

# Comparing one-step full-spectrum inversion with two-step splitting function inversion in normal mode tomography

Lisanne Jagt and Arwen Deuss

*Department of Earth Sciences, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, the Netherlands. E-mail: [e.a.a.m.jagt@uu.nl](mailto:e.a.a.m.jagt@uu.nl)*

Accepted 2021 June 21. Received 2021 June 15; in original form 2021 January 8

## SUMMARY

Earth's normal modes, or whole Earth oscillations, provide important constraints on Earth's large-scale 3-D structure. In addition to constraining shear and compressional wave velocities, they are the only seismic data sensitive to density perturbations. Density is particularly difficult to determine, and previous studies have found contradicting results, hence the method chosen to invert normal mode data for 3-D structure becomes important. In the problem of inverting the measured frequency spectra for an earth model, we can take two approaches: (i) a one-step full-spectrum inversion, where normal mode spectra are directly inverted for a mantle model and (ii) a two-step splitting function inversion, where first the spectra are inverted for splitting functions, which are then inverted for a mantle model. Here we compare the methodology and results of both approaches, continuing the work done by Li *et al.* and Durek & Romanowicz, and extending it to higher spherical harmonic degrees. Using exactly the same normal mode data set, we use both inversion approaches to make 3-D shear wave velocity mantle models. Both approaches give models consistent with previous tomographic studies, although spectral misfits are consistently lower for the one-step full-spectrum inversion. We also show that we cannot draw any conclusions on odd-degree structure in the lower mantle with the currently available normal mode data sets.

**Key words:** Structure of the Earth; Inverse theory; Seismic tomography; Surface waves and free oscillations.

## 1 INTRODUCTION

Normal modes, or free oscillations, are whole Earth oscillations excited by large earthquakes (typically  $M_w > 7.5$ ). These standing waves along the Earth's surface and radius make the Earth 'ring like a bell'. Seismic normal modes are long-period waves, typically studied in the frequency domain between 0 and 10 mHz. They constrain large-scale Earth structures, and have been used to make global tomographic models from crust to inner core (e.g. Li *et al.* 1991; Durek & Romanowicz 1999; Ishii & Tromp 1999, 2004; Resovsky & Ritzwoller 1999; Beghein & Trampert 2003; Resovsky & Trampert 2003), often in combination with body wave and surface wave data (e.g. Ritsema *et al.* 1999, 2011; Koelemeijer *et al.* 2016; Durand *et al.* 2016). In addition to providing information on shear wave velocity [ $\delta \ln(V_s)$ ] and compressional wave velocity [ $\delta \ln(V_p)$ ] anomalies, modes are sensitive to variations in density ( $\delta \ln \rho$ , Ishii & Tromp 1999; Resovsky & Trampert 2003). It has proven difficult to determine 3-D variations in density due to trade-offs with other parameters such as core-mantle boundary (CMB) topography and  $V_s$ , and their strong dependence on *a priori* constraints (Resovsky & Ritzwoller 1999; Romanowicz 2001; Kuo & Romanowicz 2002). Yet information on density carries

important implications for mantle dynamics, since density perturbations drive mantle flow (e.g. Davies & Gurnis 1986; Kellogg *et al.* 1999).

Of special interest are the two large low shear-wave velocity provinces (LLSVPs) in the lower mantle beneath the Pacific Ocean and Africa. Ever since their first appearance in mantle tomographic models as anomalously low  $V_s$  regions with a prominent degree 2 signature (Dziewonski *et al.* 1977; Giardini *et al.* 1987), their nature has been controversial. If they have a low density, then they are likely to have a thermal origin. Alternatively, if they are found to have a higher density, they need to have a compositional origin. To answer this question, relative amplitudes of density with respect to seismic velocities as well as the sign of density anomalies have been studied extensively, but nevertheless are still heavily debated (Romanowicz 2001; Resovsky & Trampert 2003). Although most normal mode studies argue for a higher than average density (Ishii & Tromp 1999; Resovsky & Trampert 2003; Ishii & Tromp 2004; Trampert *et al.* 2004; Lau *et al.* 2017), observations of light LLSVPs exist (Koelemeijer *et al.* 2017). The question of 3-D density variations in the mantle thus remains unanswered.

Modes appear as distinct peaks in frequency spectra by Fourier transforming several days to week-long seismograms from the time

to the frequency domain. The shape and position of normal mode peaks change due to departures of Earth from a spherical, non-rotating, elastic, isotropic (SNREI) Earth. Measurements of spectral peaks are data input for inversion procedures aiming to image the Earth's interior. In principle, two different inversion approaches can be used for obtaining an earth model from normal mode data, either (1) in one step, by directly inverting the spectra or (2) in two steps, via the splitting function method (see Fig. 1).

The one-step full-spectrum inversion involves a direct inversion of normal mode spectra for an earth model in terms of velocity and/or density perturbations. This approach was first used by Li *et al.* (1991) and has only been used a few times in the last decades (Li *et al.* 1991; Durek & Romanowicz 1999; Kuo & Romanowicz 2002). The problem is that it is a non-linear method involving diagonalization of large matrices and requires large computer power.

The two-step splitting function inversion was introduced by Giardini *et al.* (1987) to separate the inversion into two smaller steps which require less computer power. As the name suggests, it involves first inverting the normal mode spectra for splitting function coefficients, also called structure coefficients (e.g. Resovsky & Ritzwoller 1998; Deuss *et al.* 2013), as an intermediate step. Splitting functions describe the weighted depth average of lateral heterogeneities that a mode 'sees', and are similar to phase velocity maps for surface waves. Splitting functions are linearly dependent on 3-D variations in Earth structure. They can be visualized as splitting function maps showing where locally the frequency of a given mode is slightly lower or higher than its average frequency. The splitting function coefficients are then inverted in the second step for variations in velocity and density. Splitting functions are most commonly used when normal mode constraints are included in global tomographic models, often in combination with other data such as body waves and surface waves, for example in constructing mantle models S40RTS (Ritsema *et al.* 2011), SP12RTS (Koelemeijer *et al.* 2016) and SEISGLOB1 (Durand *et al.* 2016). Generally, measuring splitting functions and performing the two-step splitting function inversion are more common in normal mode studies (e.g. Giardini *et al.* 1987; He & Tromp 1996; Resovsky & Ritzwoller 1998, 1999; Masters *et al.* 2000; Deuss *et al.* 2013; Pachhai *et al.* 2016; Koelemeijer *et al.* 2017), than the one-step direct spectrum inversion.

So far, only Li *et al.* (1991), Kuo & Romanowicz (2002) and Durek & Romanowicz (1999) have attempted to directly invert the spectra for a model. A splitting function inversion may have multiple local, or sometimes global, minima (Megnin & Romanowicz 1995). The problem may be especially relevant for core-sensitive modes and therefore Durek & Romanowicz (1999) performed a one-step full-spectrum inversion to make an anisotropic model of the inner core to circumvent the problem of non-uniqueness in measuring splitting functions. We will discuss the theoretical differences between the two inversion methods in more detail in Section 2.4. Splitting functions predicted for their preferred model matched the measured splitting functions retrieved from spectral data well. The one exception ( ${}_{13}S_1$ ) was attributed to non-uniqueness of its measured splitting function. Here, we will investigate the difference between the two methods for mantle sensitive modes. Kuo & Romanowicz (2002) inverted synthetic spectra directly for  $V_s$ ,  $V_p$  and  $\rho$  simultaneously and looked at whether 3-D variations in mantle density could be resolved with then available normal mode data set. They found that density structure obtained from their normal mode data set was not reliable, due to contamination of seismic velocity structure into density.

Li *et al.* (1991) compared the one-step full-spectrum inversion and two-step splitting function inversion methods by making tomographic models based on the same normal mode data set. Their models included  $V_s$  mantle structure, CMB topography and inner core anisotropy. Resulting mantle models from the two approaches were in good agreement, which prompted the authors to conclude that splitting functions can serve as a convenient intermediate step in obtaining earth models from normal mode splitting. Indeed, the two-step splitting function inversion method has been the preferred choice in the last decades to include normal modes in tomographic models.

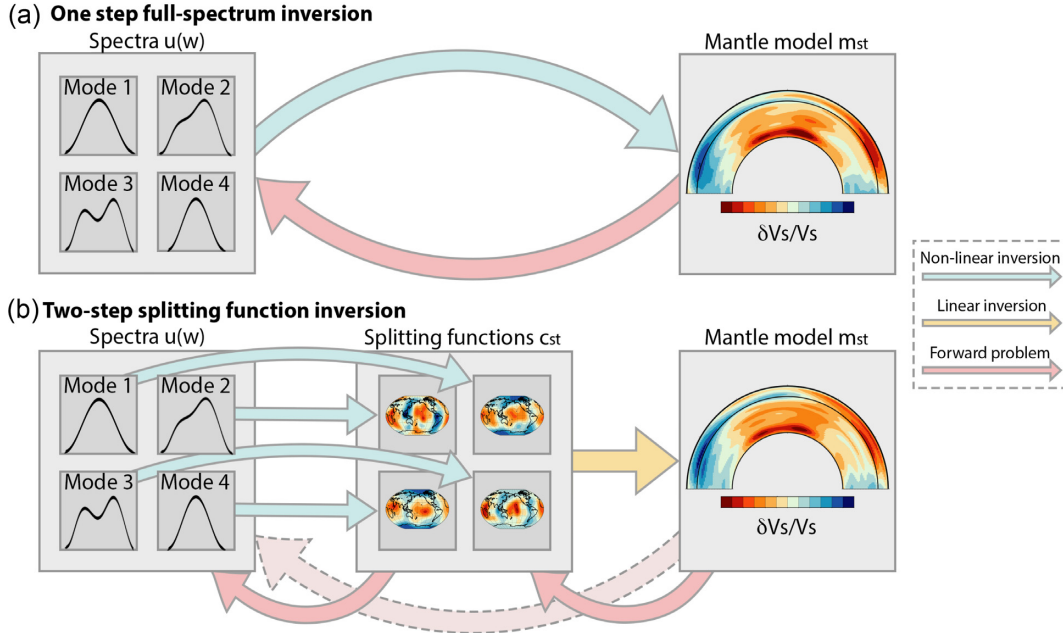
It is important to revisit this question again in light of recent theoretical developments, the expansion of data and increase in computer power. For example, Li *et al.* (1991) used the self-coupling approximation, meaning that they treated each mode in isolation. It has recently been suggested that self-coupling may be a reasonable approximation for modes below 1.5 mHz that are isolated in the frequency spectrum (Yang & Tromp 2015), but all modes exchange varying amounts of energy with other modes, especially when their frequencies are close. Deuss & Woodhouse (2001) showed that this so-called cross-coupling (i.e. resonance) in large groups of modes needs to be taken into account to accurately calculate synthetic spectra, even below 1.5 mHz (Akbarashrafi *et al.* 2018). So, the question is what happens when we take this cross-coupling between modes into account in the one-step full-spectrum inversion and two-step splitting function inversion. Here, we will extend the work done by Li *et al.* (1991), using the large normal mode data set initially compiled by Deuss *et al.* (2013) to make a new catalogue of splitting function measurements. Owing to greater computational power, we will also be able to go to higher spherical harmonic degrees, significantly increase the amount of data, and test the influence of cross-coupling.

We will make a shear wave velocity model using both the one-step full-spectrum inversion and the two-step splitting function inversion methods and investigate the effect of the two methods on the misfits and resulting models.

## 2 THEORETICAL FRAMEWORK

We will use normal modes to make tomographic models of Earth's mantle. There are two types of normal modes: spheroidal modes, which involve  $P$ -SV motion and are similar to Rayleigh waves, and toroidal modes, which involve SH-motion and are similar to Love waves. Normal modes are characterized by their overtone number  $n$  and angular order  $l$ :  ${}_nS_l$  for spheroidals and  ${}_nT_l$  for toroidals. A mode with a given  $l$  consists of  $2l + 1$  singlets, or eigenfunctions, which are labelled with azimuthal order  $m$ . In a SNREI Earth, all  $2l + 1$  singlets of a mode will be degenerate, i.e. they all have the same eigenfrequency  $\omega_0$ . Rotation, ellipticity and 3-D structure of the Earth remove the degeneracy and lead to spectral splitting of singlets, giving each singlet a different frequency. Contributions of rotation and ellipticity are known exactly and only depend on the 1-D earth model (Woodhouse & Dahlen 1978), so we can invert the observed remaining spectral splitting of normal modes for the contribution due to 3-D heterogeneity in  $V_s$ ,  $V_p$  and  $\rho$ . We will use both the one-step full-spectrum inversion approach and the two-step approach with the intermediate splitting functions estimation. A detailed explanation for both approaches is given in the sections later.

