^\perp$ ] both  $S_R$  and  $S_N$  will display a sinusoidal shape. Now  $v_A$  and  $v_B$  are

$$v_A = v_{PA} \left( 1 + \delta \sin^2 \theta \cos^2 \theta + \left( \frac{v^\perp}{v_{PA}} - 1 \right) \sin^4 \theta \right)$$

$$v_B = v_{PB} \left( 1 + \delta \sin^2 \theta \cos^2 \theta + \left( \frac{v^\perp}{v_{PB}} - 1 \right) \sin^4 \theta \right)$$

1130 and the sine and cosine functions do not cancel out in  $S_R$ , as they did in the case of an anomaly in  $v$  in P[ $\varepsilon$ ], nor in  $S_N$ . For the latter it is trivial to see that the sinusoidal dependencies are identical to the case in P[ $\varepsilon$ ].

As for  $\delta$ , we computed its sensitivity in both parametrizations and checked that its pattern is independent of the parametrization used as expected. Indeed, in  $P[\varepsilon]$

$$v_A = v_P(1 + \delta_A \sin^2 \theta \cos^2 \theta + \varepsilon \sin^4 \theta)$$

$$v_B = v_P(1 + \delta_B \sin^2 \theta \cos^2 \theta + \varepsilon \sin^4 \theta)$$

1135 and in  $P[v^\perp]$

$$v_A = v_P(1 + \delta_A \sin^2 \theta \cos^2 \theta + \left(\frac{v^\perp}{v_P} - 1\right) \sin^4 \theta)$$

$$v_B = v_P(1 + \delta_B \sin^2 \theta \cos^2 \theta + \left(\frac{v^\perp}{v_P} - 1\right) \sin^4 \theta)$$

and it is trivial to prove that these expressions are equivalent since

$$\varepsilon = \left(\frac{v^\perp}{v_P} - 1\right)$$

-Inversion tests:

The inversion tests are done properly, assuming several variants and various strategies, and results (especially concerning the  
1140 poor recovery of  $\delta$  parameter) are interesting, but their main drawback is assumption of near-ideal experimental conditions:

- the measurements geometry (spherical geometry and uniform angular spacing of sources/receivers locations, which results in unrealistically uniform ray coverage).

1145 -no noise assumed.

Such conditions are almost never possible in case of seismic in-situ experiments, where sources and receivers locations are usually limited to the earth surface, resulting in quite unfavorable ray geometry for solving inversion problem. Therefore, presented tests provide good estimate of ‘maximum capabilities’ of the code. For the case of modeling the VTI medium in  
1150 ideal conditions (which is valuable because it shows ‘weak points’ of the code and parametrization assumed – if the method fails in some aspect in ideal conditions, it will fail even more in case of real data). But such tests give no or very little information about reliability (expected resolution, dependence on the noise level, dependence on the initial model etc.) of a typical seismic experiment.

1155 I think that the manuscript would improve a lot if the authors could add more realistic synthetic tests: in order to properly check behavior of the code in case of typical data form seismic experiment, the tests should assume noisy data, surface location of sources/receivers, and also the dependence of the result on various initial models should be studied.

We are aware that the synthetic experiment that we designed to perform our tests is not realistic. In fact, our goal is to use this canonical benchmark to assess accuracy, sensitivity, and performance of the four inversion strategies under ideal and equal conditions for all anisotropic parameters, avoiding the specificities and biases of a synthetic experiment simulating a particular field case study. These tests on a canonical benchmark provide conclusions that are generally informative of the code's performance. As the referee correctly points out "presented tests provide good estimate of 'maximum capabilities' of the code which is valuable because it shows 'weak points' of the code and parametrization assumed". In other words, here we are not interested in the performance of the code in a more or less specific geological and experimental context, but in obtaining an upper limit to the code's capabilities, an ideal but generalizable estimation for the code's performance.

We have modified the manuscript to clarify the purpose of our testing approach and the reasons behind it. Specifically, we have edited the Abstract (lines 10 and 11), the third paragraph in the Introduction, the first paragraph to section 3 Synthetic tests, and the second paragraph in the Conclusions.