An overview of the main steps and the major differences between the two inversion schemes is outlined in Fig. 1. The one-step



**Figure 1.** Inversion procedure for (a) one-step full-spectrum inversion with only one non-linear inversion (blue arrow) inverting all spectra for all modes simultaneously for a mantle model; (b) two-step splitting function inversion with non-linear inversions (blue arrows) to obtain splitting functions from spectra for all modes separately; and then inverting all those splitting functions simultaneously for a mantle model in a linear inversion (yellow arrow). All of the forward (synthetic) calculations are represented by red arrows.

full-spectrum inversion (Fig. 1a) has only one inversion step (blue arrow), in which we invert the normal mode spectra non-linearly for a mantle model. The fit to the original spectra is determined by computing synthetic spectra for the resulting mantle model (red arrow). The two-step splitting function inversion (Fig. 1b) first involves inverting the normal mode spectra for splitting functions for each mode separately (blue arrows), which is a non-linear process. Secondly, the splitting functions are linearly inverted simultaneously for a mantle model (yellow arrow). Different data fits are determined for the resulting mantle model, both to the measured splitting functions (red arrow) but also to the original spectra by computing synthetic spectra (light red arrow).

## 2.1 Forward problem: synthetic spectra

In both inversion approaches we need to calculate synthetic spectra for given mantle model coefficients  $m_{st}$  or splitting function coefficients  $c_{st}$  (see red arrows in Fig. 1), in order to minimize the spectral misfit between data and synthetics. We use the ‘Coriolis and kinetic energy approximation’ explained in Deuss & Woodhouse (2001) to compute synthetics. Synthetic normal mode seismograms in the time domain for a given Earth structure are harmonic functions of time  $t$  given by:

$$u(t) = \text{Re} \left[ \mathbf{r} \cdot \mathbf{e}^{i\sqrt{\mathbf{M}}t} \cdot \mathbf{s} \right], \quad (1)$$

where  $u(t)$  is the synthetic seismogram,  $\mathbf{s}$  is the source vector depending on the moment tensor,  $\mathbf{r}$  is the receiver vector depending on the instrument location and orientation, and  $\mathbf{M}$  is the splitting matrix which includes contributions due to 3-D variations in Earth structure, ellipticity and rotation (Woodhouse 1980), either in terms of the model coefficients  $m_{st}$  or splitting function coefficients  $c_{st}$ . We obtain the moment tensors from the Global CMT catalogue (Ekström *et al.* 2012) and treat the source equally in both inversion methods. Hence we will not discuss the effects of uncertainties and

errors in the source parameters on the  $\delta \ln(V_s)$  models. In the presence of ellipticity, rotation and asphericities,  $\mathbf{M}$  is not diagonal, so we need to diagonalize this matrix to compute the exponential. This is done by eigenvalue decomposition:  $\mathbf{M}\mathbf{U} = \mathbf{U}\mathbf{\Lambda}$ , where  $\mathbf{U}$  contains the eigenvectors and  $\mathbf{\Lambda}$  contains the eigenvalues on the diagonal. The synthetic seismogram (eq. 1) then becomes:

$$u(t) = \text{Re} \left[ (\mathbf{r} \cdot \mathbf{U}) \mathbf{e}^{i\sqrt{\mathbf{\Lambda}}t} (\mathbf{U}^{-1} \cdot \mathbf{s}) \right]. \quad (2)$$

The splitting matrix  $\mathbf{M}$  contains the degenerate frequencies  $\omega_0^2$  on the diagonal and additional diagonal and non-diagonal contributions from ellipticity, rotation, 3-D heterogeneity in  $V_s$ ,  $V_p$ ,  $\rho$  and anisotropy. For a mode pair  $k, k'$  with degenerate frequencies  $\omega_k, \omega_{k'}$  in the reference model, we can write  $\mathbf{M}$  as:

$$\mathbf{M}_{mm'}^{(kk')} = \omega_0^2 \delta_{kk'} + \omega_0 \mathbf{W}_{mm'}^{(kk')} + \mathbf{H}_{mm'}^{(kk')}, \quad (3)$$

where  $s$  is the angular order or structural degree,  $t$  is azimuthal order,  $\omega_0 = (\omega_k + \omega_{k'})/2$ , matrix  $\mathbf{W}$  describes the Coriolis force due to rotation, and matrix  $\mathbf{H}$  describes ellipticity, 3-D heterogeneity in  $V_s$ ,  $V_p$  and  $\rho$  and anisotropy. Please note that the dimension of the splitting matrix is  $\omega^2$  (Woodhouse 1980), hence the square root in eq. (1). In the self coupling case,  $\mathbf{M}$  is a square matrix of size  $(2l+1)(2l+1)$  and  $k = k'$  so  $\delta_{kk'} = 1$ , and in the absence of 3-D structure, the splitting matrix is diagonal. The cross coupling is described by matrix contributions of size  $(2l+1)(2l'+1)$  and  $k \neq k'$  so  $\delta_{kk'} = 0$ .

For modes that are sufficiently isolated in the spectrum, we use the self-coupling approximation, meaning that only singlets belonging to the same mode exchange energy. We use the group-coupling approximation for small groups of 2–3 modes that interact strongly. The modes that we allowed to couple varies for each method. It is an advantage of the one-step inversion that we can easily include coupling between all modes in a particular spectral segment. In the two-step inversion, this is more complicated, since we cannot always include all cross-coupling, since that would result in too many unknown model parameters compared to the limited data

availability per mode. It has been shown that full-coupling would probably be a more accurate approach (Deuss & Woodhouse 2001; Al-Attar *et al.* 2012; Yang & Tromp 2015), but the aim of the present paper is to compare  $V_s$  models obtained using either direct spectrum inversion or splitting functions, and therefore we need to use the same theoretical approximations in both. In addition, full-coupling is computationally very expensive and will be the subject of a further paper.

### 2.1.1 One-step full-spectrum inversion

The two inversion methods differ in the calculation of  $\mathbf{H}$ . In the one-step full-spectrum method the matrix  $\mathbf{H}$  depends on the model coefficients  $m_{st}$ . The splitting matrix elements are then computed using the ‘Woodhouse kernels’ in eq. (A.17) of Woodhouse (1980), which can also be found in appendix D of Dahlen & Tromp (1998). Matrix  $\mathbf{H}$  is in this case a function of  $m_{st}$ , which include 3-D perturbations in  $V_s$ ,  $V_p$ ,  $\rho$ , anisotropy and discontinuity topography:

$$\mathbf{H}_{mm'}^{(kk')} = \sum_{s=-l'}^{l+l'} \sum_{t=-s}^s \gamma_{ll's}^{mm't} \left[ \int_0^a \delta m_{st}(r) K_{s(kk')}(r) dr + \sum_d \delta h_{st} H_{s(kk')}^d \right], \quad (4)$$

where the coefficients  $\gamma_{ll's}^{mm't}$  are given by:

$$\gamma_{ll's}^{mm't} = \int_0^{2\pi} \int_0^\pi Y_l^{m*}(\theta, \phi) Y_s^t(\theta, \phi) Y_l^{m'}(\theta, \phi) \sin \theta d\theta d\phi, \quad (5)$$

with  $Y_l^m$  as the fully normalized spherical harmonics according to Edmonds (1960). Woodhouse (1980) and Dahlen & Tromp (1998) describe how to evaluate this integral.  $\delta m_{st}$  and  $\delta h_{st}$  describe Earth’s elastic heterogeneity ( $\delta \ln(V_s)$ ,  $\delta \ln(V_p)$ ,  $\delta \ln \rho$ ) and discontinuity topography, respectively, parametrized in spherical harmonics with angular order  $s$  and azimuthal order  $t$ .  $K_s(r)$  and  $H_s^d$  are known kernels (Woodhouse 1980).

### 2.1.2 Two-step splitting function inversion

In the two-step splitting function inversion, the matrix  $\mathbf{H}$  depends on the splitting function coefficients  $c_{st}$  instead. The heterogeneity matrix in  $\mathbf{H}$  can then be readily described using splitting function coefficients  $c_{st}$  as:

$$\mathbf{H}_{mm'}^{(kk')} = \sum_{s=-l'}^{l+l'} \sum_{t=-s}^s \gamma_{ll's}^{mm't} c_{st(kk')}. \quad (6)$$

It is interesting to note that the splitting matrix now only depends on  $c_{st}$  values and these can be estimated without any knowledge of the underlying 3-D heterogeneity. This is what is being used in the first step of the splitting function inversion. Only in the second step we use the fact that the splitting function coefficients  $c_{st}$  do linearly depend on 3-D heterogeneity and are given by:

$$c_{st(kk')} = \int_0^a \delta m_{st}(r) K_{s(kk')}(r) dr + \sum_d \delta h_{st} H_{s(kk')}^d. \quad (7)$$

## 2.2 Partial derivatives

Synthetic seismograms  $u(t)$  and the corresponding spectra are non-linearly dependent on either (1) the model parameters  $m_{st}$  in the one-step inversion or (2) the splitting function coefficients  $c_{st}$  in

the two-step inversion. Therefore we need to linearize these two problems using partial derivatives.

### 2.2.1 One-step full-spectrum inversion

In the one-step inversion we need to calculate derivatives of the synthetic spectra with respect to the model parameters  $m_{st}$  (see blue arrow in Fig. 1a). The partial derivatives of seismogram  $u(t)$  (see eq. 2) with respect to model parameters  $m_{st}$  are given by:

$$\frac{\partial u(t)}{\partial m_{st}} = \text{Re} \left[ e^{i\omega t} \left( \mathbf{r} \cdot \frac{\partial \mathbf{U}}{\partial m_{st}} \right) (\mathbf{U}^{-1} \cdot \mathbf{s}) + (\mathbf{r} \cdot \mathbf{U}) \left( \frac{\partial \mathbf{U}^{-1}}{\partial m_{st}} \cdot \mathbf{s} \right) + (\mathbf{r} \cdot \mathbf{U}) i t \frac{\partial \omega}{\partial m_{st}} (\mathbf{U}^{-1} \cdot \mathbf{s}) \right], \quad (8)$$

where  $\omega$  is diagonal matrix  $\sqrt{\Lambda}$ . We use Rayleigh’s principle to find perturbations in the eigenvectors  $\mathbf{U}$  and eigenvalues  $\Lambda$  of splitting matrix  $\mathbf{M}$  resulting from a perturbation  $\delta m_{st}$  in the model coefficients. The eigenvalue correction is given by:

$$\delta \omega_n^2 = \delta \Lambda = \mathbf{u}_n^{-1} \delta \mathbf{M} \mathbf{u}_n, \quad (9)$$

where  $\mathbf{u}_n$  is a column vector of  $\mathbf{U}$  and  $\mathbf{u}_n^{-1}$  is a row vector of  $\mathbf{U}^{-1}$ . The eigenvector corrections are given by:

$$\delta \mathbf{u}_n = \sum_{l \neq n} \frac{\mathbf{u}_l^{-1} \delta \mathbf{M} \mathbf{u}_n}{\omega_n^2 - \omega_l^2} \mathbf{u}_l \quad (10)$$

$$\delta \mathbf{u}_n^{-1} = \sum_{l \neq n} \frac{\mathbf{u}_n^{-1} \delta \mathbf{M} \mathbf{u}_l}{\omega_n^2 - \omega_l^2} \mathbf{u}_l^{-1}, \quad (11)$$

where  $\omega_n$  are the diagonal elements of  $\omega = \sqrt{\Lambda}$ . These eigenvalue and eigenvector corrections are substituted in eq. (8) for  $\frac{\partial \omega}{\partial m_{st}}$ ,  $\frac{\partial \mathbf{U}}{\partial m_{st}}$  and  $\frac{\partial \mathbf{U}^{-1}}{\partial m_{st}}$  to evaluate the derivatives. Here, we will only calculate derivatives with respect to  $V_s$ , so  $m_{st}$  only contains  $\delta \ln(V_s)$ . If we extend this to  $V_p$  and  $\rho$  then  $m_{st}$  contains three sets of independent parameters for which derivatives need to be calculated.

### 2.2.2 Two-step splitting function inversion

**First step - splitting function measurement** In the first step of the splitting function inversion method, we need to calculate derivatives of the synthetic spectra with respect to the splitting function coefficients  $c_{st}$  (see blue arrows in Fig. 1b). We can easily obtain the equation for the partial derivatives by replacing every instance of  $m_{st}$  in eq. (8) by  $c_{st}$ . These derivatives are then used in the first step of estimating the splitting function coefficients in the two-step inversion method. This equation is also given in Deuss *et al.* (2013) as eq. (12). These are then the derivatives needed to measure the splitting function coefficients from the normal mode spectra in the first step of the splitting function inversion. We will not do this first step in this paper, because it was already done by Deuss *et al.* (2013) and we will use their measured  $c_{st}$ ’s for the second step.

**Second step - splitting function inversion** We then need to calculate the derivatives of the splitting function coefficients  $c_{st}$  with respect to the model parameters  $m_{st}$  (see yellow arrow in Fig. 1b). Because it is a linear problem, the partial derivatives are the same as what we use to do the forward calculation (eq. 7):

$$\frac{\partial c_{st}}{\partial m_{st}} = K_s(r), \quad (12)$$

where  $K_s(r)$  appear once again as known kernels (Woodhouse 1980). This is the step we will perform in this paper, represented by the yellow arrow in Fig. 1(b).



### 2.3 Inversion method

We use the iterative least squares approach of Tarantola & Valette (1982) to estimate the model ( $m_{st}$ ) or splitting function coefficients ( $c_{st}$ ) from the normal mode spectra (see blue arrows in Fig. 1). In both approaches we minimize the same objective function, but the model parameters are  $m_{st} [\delta \ln(V_s)]$  in the one-step inversion (Section 2.3.1). The model parameters are splitting function coefficients  $c_{st}$  in the first step (blue arrow in Fig. 1b) and  $m_{st} [\delta \ln(V_s)]$  in the second step of the two-step inversion (Section 2.3.2, yellow arrow in Fig. 1b). We will compare the two inversion approaches based on their fit to the original spectral data, shape of their L-curves, robustness to choices in weighting and similarity of the resulting models.

#### 2.3.1 One-step full-spectrum inversion

The total objective function  $\Phi$  that we try to minimize in both methods is:

$$\Phi(\mathbf{m}) = \frac{1}{2} [\mathbf{d} - \mathbf{u}(\mathbf{m})]^T \mathbf{C}_d^{-1} [\mathbf{d} - \mathbf{u}(\mathbf{m})] + \frac{1}{2} [\mathbf{m} - \mathbf{m}_0]^T \mathbf{C}_m^{-1} [\mathbf{m} - \mathbf{m}_0] \quad (13)$$

in which  $\mathbf{d}$  is observed data,  $\mathbf{u}(\mathbf{m})$  is synthetic data given a model  $\mathbf{m}$ ,  $\mathbf{m}_0$  is the starting model, and  $\mathbf{C}_d^{-1}$  and  $\mathbf{C}_m^{-1}$  are the *a priori* data and model covariance matrices, respectively. In the one-step method the model  $\mathbf{m}$  contains the  $m_{st}$  model parameters for  $\delta \ln(V_s)$ . Function  $\Phi$  minimizes both misfit to the data and size of the model parameters in a least-squares sense (Tarantola 2005).

The spectra  $\mathbf{u}(\mathbf{m})$  depend non-linearly on the model parameters  $\mathbf{m}$ , so we need to iterate the inversion. The model that minimizes this objective function is obtained through the Gauss–Newton (or quasi-Newton) method, a gradient-based approach (Tarantola & Valette 1982; Pratt *et al.* 1998). The final expression for the iterative damped least-squares inversion is given by:

$$\mathbf{m}^{i+1} = \mathbf{m}^i + (\mathbf{A}_i^T \mathbf{C}_d^{-1} \mathbf{A}_i + \mathbf{C}_m^{-1})^{-1} \times [\mathbf{A}_i^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{u}(\mathbf{m}^i)) - \mathbf{C}_m^{-1} (\mathbf{m}^i - \mathbf{m}_0)], \quad (14)$$

where  $\mathbf{A} = [\frac{\partial \mathbf{u}}{\partial \mathbf{m}}]$  is the matrix of partial derivatives (eq. 8),  $\mathbf{m}^0$  is the starting model,  $\mathbf{m}^i$  the vector containing the model coefficients  $m_{st}$  for  $\delta \ln(V_s)$  for the previous iteration and  $\mathbf{m}^{i+1}$  the updated model coefficients for the current iteration,  $\mathbf{d}$  the measured spectra, and  $\mathbf{u}(\mathbf{m}^i)$  the synthetic spectra (eq. 2).

#### 2.3.2 Two-step splitting function inversion

**First step** The first step of the splitting function method is very similar to the one-step method, but instead of inverting for  $m_{st}$  we now invert for splitting function coefficients  $c_{st}$ . The main difference is that we estimate the  $c_{st}$ 's for each mode separately, so we run a non-linear iterative inversion for every mode. Because the same inverse theory applies, we obtain the non-linear splitting function inversion from the spectra by replacing  $\mathbf{m}$  by  $\mathbf{c}$  in eqs (13) and (14) to obtain splitting function coefficients in the iterative inversion of this first step.

**Second step** Once the  $c_{st}$ 's have been measured for all the different modes, we then linearly invert the  $c_{st}$ 's for model parameters  $m_{st}$  in the second inversion step. The  $c_{st}$  are linearly related to the  $m_{st}$  through known kernels (eq. 7). A collection of splitting function coefficients for a large number of modes is inverted

jointly for a model of 3-D elastic structure, using the fact that each mode has a different depth sensitivity. This is now a linear inversion. We solve for the model coefficients in a damped least squares inversion:

$$\mathbf{m} = \mathbf{m}^0 + (\mathbf{A}^T \mathbf{C}_d^{-1} \mathbf{A} + \mathbf{C}_m^{-1})^{-1} \mathbf{A}^T \mathbf{C}_d^{-1} (\mathbf{c} - \mathbf{c}(\mathbf{m}^0)) \quad (15)$$

where  $\mathbf{A} = [\frac{\partial \mathbf{c}}{\partial \mathbf{m}}]$  is the matrix of partial derivatives. Comparing eq. (15) for the second step of the splitting function inversion with eq. (14), we see the advantage of eq. (15): it does not have to be iterated because it is a linear inversion.

#### 2.3.3 Choices in weighting

The choice of weighting the data is very important in any inversion scheme and is included by assigning varying values to the *a priori* data covariance matrix  $\mathbf{C}_d^{-1}$ . We found that weighting in the one-step direct spectrum inversion is done automatically, since modes with more data have a larger weight by virtue of contributing more data. Hence those modes are more dominant in determining the updated model. Weighting in the one-step inversion is therefore implicit (i.e. no variations in  $\mathbf{C}_d^{-1}$  required for the different modes). The first step in the two-step splitting function inversion does not require variations in  $\mathbf{C}_d^{-1}$  either, because only one mode (or a pair) is measured at a time (see Deuss *et al.* 2013). However, the second step in the two-step splitting function inversion depends much more strongly on weighting of the different splitting function coefficients. We will apply weighting by using the errors assigned to the coefficients in the first step of the splitting function inversion. Other weighting schemes have also been used. Li *et al.* (1991), for example, included all splitting function coefficients with equal weight. They found that their splitting function coefficient errors for modes with little splitting were unrealistically small, and decided to assume the same standard error for all splitting function coefficients. We require non-uniform weighting of the splitting function coefficients to get reasonable models. We decided to use the errors  $\sigma_d$  estimated by Deuss *et al.* (2013) on the diagonal of the data covariance matrix  $\mathbf{C}_d^{-1}$  as  $1/\sigma^2$  with  $\sigma = \sqrt{\sigma_d}$ , leading to a weighting range from 0.15 to 51. Using  $\sigma = \sigma_d$  instead would give a much larger range of weighting coefficients: 0.02–2646, making modes with the smallest errors too dominant.

A spectral segment is defined by a time and frequency window in the spectrum corresponding to a single seismogram. Modes with more segments generally have lower error estimates of their splitting function coefficients, so the error-based weighting of the two-step splitting function inversion is comparable to the implicit weighting by data amount in the one-step direct spectrum inversion. Contrary to the procedure outlined so far, we could choose to weigh each mode in the one-step inversion equally. The expected outcome is that modes with less segments will get a lower spectral misfit in the equal weights procedure. This is indeed what we see. The resulting models are still highly similar, with a correlation coefficient of 0.96. This indicates that the one-step inversion is robust to changes in the weighting scheme. The two-step inversion is more dependent on the weighting scheme, since treating all modes as equally important in the second inversion of the two-step inversion does not give a  $V_s$  model that looks similar to models in previous studies. Applying equal weights worked for Li *et al.* (1991), but in our case it resulted in very large  $c_{st}$  misfits. Furthermore, the models did not look like  $\delta \ln(V_s)$  obtained in previous studies, possibly due to the higher maximum degree of the model and/or including more data.

## 2.4 Misfit definition

For the two-step inversion we can compute two kinds of misfit for a model:  $c_{st}$  misfit and spectral misfit (red arrows in Fig. 1b). For the one-step inversion we only compute the spectral misfit (red arrow in Fig. 1a).

### 2.4.1 One-step direct spectrum inversion

The spectral misfit per mode segment is defined as:

$$\text{spectral misfit} = \frac{1}{N} \sum \frac{\sum_{i=1}^N (\mathbf{d}_i - \mathbf{u}_i(m_{st}))^2}{\sum_{i=1}^N (\mathbf{d}_i)^2} \quad (16)$$

which is the squared difference between  $n$  spectral data points in observed spectra  $\mathbf{d}_i$  and synthetic spectra  $\mathbf{u}_i(m_{st})$ , divided by the norm of the data, averaged over the number of spectral segments  $N$ . The total average spectral misfit of all modes combined is calculated by weighing each spectral segment misfit by the number of segments per mode.

### 2.4.2 Two-step splitting function inversion

The first step in the two-step method was already performed by Deuss *et al.* (2013), using the same spectral misfit definition as in eq. (16), with  $m_{st}$  replaced by  $c_{st}$ . Here we performed the second inversion step of fitting the splitting functions measured by Deuss *et al.* (2013). The  $c_{st}$  misfit of splitting function coefficients (i.e. the misfit that is minimized in the second step of the splitting function inversion) is defined as:

$$c_{st} \text{ misfit} = \frac{1}{n_{tot}} \sum n \cdot \frac{\sum_{s,t} |c_{st}^{obs} - c_{st}^{mod}|}{\sum_{s,t} |c_{st}^{obs}|}, \quad (17)$$

where  $c_{st}^{obs}$  are the measured splitting function coefficients,  $c_{st}^{mod}$  are the predicted coefficients for the new model,  $n$  is the number of  $c_{st}$ 's per mode,  $N$  the total number of modes and  $n_{tot}$  the total number of  $c_{st}$ 's. This begs the question of whether our final model still fits the original spectral data from which those splitting functions were derived in the first step (dashed red arrow in Fig. 1b)? We have a unique opportunity to compare the two inversion methods, since we use the same initial spectral data set and can therefore calculate the spectral data fit of all resulting models. Another important question for the two-step inversion is: would the model that we choose based on the splitting function misfits be the same model that we choose based on spectral misfits? Hence, we also use eq. (16) to calculate the spectral misfit for the model obtained in the two-step splitting function inversion.

## 2.5 Theoretical comparison of the two methods

Based on the overview of the two inversion schemes in previous sections, we now list the advantages and disadvantages of both methods. The advantage of the one-step method is that all data is fitted with the best model. The disadvantage is that it is computationally intensive, amounting to approximately 4–5 d on a High-Performance Computing cluster to reach convergence. The advantage of the two-step inversion is its lower computational cost, since the problem is decoupled into a number of smaller problems. The first step of the two-step inversion takes up to a few hours until reaching convergence, and the second step only a few minutes. The decoupling also leads to splitting functions in the first inversion step being obtained for each mode or mode group separately. These splitting functions

are therefore not required to be consistent with a single earth model (e.g. Kuo & Romanowicz 2002). If a measured splitting function does not represent the unique solution corresponding to true properties of the Earth, the error will propagate into the final model (Li *et al.* 1991). Weighting in the second inversion step is necessary to make measurements with large uncertainties less important, otherwise erroneous splitting function measurements will dominate the model. Furthermore, regularization is applied twice in the two-step inversion, hence information is lost twice.