Nonetheless, in an upcoming paper presenting an application of the code to an anisotropic field case, we will first evaluate the code's performance on a realistic synthetic experiment simulating this particular field study. This simulation should yield an estimation of the potential quality of the results as well as information on the best modeling strategy to approach this specific case, and it will include noise, an initial model obtained from isotropic tomography, and will replicate the same exact acquisition geometry used in the field. By synthetically reproducing a more or less specific type of seismic experiment under some realistic circumstances, e.g. a certain noise level, we will get an idea of what to expect in that particular case with the selected noise level, receiver and source densities and distributions, geological features and anomalies, a priori information available in the initial models, etc. Therefore, this sort of realistic synthetic testing and the conclusions that can be drawn from it become relevant and meaningful as preliminary work linked to a particular field data application.

We believe that adding these other type of tests here would result in an exceedingly long manuscript covering too many aspects. Also, we think that realistic synthetic tests are better presented along with the field case that they are simulating. Thus, we prefer to separate our work into two publications, this first one of technical and methodological content, and a second one focusing on the field data application. Figure R1 corresponds to this anisotropic field case. In Sallarès et al. [2013] we obtained an isotropic  $V_p$  model from a refraction and wide-angle reflection seismic (WAS) data set (sub-horizontal propagation) and compared it to the image obtained from multichannel seismic (MCS) reflection data (near-vertical propagation). The top plot shows the isotropic  $V_p$  model, with the vertical coordinate converted from depth to two-way time (TWT), superimposed onto the MCS image. The white circles inside the model delineate the geometry of inter-plate reflection imaged by MCS data. The thick red line corresponds to the TWT-converted inter-plate boundary obtained from WAS data. The mismatch between the two locations of the inter-plate boundary is most likely due to some degree of seismic anisotropy between near-vertical and sub-horizontal propagations. In the bottom plot we increase the  $V_p$  values by

15%, and with this the MCS and WAS locations of the inter-plate boundary now display a good match. Thus, this 15% increase is a good initial estimate of  $\varepsilon$ , and it indicates, as a general trend, that near-vertical propagation is ~15% faster than sub-horizontal propagation in this area of the subsurface ( $\varepsilon \approx 0.15$ ). For an estimate of  $\delta$  we will use the  $V_{NMO}$  model from the normal move-out correction of MCS data processing and the isotropic  $V_p$  model.

Figs 4-9: It should be marked which column represents which parameter.

The corresponding parameter is indicated within each of the figures in the first row. We have enlarged the four symbols to make them more visible (new Figures 5 to 10).