What are the practical implications of these theoretical differences? That is what we will explore in the rest of this paper.

## 3 DATA

We use normal mode spectra as starting point of our inversions. These spectra are obtained by taking vertical component seismic data of tens of hours length, removing the first few hours since the earthquake onset, and then transforming them to the frequency domain. For the one-step inversion (blue arrow in Fig. 1a) we use the spectral segments of Deuss *et al.* (2013), consisting of vertical component data for 91 large events since 1976 with  $M_w \geq 7.4$ . For the two-step approach we use the spheroidal mode splitting functions obtained by Deuss *et al.* (2013), and only execute the second inversion step in this paper (yellow arrow in Fig. 1b). So, both inversion schemes use exactly the same spectral data. We exclude spheroidal modes that are sensitive to the inner core or couple strongly to an inner core mode, since our focus is on the mantle. This makes the total number of spectral segments 84 248, for 76 modes and mode groups in the one-step inversion and 104 self-coupled and 8 cross-coupled splitting functions for the two-step inversion (Tables 1 and 2). Our data set displays a significant increase with respect to only 10 events, 33 self-coupled spheroidal mode splitting functions and approximately 1000 spectral segments of Li *et al.* (1991). We invert for  $V_s$ , which is the observable that modes are most sensitive to. To reduce the number of independent parameters, we assume that the aspherical perturbations in compressional velocity,  $\delta \ln(V_p)$ , and density,  $\delta \ln \rho$ , are proportional to  $\delta \ln(V_s)$  with factors 0.5 (Li *et al.* 1991) and 0.3 (Karato 1993), respectively. We will remove these scaling factors in a future study. Crustal contributions are included using model CRUST 5.1 (Mooney *et al.* 1998). In the one-step inversion, crustal effects are included in the splitting matrix for computing synthetic spectra. Splitting function predictions for the crust are subtracted from the measured  $c_{st}$ 's, effectively removing the crust.

We use the same parametrization as global mantle shear wave velocity models S20RTS (Ritsema *et al.* 1999) and S40RTS (Ritsema *et al.* 2011), using 21 splines for depth, and spherical harmonics laterally. Models S20RTS and S40RTS go up to spherical harmonic degree 20 and 40, respectively, but here we will use maximum degree 8. We will initially invert for even degrees only, but explore the odd degrees as well. We use 1-D model PREM (Dziewonski & Anderson 1981) as starting model for the one-step inversion and for the second step of the two-step inversion.

## 4 RESULTS

We invert normal mode data for a shear velocity mantle model using the one-step direct spectrum inversion and the two-step splitting function inversion. We will compare the resulting models and misfits in detail below.

**Table 1.** Spectral segments containing the mode groups that were used in the one-step full-spectrum inversion, with the number of events as  $N_{ev}$ .

Spectral segment	$N_{ev}$
0S <sub>2</sub>	8
0S <sub>3</sub> – 0T <sub>2</sub> – 2S <sub>1</sub>	37
0S <sub>4</sub> – 0T <sub>3</sub> – 1S <sub>2</sub>	56
0S <sub>5</sub>	65
0S <sub>6</sub>	82
0S <sub>7</sub>	87
0S <sub>8</sub> – 4S <sub>1</sub> – 0T <sub>9</sub>	89
0S <sub>9</sub> – 0T <sub>10</sub>	91
0S <sub>11</sub> – 2S <sub>7</sub> – 0T <sub>12</sub>	86
0S <sub>13</sub> – 0T <sub>14</sub>	91
0S <sub>14</sub> – 2S <sub>9</sub> – 0T <sub>15</sub>	91
0S <sub>15</sub> – 0T <sub>16</sub>	91
0S <sub>16</sub> – 0T <sub>17</sub>	91
0S <sub>17</sub> – 2S <sub>11</sub> – 0T <sub>18</sub>	91
0S <sub>19</sub> – 0T <sub>20</sub>	90
0S <sub>21</sub> – 1S <sub>14</sub> – 0T <sub>22</sub>	91
1S <sub>2</sub> – 0S <sub>4</sub> – 0T <sub>3</sub>	9
1S <sub>4</sub>	77
1S <sub>5</sub> – 2S <sub>4</sub>	84
1S <sub>6</sub> – 2S <sub>5</sub>	82
1S <sub>7</sub>	74
1S <sub>8</sub>	75
1S <sub>9</sub>	74
1S <sub>10</sub>	70
2S <sub>1</sub> – 0S <sub>3</sub> – 0T <sub>2</sub>	3
2S <sub>6</sub>	84
2S <sub>8</sub> – 4S <sub>3</sub>	91
2S <sub>10</sub> – 4S <sub>5</sub>	89
2S <sub>12</sub>	90
2S <sub>13</sub>	90
3S <sub>6</sub>	87
3S <sub>7</sub> – 5S <sub>5</sub>	86
3S <sub>9</sub>	90
4S <sub>2</sub> – 0S <sub>10</sub> – 0T <sub>11</sub>	90
4S <sub>4</sub> – 1T <sub>8</sub>	86
5S <sub>3</sub>	86
5S <sub>4</sub> – 2T <sub>4</sub>	89
5S <sub>6</sub> – 0S <sub>21</sub> – 0T <sub>22</sub>	86
5S <sub>7</sub>	78
5S <sub>8</sub>	87
5S <sub>11</sub> – 7S <sub>8</sub>	73
5S <sub>12</sub>	63
5S <sub>14</sub> – 9S <sub>8</sub>	79
5S <sub>16</sub> – 8S <sub>10</sub>	80
5S <sub>17</sub>	65
6S <sub>9</sub> – 7S <sub>6</sub>	87
6S <sub>10</sub>	84
6S <sub>15</sub> – 9S <sub>10</sub>	62
6S <sub>18</sub>	78
7S <sub>5</sub>	89
7S <sub>7</sub>	77
7S <sub>9</sub> – 9S <sub>6</sub>	57
8S <sub>6</sub>	56
8S <sub>7</sub>	59
9S <sub>11</sub>	72
9S <sub>12</sub> – 10S <sub>10</sub>	66
9S <sub>13</sub> – 5S <sub>22</sub>	58
9S <sub>14</sub> – 14S <sub>7</sub>	49
9S <sub>15</sub> – 14S <sub>8</sub>	39
10S <sub>20</sub> – 15S <sub>12</sub> – 16S <sub>10</sub>	57
11S <sub>9</sub>	40
11S <sub>10</sub> – 4S <sub>28</sub>	42

**Table 1.** Continued

Spectral segment	$N_{ev}$
11S <sub>23</sub> – 13S <sub>18</sub> – 19S <sub>10</sub>	50
11S <sub>24</sub> – 15S <sub>15</sub>	42
11S <sub>25</sub>	35
12S <sub>6</sub>	36
12S <sub>7</sub>	56
12S <sub>13</sub>	49
12S <sub>14</sub>	42
12S <sub>15</sub>	41
13S <sub>16</sub> – 14S <sub>13</sub> – 16S <sub>11</sub>	50
13S <sub>19</sub> – 19S <sub>11</sub>	43
13S <sub>20</sub>	39
14S <sub>14</sub>	26
15S <sub>16</sub> – 17S <sub>15</sub>	49
17S <sub>14</sub>	29

#### 4.1 L-curves and misfits

We investigate and compare the L-curves and corresponding misfits of the models resulting from the one-step and two-step inversions (Fig. 2). All L-curves show the characteristic shape of decreasing misfit as a function of model size (Fig. 2a) and also as a function of number of effective unknowns or independent model parameters (Fig. 2b). Model size is defined as

$$\text{msize} = \sum_{s=0}^{s_{\max}} \sum_{t=-s}^s m_{st}^2 \quad (18)$$

and the number of effective unknowns as the trace of the resolution matrix. The most optimum model is usually defined at the ‘kink’ in the L-curve, at which the misfit does not significantly decrease anymore but model size and the number of effective unknowns still increase, filling the null space. As mentioned in Section 2.4.2, we can compute two kinds of misfit for the two-step inversion: misfit to the measured splitting function coefficients and misfit to the spectra. We find that the shapes of the L-curves for the two-step inversion are similar for both the  $c_{st}$  misfit (dotted line in Fig. 2a) and the spectral misfit (dashed line in Fig. 2a) and that the kink in the curve is found in the same place for both misfits (black dots in Fig. 2a). So, the chosen best model based on spectral misfit is the same model we would choose based on  $c_{st}$  misfit. The  $c_{st}$  misfit is lower than the spectral misfits, but it cannot be directly compared to the spectral misfits, since it is a different kind of misfit for a different kind of data (i.e.  $c_{st}$ ’s instead of spectra). The two chosen  $\delta \ln(V_s)$  models have approximately the same number of effective unknowns (Fig. 2b), which allows for direct comparison between the models. It is interesting to note that for the same number of effective unknowns, the one-step model has a slightly larger model size than the two-step model, which is probably a direct consequence of the inversion method. The same feature is also often seen in full-waveform model inversions, which usually have stronger anomalies than simple arrival time inversions.

The most important observation we make from the L-curves is that the total average spectral misfit for the one-step inversion is consistently lower than for the two-step inversion, for all model sizes and numbers of effective unknowns (compare solid curve to dashed curve in Fig. 2). Thus, the one-step inversion results in a better fit to the original data making the one-step model the preferred choice. Also, it shows that the worries initially voiced by Li *et al.* (1991) and Durek & Romanowicz (1999) that information might be lost in the two-step inversion resulting in a worse fit to the original data,

**Table 2.** Modes and mode pairs for which their measured splitting functions were used in the two-step inversion, with the number of events as  $N_{\text{ev}}$ . Modes between brackets were included in the forward computations, but not inverted for.

Splitting function	$N_{\text{ev}}$
$0S_2$	8
$0S_3(-0T_2-2S_1)$	37
$0S_4(-0T_3-1S_2)$	56
$0S_5$	65
$0S_6$	82
$0S_7$	87
$0S_8(-4S_1-0T_9)$	89
$0S_9(-0T_{10})$	91
$0S_{11}(-2S_7-0T_{12})$	86
$0S_{11}-2S_7(-0T_{12})$	86
$0S_{13}(-0T_{14})$	91
$0S_{14}(-2S_9-0T_{15})$	91
$0S_{15}(-0T_{16})$	91
$0S_{16}(-0T_{17})$	91
$0S_{17}(-2S_{11}-0T_{18})$	91
$0S_{17}-2S_{11}(-0T_{18})$	91
$0S_{19}(-0T_{20})$	90
$0S_{21}(-1S_{14}-0T_{22})$	91
$1S_2(-0S_4-0T_3)$	9
$1S_4$	77
$1S_5(-2S_4)$	84
$1S_6(-2S_5)$	82
$1S_7$	74
$1S_8$	75
$1S_9$	74
$1S_{10}$	70
$1S_{14}(-0S_{21}-0T_{22})$	91
$2S_1$	3
$2S_4(-1S_5)$	84
$2S_5(-1S_6)$	82
$2S_6$	84
$2S_7(-0S_{11}-0T_{12})$	86
$2S_8(-4S_3)$	91
$2S_9(-0S_{14}-0T_{15})$	91
$2S_{10}(-4S_5)$	89
$2S_{11}(-0S_{17}-0T_{18})$	91
$2S_{12}$	90
$2S_{13}$	90
$3S_6$	87
$3S_7(-5S_5)$	86
$3S_7-5S_5$	86
$3S_9$	90
$4S_1(-0S_8-0T_9)$	89
$4S_2(-0S_{10}-0T_{11})$	90
$4S_3(-2S_8)$	91
$4S_4(-1T_8)$	86
$4S_5(-2S_{10})$	89
$5S_3$	86
$5S_4(-2T_4)$	89
$5S_5(-3S_7)$	86
$5S_6(-0S_{21}-0T_{22})$	86
$5S_7$	78
$5S_8$	87
$5S_{11}(-7S_8)$	73
$5S_{12}$	63
$5S_{14}(-9S_8)$	79
$5S_{14}-9S_8$	79
$5S_{16}(-8S_{10})$	80
$5S_{16}-8S_{10}$	80

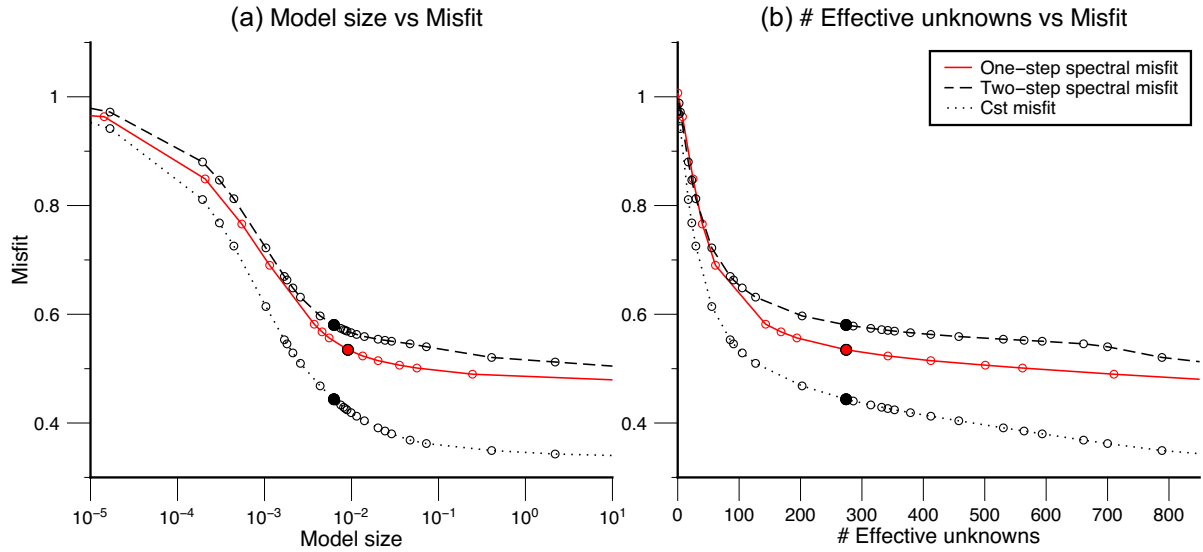
**Table 2.** Continued

Splitting function	$N_{\text{ev}}$
$5S_{17}$	65
$6S_9(-7S_6)$	87
$6S_{10}$	84
$6S_{15}(-9S_{10})$	62
$6S_{18}$	78
$7S_5$	89
$7S_6(-6S_9)$	87
$7S_7$	77
$7S_8(-5S_{11})$	73
$7S_9(-9S_6)$	57
$8S_6$	56
$8S_7$	59
$8S_{10}(-5S_{16})$	80
$9S_6(-7S_9)$	57
$9S_8(-5S_{14})$	79
$9S_{10}(-6S_{15})$	62
$9S_{11}$	72
$9S_{12}(-10S_{10})$	66
$9S_{12}-10S_{10}$	66
$9S_{13}(-5S_{22})$	58
$9S_{14}(-14S_7)$	49
$9S_{15}(-14S_8)$	39
$10S_{10}(-9S_{12})$	66
$10S_{20}(-15S_{12}-16S_{10})$	57
$11S_9$	40
$11S_{10}(-4S_{28})$	42
$11S_{23}(-13S_{18}-19S_{10})$	50
$11S_{24}(-15S_{15})$	42
$11S_{25}$	35
$12S_6$	36
$12S_7$	56
$12S_{13}$	49
$12S_{14}$	42
$12S_{15}$	41
$13S_{16}(-14S_{13}-16S_{11})$	50
$13S_{18}(-11S_{23}-19S_{10})$	50
$13S_{19}(-19S_{11})$	43
$13S_{20}$	39
$14S_7(-9S_{14})$	49
$14S_8(-9S_{15})$	39
$14S_{13}(-16S_{11}-13S_{16})$	50
$14S_{13}-16S_{11}(-13S_{16})$	50
$14S_{14}$	26
$15S_{12}(-16S_{10}-10S_{20})$	57
$15S_{12}-16S_{10}(-10S_{20})$	57
$15S_{15}(-11S_{24})$	42
$15S_{16}(-17S_{15})$	49
$16S_{10}(-15S_{12}-10S_{20})$	57
$16S_{11}(-14S_{13}-13S_{16})$	50
$17S_{14}$	29
$17S_{15}(-15S_{16})$	49
$19S_{10}(-13S_{18}-11S_{23})$	50
$19S_{11}(-13S_{19})$	43

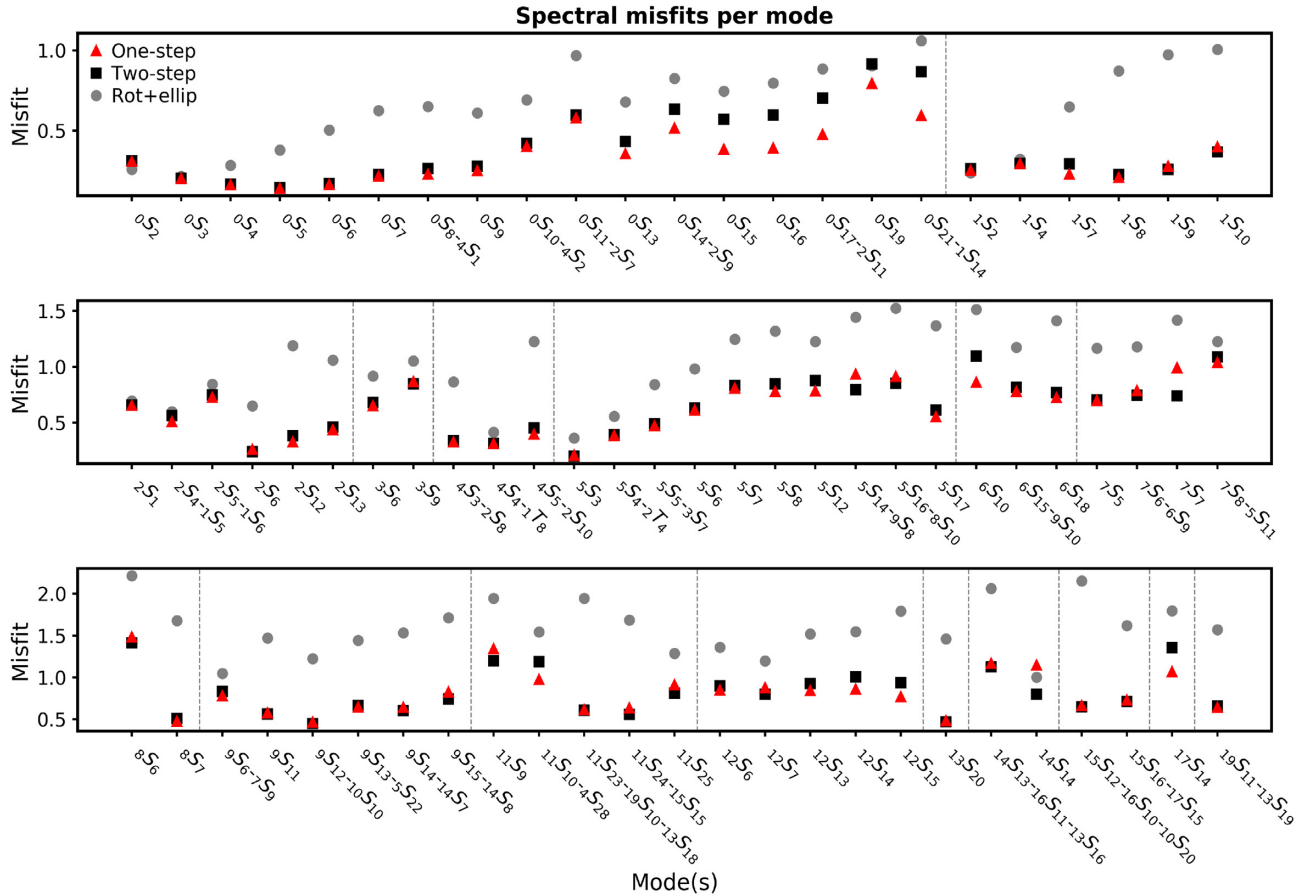
are valid. It is possible that by applying a different splitting function weighting procedure we might succeed to shift the two-step spectral misfit L-curve to lower misfits, but we think weighting by error is the most logical procedure and most similar to implicit weighting by data amount in the one-step inversion.

All our inversions are started from PREM, only taking rotation and ellipticity into account. If we examine the average spectral misfit per mode or small group of modes (Fig. 3), we see that both our models improve the spectral data fit with respect to PREM





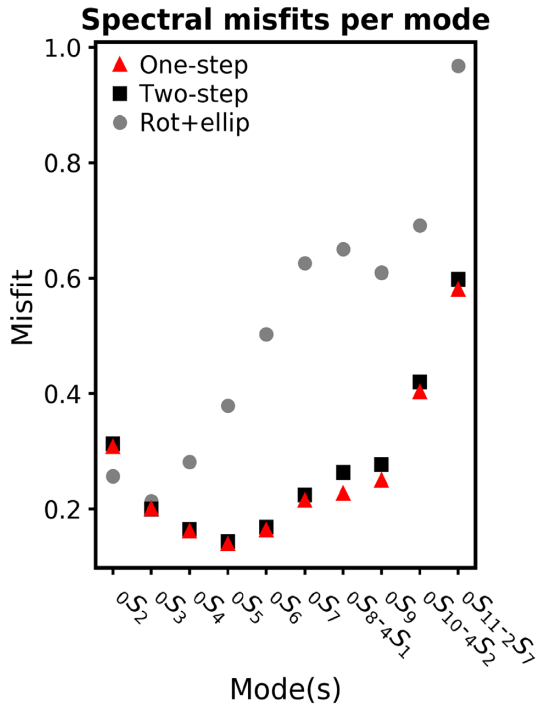
**Figure 2.** L-curve for the one-step inversion spectral misfits (red), and L-curves for the  $c_{st}$  misfits (dotted) and spectral misfits (dashed) for the models resulting from the two-step inversion, as a function of (a) model size and (b) number of effective unknowns. Red dot: best model for one-step inversion. Black dots: best model for the two-step inversion.



**Figure 3.** Average spectral misfit per mode or mode group for the best models from the one-step (red triangles) and two-step (black squares) inversion, compared to predicted misfits for PREM plus rotation and ellipticity (grey dots).

plus rotation and ellipticity across almost the entire mode range. For some low angular order modes the misfit neither decreases nor increases significantly. The lower average spectral misfit of the one-step model is mostly due to lower misfits for the higher angular order

fundamental modes ( $0S_{13} - 0S_{21}$ ), however, the average spectral misfit without these modes of the one-step inversion is still lower (0.55) than for the two-step inversion (0.56). The one-step model fits the spectra better than or equally well as the two-step model for



**Figure 4.** Zooming in on the lower angular order fundamentals of Fig. 3.

the majority of modes. Even if it seems that there is no difference between the one-step and two-step spectral misfit, such as for the lower angular order fundamentals, for some modes the one-step spectral misfit is lower (Fig. 4). The few exceptions are mainly  $V_p$  sensitive modes.

The difference in misfit between the two inversion methods is also clearly visible in synthetic spectra based on the two best models (Fig. 5). Although both our models fit the spectra better than PREM + rotation and ellipticity, the model resulting from the one-step inversion fits the position and amplitude of the peaks better, especially for modes with lower overtone numbers such as the fundamentals shown in Fig. 5. Not all individual spectra can be fitted better when including 3-D  $V_s$  variations, with scaled  $V_p$  and density (Fig. 6). In these cases, adding independent constraints on  $V_p$  and density might improve the fit.

## 4.2 Model characteristics

We first look at the general features of the two models. Model slices at several depths in the mantle show the similarities in pattern and amplitude of  $\delta \ln(V_s)$  (Figs 7a and b) of the models made using the two different approaches. At first glance, the two models look very similar. Both show the fast slab anomalies in the upper mantle around 500 km depth, and the Large Low Shear Velocity Provinces (LLSVPs) in the lowermost mantle. If you look in more detail, small differences become visible, for example in the low velocity anomalies in the LLSVPs around 2000 and 2800 km depth which show stronger negative velocity anomalies in the one-step than in the two-step model. We also compare our models to SP12RTS- $V_s$  (Koelemeijer *et al.* 2016) and S40RTS (Ritsema *et al.* 2011). Both of these models include measurements of body wave traveltimes, and Rayleigh surface wave phase velocities, in addition to normal mode splitting function coefficients, providing more data sensitive to the uppermost and lowermost parts of the mantle. From a depth of 500 km onward, our models start to look similar to SP12RTS and

S40RTS. The differences around 200 km depth are due to the lack of surface waves in our models which are based only on normal mode data.