1200

Figures

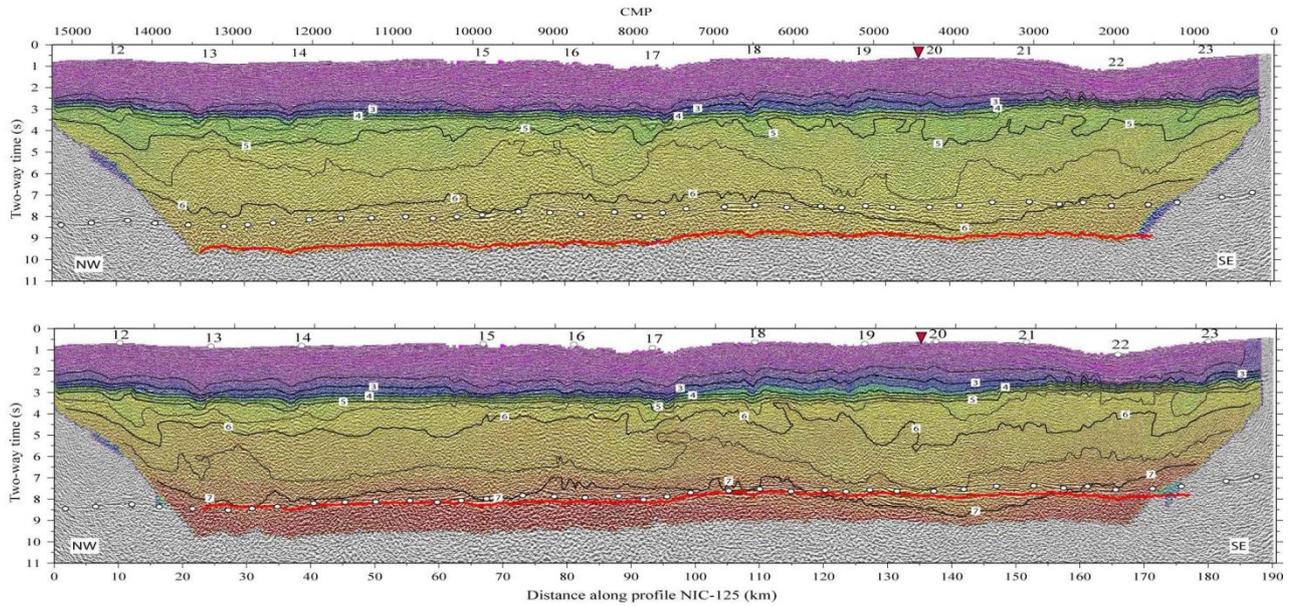
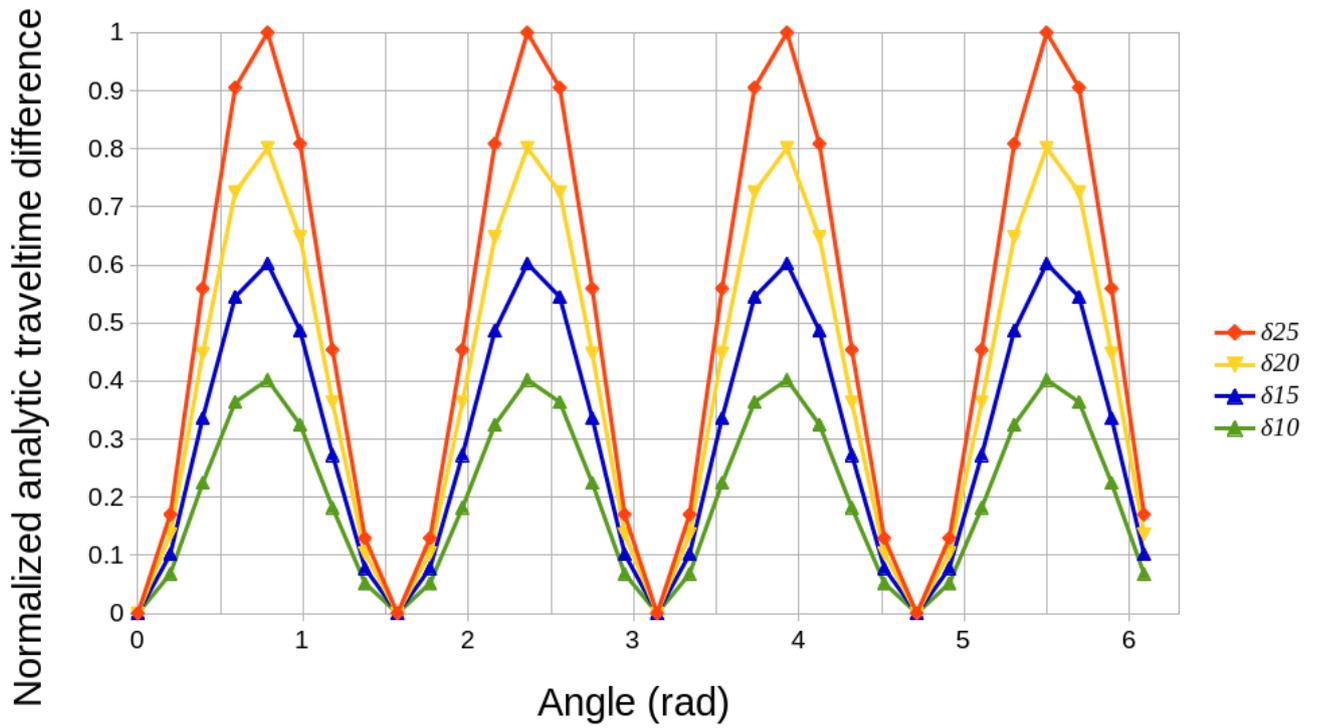


Figure R1. The top plot shows the isotropic  $V_p$  model, with the vertical coordinate converted from depth to two-way time (TWT), superimposed onto the MCS image. The white circles inside the model delineate the geometry of inter-plate reflection imaged by MCS data. The thick red line corresponds to the TWT-converted inter-plate boundary obtained from WAS data. The mismatch between the two locations of the inter-plate boundary is most likely due to some degree of seismic anisotropy between near-vertical and sub-horizontal propagations. In the bottom plot we increase the  $V_p$  values by 15%, and with this the MCS and WAS locations of the inter-plate boundary now display a good match. Ocean bottom receivers are numbered, and the red triangle represents the intersection point with a perpendicular WAS profile.



1205 **Figure R2:** Normalized analytic  $\delta$  sensitivity for anomalies of 10%, 15%, 20%, and 25%. Sensitivity and anomaly increment display a linear relationship that is particularly clear at the maxima of the curves.