The differences between our two models become more apparent in cross-sections through the mantle (Fig. 8). We see in the top panel of Fig. 8 that the Pacific LLSVP in the one-step model is more stretched in a Northwest-Southeast orientation, compared to the two-step model. The slab subducting beneath Japan seems more continuous in the two-step model. In the bottom panel of Fig. 8, the African LLSVP in the two-step model seems to be more connected to the East African Rift anomaly in the upper mantle than in the one-step model.

The one-step and two-step models are similar in their power spectrum per structural degree (Figs 9a and b), though the power distribution as a function of depth appears to be smoother in the one-step than in the two-step inversion. Any differences in power spectra per degree could arise from the inversion method, or the difference in the amount of mode coupling between the two methods. Our models share the characteristic of SP12RTS and S40RTS (Figs 9c and d) of a dominant degree two structure, especially in the upper and lower parts of the mantle. Degree two structure in the uppermost mantle is the consequence of tectonic processes, where subducting slabs form the ‘ring around the Pacific’. LLSVPs result in a large degree two signature in the lowermost mantle. As we already saw in the depth slices (Figs 7a and b), the power in the uppermost 200 km of the mantle is weaker in our models compared to SP12RTS and S40RTS, probably due to a lack of surface waves in our inversions. Lower power in the CMB region could be the result of omitting Stoneley modes and core diffracted body waves from our inversions.

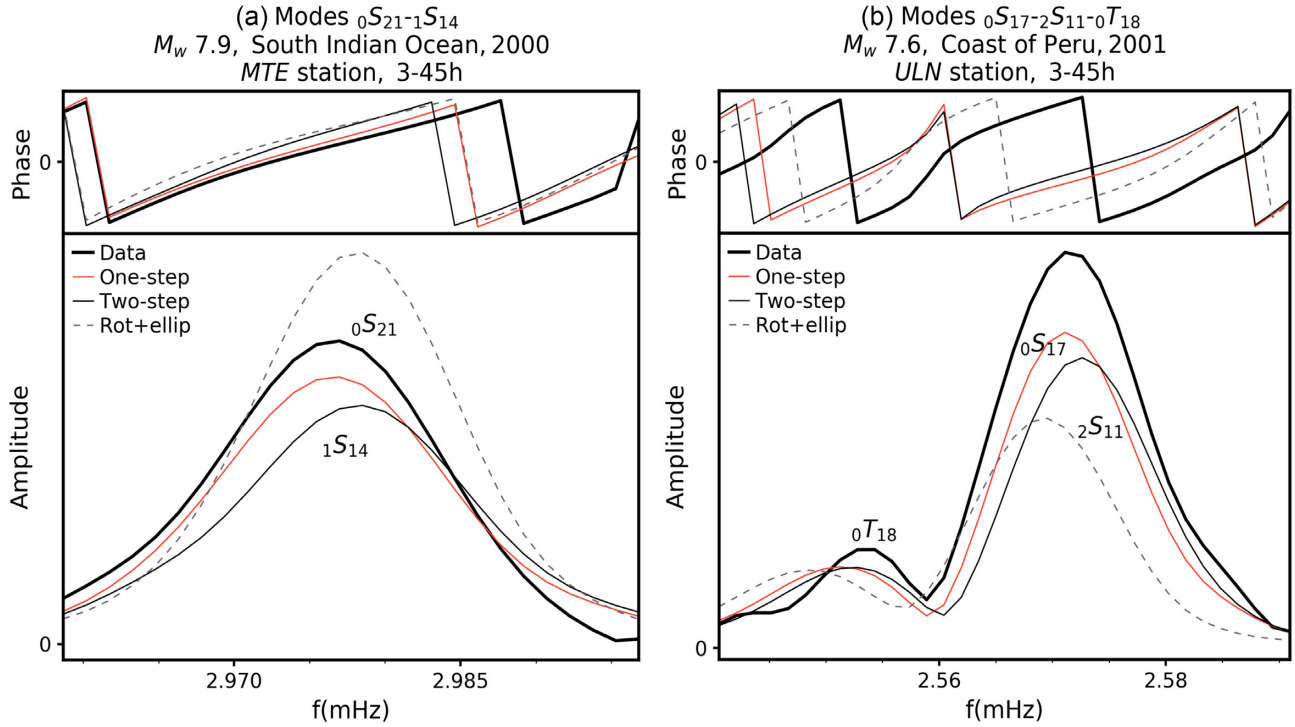
By visualizing data sensitivity of our normal mode data set, we get a better understanding of our resulting models. Data sensitivity to each depth spline is related to the diagonal elements of the  $\mathbf{A}^T \mathbf{A}$  matrix. Following Gu *et al.* (2001) and Koelemeijer *et al.* (2016), data sensitivity for the  $k$ th spline is defined by a horizontal average:

$$\widehat{(\mathbf{A}^T \mathbf{A})}_k = \tilde{a}_k = \sqrt{\frac{1}{4\pi} \sum_{st} a_{st,k}^2}, \quad (19)$$

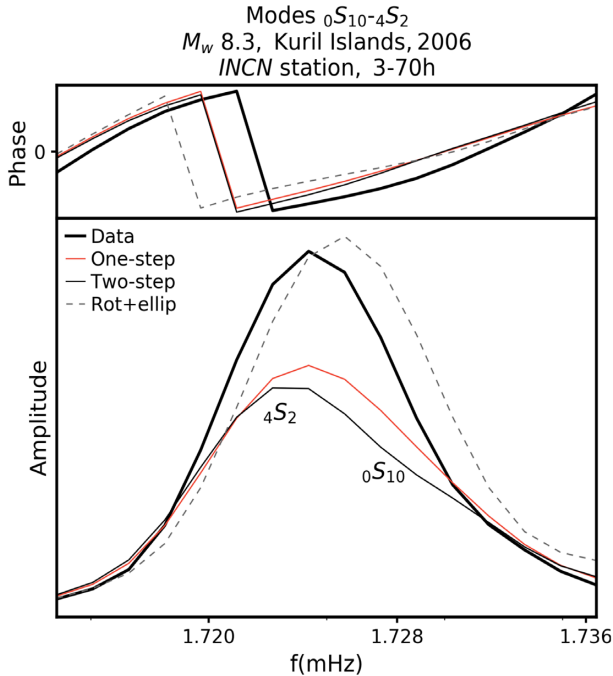
where  $a_{st,k}$  are diagonal elements of  $(\mathbf{A}^T \mathbf{A})_k$ . Both the one-step and two-step models have peak sensitivity at the second to deepest spline (2600 km depth) and least sensitivity in the most shallow part of the mantle (Fig. 10a). Note that the sensitivity in Fig. 10(a) is normalized. We cannot draw any conclusions from the absolute sensitivities in the two inversion approaches, because it is not possible to compare the  $\mathbf{A}^T \mathbf{A}$  of the linear two-step inversion to that of the non-linear one-step inversion.

The root mean square (RMS) amplitudes of the four models paint the same picture as the power spectra: higher amplitudes are seen in the uppermost mantle and CMB region for SP12RTS and S40RTS than for our two models (Fig. 10b). In the depth range of 700–2500 km all  $V_s$  models have similar RMS amplitudes. It is the depth range dominated by the normal mode data in SP12RTS and S40RTS, which explains the similarity in RMS amplitudes across all four models. Comparing our own models, we again see that the RMS amplitude of the one-step model varies more smoothly as a function of depth than the RMS amplitude of the two-step model.

The cross-correlations between the four models as a function of depth are shown in Fig. 10(c). The correlation between our two models is close to 1.0 in the upper 1000 km of the mantle and drops to values of less than 0.9 around 1250 km depth and the



**Figure 5.** Synthetic spectra for 1-D structure plus coupling for rotation and ellipticity (grey dashes), and for the best models from the one-step inversion (red) and from the two-step inversion (thin black). Thick black spectra represent real data of events (a) in the South Indian Ocean and (b) off the coast of Peru.



**Figure 6.** Synthetic spectra for 1-D structure plus coupling for rotation and ellipticity (grey dashes), and for the best models from the one-step inversion (red), and from the two-step inversion (thin black). Thick black spectrum represents real data of an event near the Kuril Islands.

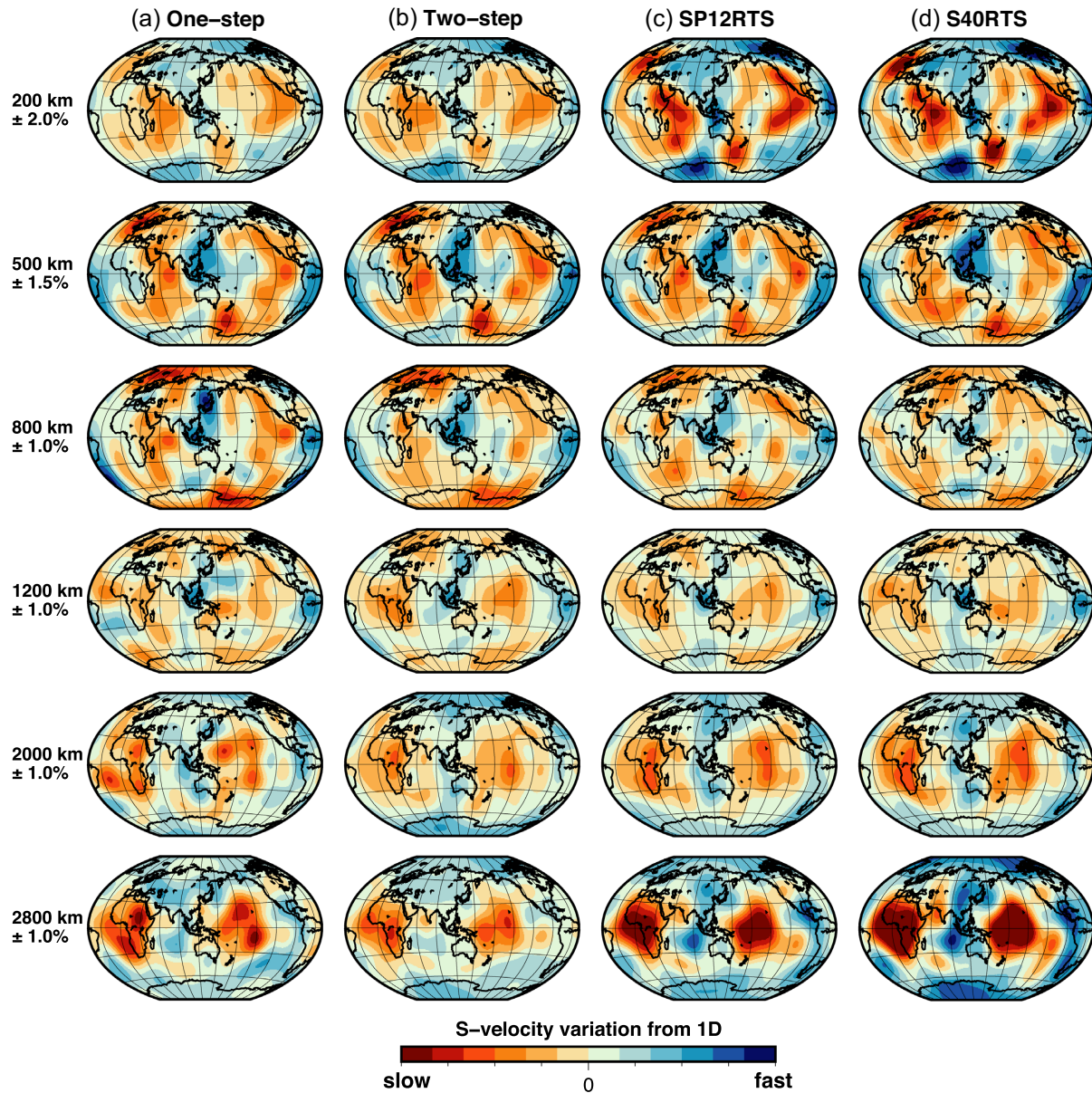
CMB. The drop around 1250 km depth can be explained by looking at the 1200 km slice in Fig. 7 and the power spectra in Fig. 9. Here, the one-step model has a higher power for degree 6, resulting in more dominant small-scale structures than in the two-step model. The drop around the CMB is caused by differently shaped

LLSVPs, especially the Pacific LLSVP. In the one-step model both LLSVPs consist of two elongated slow regions in a north–south configuration, whereas in the two-step model both LLSVPs are more stretched in an east–west direction. These drops in correlation between the one-step and two-step models are probably the consequence of the inversion method and might be due to errors propagating from the intermediate splitting function measurement into the final model in the two-step inversion method, since we used the same spectral data in both methods.

Our two models correlate better with SP12RTS than with S40RTS. Correlation of all other three models with S40RTS is especially poor at 800 and 1400 km depth, although their RMS amplitudes are similar. The poor correlation could be caused by S40RTS not including modes with overtone number  $>5$ , in contrast to our models and SP12RTS which do include higher overtones.

We have started the one-step inversion and the second step of the two-step inversion from PREM. To check whether or not we converge to a local minimum, we also experimented with starting our inversions from S20RTS (Ritsema *et al.* 1999) and a random model (Figs S1–S4). If we apply sufficiently low damping, we end up with models that look very similar to the ones obtained starting from PREM, for both inversion methods and all starting models, converging to comparable misfits and model sizes (Fig. S5). The only difference in observed patterns is in the uppermost 200 km of the mantle, implying that this part of the mantle is most poorly constrained by the normal mode data. The data sensitivity distribution in our models (Fig. 10a) agrees with this observation. Including spheroidal fundamentals with higher angular order, which are similar to Rayleigh surface waves, might improve the data coverage of this part of the Earth. An important observation for this study is that the spectral misfit for the one-step inversion is always lower than for the two-step inversion, independent of the starting model (Fig. S5).





**Figure 7.** Model slices showing even degrees 2, 4, 6, 8 for a) the one-step inversion; b) two-step inversion; c) SP12RTS (Koelemeijer *et al.* 2016); d) S40RTS (Ritsema *et al.* 2011).

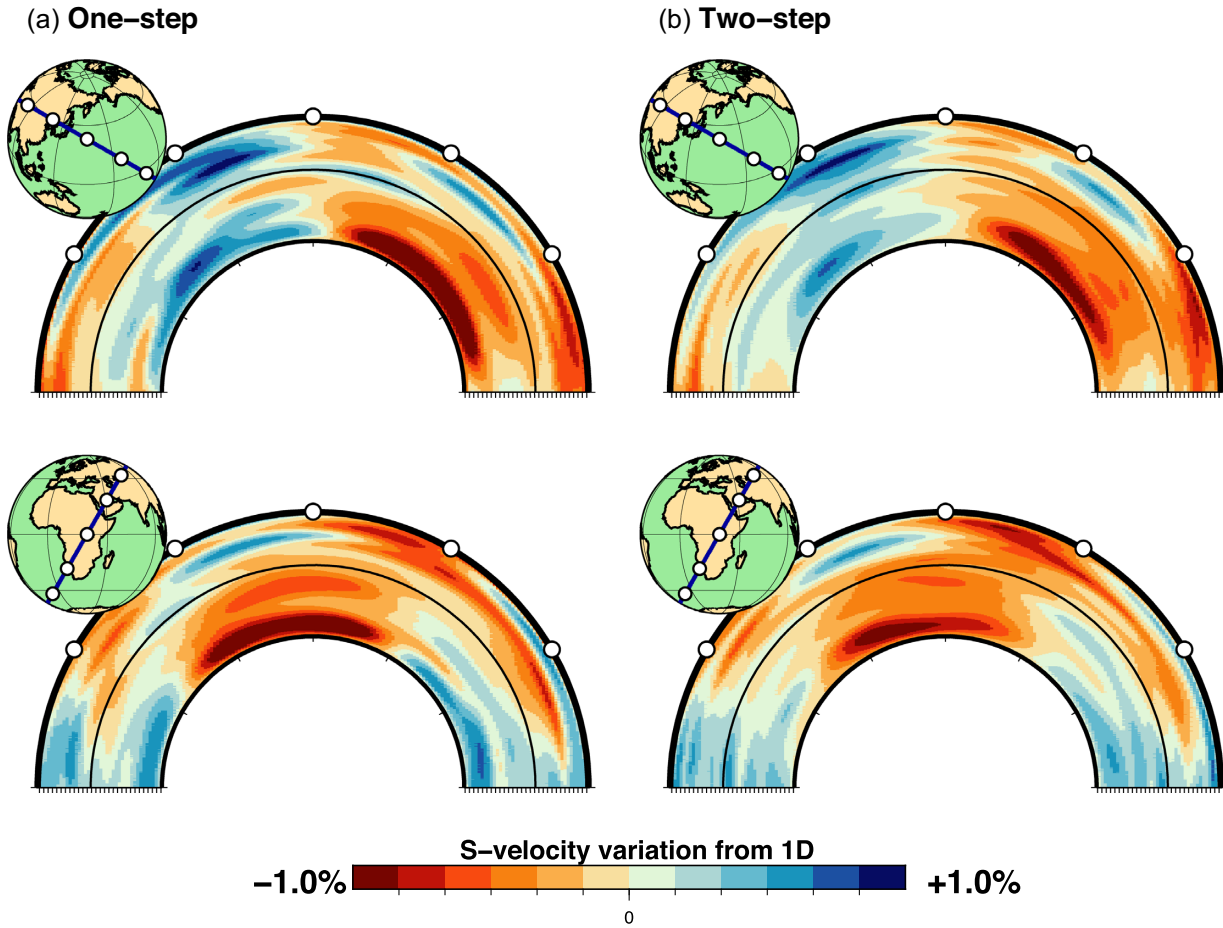
### 4.3 Odd degrees

We have so far only considered even degree structure, but we are wondering if odd-degree structure can also be resolved and if there is also a difference between the two inversion methods for odd-degree structure. Sensitivity of normal modes in self-coupling is limited to even degrees, since waves travelling in opposite directions destructively interfere for odd-degree structure. This can also be understood from the coupling rules, which state that two spheroidal modes  ${}_nS_l$  and  ${}_nS_{l'}$  only couple for structure  $|l - l'| \leq s \leq |l + l'|$ . In the case of cross-coupling, we will get sensitivity to odd-degree structure if  $l - l'$  is an odd number. Therefore we are bound to mode pairs that couple for odd-degree structure according to the coupling rules and are close enough in frequency to form a group or ‘hybrid’ mode pair. Unfortunately, we will show below that only a very small number of spheroidal modes pairs differ in their angular order  $l$  by an odd-number and are close enough in frequency, limiting the number

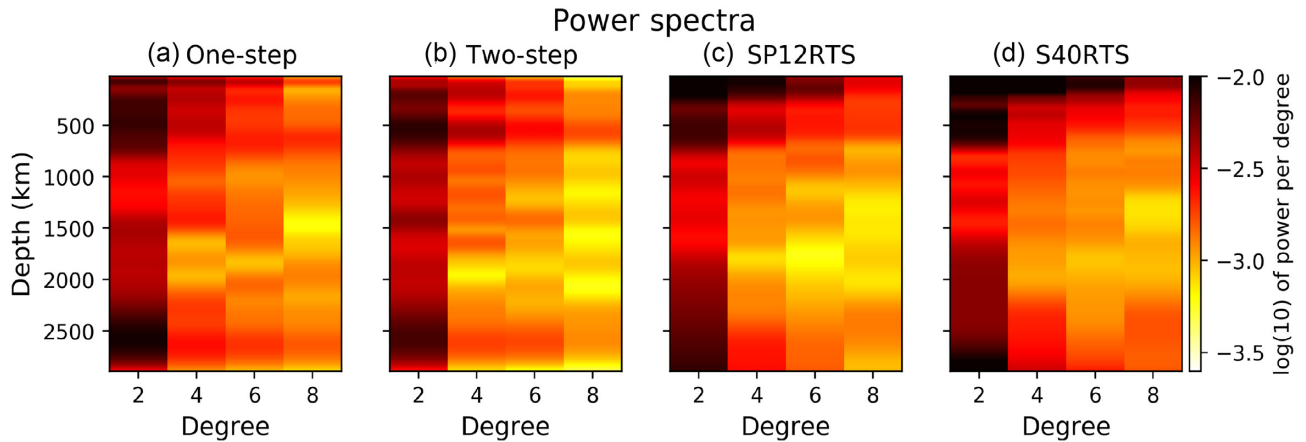
of cross-coupled splitting functions and spectral data available to constrain odd-degree structure.

We extend the one-step and two-step inversion to include odd degrees. Allowing modes in the small mode groups of Deuss *et al.* (2013) to couple for odd degrees yields a one-step model with a power spectrum containing significantly less power in odd degrees than in even degrees (Fig. 11a). Adding odd-degree splitting functions to the even-degree two-step inversion gives the same observation in the power spectrum (Fig. 11b), with less power for degree 1, 3 and 7. This is because measurements of odd-degree cross-coupled splitting functions by Deuss *et al.* (2013) are limited, especially compared to the large number of even-degree splitting functions available (blue bars in Fig. 12). Odd and even degrees were given the same norm damping to illustrate the lack of normal mode data sensitive to odd degrees compared to even degrees. The extra data added in this manner does not contribute greatly to the data sensitivity kernels (Fig. 13), though the relative increase in





**Figure 8.** Cross-sections through the models from the a) one-step inversion; b) two-step inversion. Top panel shows the subducted plate near Japan and part of the Pacific LLSVP; bottom panel shows the African LLSVP.

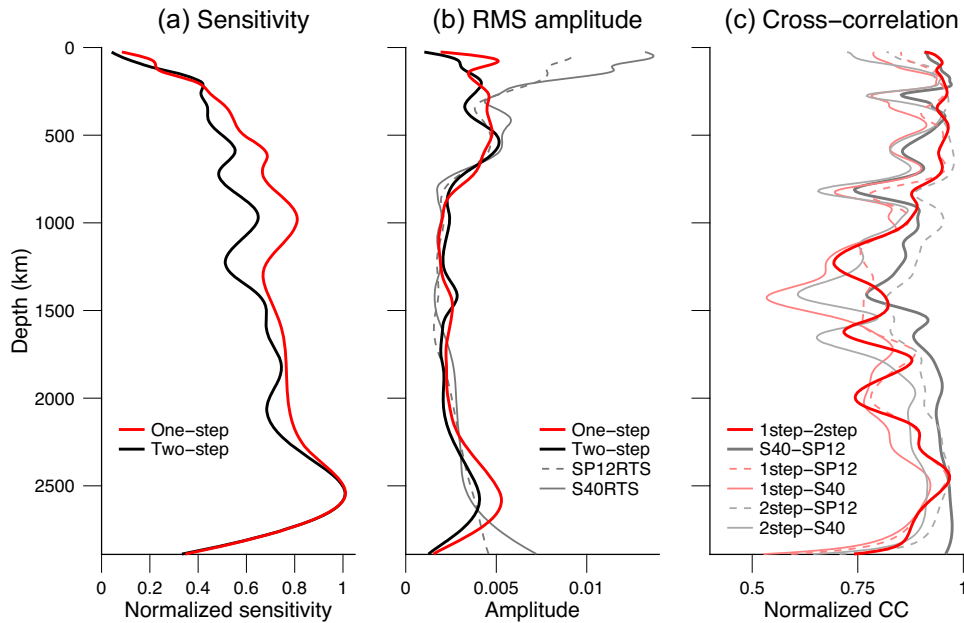


**Figure 9.** Power spectra per spherical harmonic degree for the four  $\delta \ln(V_s)$  models shown in Fig. 7.

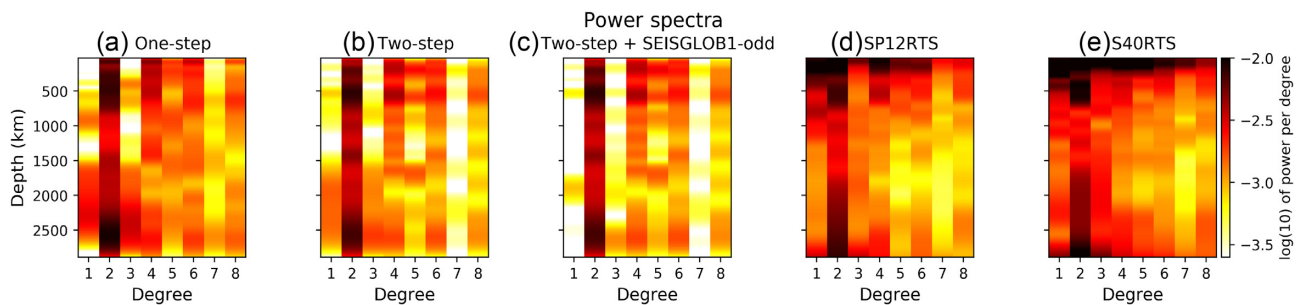
sensitivity is larger in the one-step inversion than in the two-step inversion. Thus there is only little information contained in the small amount of cross-coupled splitting functions for odd-degree structure. Group coupling in the one-step inversion, which includes a small number of mode pairs sensitive to odd degrees, does not add much information either.

Nonetheless, a recent study by Durand *et al.* (2016) argues that adding odd-degree splitting function coefficients results in odd-degree structure in the lowermost mantle becoming stronger than

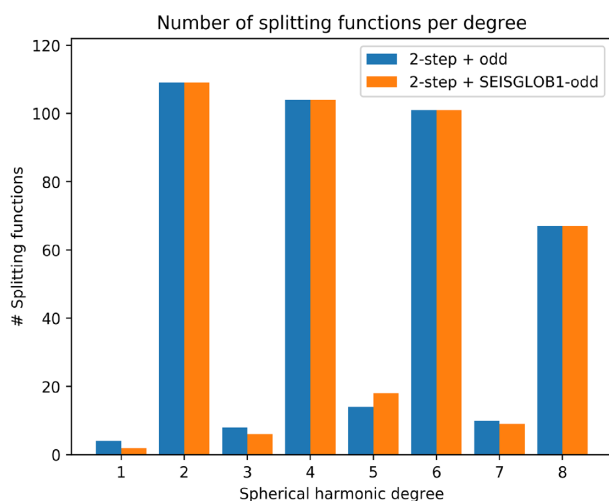
previously believed. Their mantle model SEISGLOB1 contains a higher spectral amplitude for degree 3 than degree 2 at 2800 km depth. Their cross-coupled normal mode data comprises measurements of Resovsky & Ritzwoller (1998) and Deuss *et al.* (2013). The odd-degree cross-coupled splitting functions of Deuss *et al.* (2013) and Resovsky & Ritzwoller (1998) comprises 16 and 3 spheroidal-spheroidal measurements, respectively, with some overlap. We also run our two-step inversion using their odd-degree mode pairs with equal damping for all degrees, which results in the power spectrum



**Figure 10.** (a) Normalized sensitivity for one-step and two-step models; (b) Root mean square amplitude of the four models in Fig. 7; (c) normalized cross-correlation coefficients of all combinations of the same four models (SP12 = SP12RTS, S40 = S40RTS).



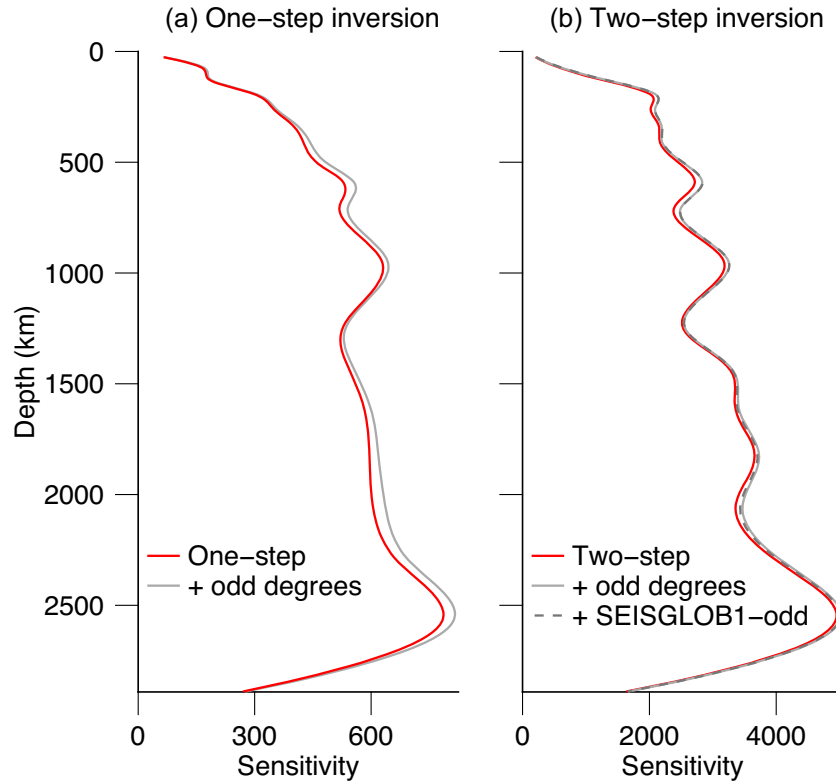
**Figure 11.** Power spectra per spherical harmonic degree, including odd degrees, showing in (a), (b), (d), (e) the four  $\delta \ln(V_s)$  models shown in Fig. 7 and (c) two-step model including odd degree couples from Durand *et al.* (2016).



**Figure 12.** Number of splitting functions (self- and cross-coupled) that constrain the spherical harmonic degrees for a two-step inversion including the odd-degree  $c_{st}$  coefficients of Deuss *et al.* (2013) (in blue) and Durand *et al.* (2016) (in orange).

of Fig. 11(c). Because they actually included a few modes less than we did, there is still some power around 2800 km for degree 3, but in fact less than in our two-step inversion (Fig. 11b) and in all cases the degree 3 structure stays smaller than the degree 2 structure. Again, the data sensitivity kernels (Fig. 13b) also show that there is negligible data sensitivity to odd degrees in the model made using the same cross-coupled splitting functions as Durand *et al.* (2016).

Similar to Durand *et al.* (2016), who increased the weights of odd-degree splitting functions to fit these data, the only way in which we are able to obtain strong odd-degree lowermost mantle patterns as in SEISGLOB1 is by applying degree-dependent damping, even though the L-curves of these odd degrees do not show any indication to do so. The non-characteristic shape of the L-curves per degree illustrates that the amount of splitting function coefficients used by Durand *et al.* (2016) is not enough to constrain whole-mantle odd-degree structure. To support this statement, we would like to add that the six mode pairs in the SEISGLOB1 data set sensitive to degree 3 (i.e.  $6 \times 7 = 42$  splitting function ‘data’ coefficients) are not sufficient to constrain all model parameters at degree 3 for 21 depth splines (i.e.  $21 \times 7 = 147$  ‘unknown’ model coefficients). Overall, we find that the two-step inversion would need more cross-coupled splitting functions that are sensitive to odd-degree structure to make robust  $V_s$  models of odd-degree structure. Although there



**Figure 13.** Data sensitivity kernels (not normalized) for (a) one-step model (red), plus odd degrees by Deuss *et al.* (2013) (grey); (b) two-step model (red), plus odd degrees by Deuss *et al.* (2013) (grey), plus odd degrees in SEISGLOB1 (Durand *et al.* 2016) (dashed).

is stronger odd-degree sensitivity in the one-step inversion, even there we also need to include more mode couples sensitive to odd degrees to make robust models of odd-degree structure.

## 5 DISCUSSION AND CONCLUSION

We have explored two different inversion methods that have been used in normal mode tomography: (i) the computationally intensive one-step full-spectrum inversion and (ii) the commonly used two-step splitting function inversion, using the same spectral data set of Deuss *et al.* (2013) as a starting point. Theoretically, the two-step inversion is more sensitive to data weighting, and information might be lost in the intermediate step of measuring splitting functions. We find that we need to weigh the splitting function coefficients by their error or uncertainties, otherwise the resulting mantle models will not be similar to  $V_s$  models obtained in previous studies, which were made with other data types including body waves and/or surface waves. The one-step inversion is more robust, both from a theoretical and a practical point of view, since all spectral data are inverted at once. Neither of the two inversion methods depends on the starting model in our inversion for  $V_s$  only, except in the uppermost mantle, where normal mode sensitivity is lacking in our data set.

The average spectral misfit for the one-step inversion is lower for all model sizes and numbers of effective unknowns, as displayed in the inversion L-curves. The model resulting from the one-step inversion is smoother than the model resulting from the two-step inversion, even when the model size and number of unknown parameters are the same between the two models. Moreover, the anomalies in  $V_s$  are stronger in the one-step model for the same number of effective unknowns, similar to stronger anomalies being observed in a full waveform inversion compared to conventional traveltime

tomography (Rickers *et al.* 2013). Whether this statement still holds when we invert for other parameters that modes are less sensitive to, such as  $V_p$  and  $\rho$ , will be addressed in a future paper. The two-step splitting function inversion for  $V_s$  suffices when combined with other data types (i.e. surface waves and body waves) and could be suitable for Monte Carlo model space search techniques (e.g. Fichtner *et al.* 2019).

We cannot draw any conclusions yet on odd-degree structure in the mantle based on the currently available small number of cross-coupled odd-degree normal mode splitting functions. It will be essential to expand the data set of odd-degree cross-coupled splitting functions in order to be able to robustly constrain odd-degree structure, especially in the lower mantle, where the odd degrees are currently poorly constrained by normal modes. To conclude, the one-step inversion results in lower misfits to the spectral data. Furthermore, the resulting model is more smooth and has stronger anomalies for the same number of effective unknowns. This makes the one-step inversion, even though more computational intensive, the preferred choice, when possible.

## ACKNOWLEDGEMENTS

We thank the editor Frederik Simons, Stephanie Durand and an anonymous reviewer for their constructive comments which helped to improve the manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 681535 - ATUNE) and a Vici award number 016.160.310/526 from the Netherlands organization for scientific research (NWO). The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related

metadata and/or derived products used in this study. Data analysis was done using ObsPy (Beyreuther *et al.* 2010) and Python package FrosPy developed by Simon Schneider, Sujania Talavera-Soza and L.J. Most figures in this paper were constructed using GMT (Wessel *et al.* 2013).

## DATA AVAILABILITY

The  $c_{st}$  measurements underlying this article are available in the supplementary materials of Deuss *et al.* (2013), and through the GitHub of FrosPy at <https://github.com/s-schneider/frosipy>.

## REFERENCES

- Akbarashrafi, F., Al-Attar, D., Deuss, A., Trampert, J. & Valentine, A., 2018. Exact free oscillation spectra, splitting functions and the resolvability of Earth's density structure, *Geophys. J. Int.*, **213**(1), 58–76.
- Al-Attar, D., Woodhouse, J.H. & Deuss, A., 2012. Calculation of normal mode spectra in laterally heterogeneous Earth models using an iterative direct solution method, *Geophys. J. Int.*, **189**(2), 1038–1046.
- Beghein, C. & Trampert, J., 2003. Robust normal mode constraints on inner-core anisotropy from model space search, *Science*, **299**(5606), 552–555.
- Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J., 2010. Obspy: a Python toolbox for seismology, *Seismol. Res. Lett.*, **81**(3), 530–533.
- Dahlen, F. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton Univ. Press.
- Davies, G.F. & Gurnis, M., 1986. Interaction of mantle dregs with convection: Lateral heterogeneity at the core-mantle boundary, *Geophys. Res. Lett.*, **13**(13), 1517–1520.
- Deuss, A. & Woodhouse, J.H., 2001. Theoretical free-oscillation spectra: the importance of wide band coupling, *Geophys. J. Int.*, **146**(3), 833–842.
- Deuss, A., Ritsema, J. & van Heijst, H., 2013. A new catalogue of normal-mode splitting function measurements up to 10 mHz, *Geophys. J. Int.*, **193**(2), 920–937.
- Durand, S., Debayle, E., Ricard, Y. & Lambotte, S., 2016. Seismic evidence for a change in the large-scale tomographic pattern across the D'' layer, *Geophys. Res. Lett.*, **43**(15), 7928–7936.
- Durek, J.J. & Romanowicz, B., 1999. Inner core anisotropy inferred by direct inversion of normal mode spectra, *Geophys. J. Int.*, **139**(3), 599–622.
- Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25**(4), 297–356.
- Dziewonski, A.M., Hager, B.H. & O'Connell, R.J., 1977. Large-scale heterogeneities in the lower mantle, *J. geophys. Res.*, **82**(2), 239–255.
- Edmonds, A., 1960. *Angular Momentum in Quantum Mechanics*, Princeton Univ. Press.
- Ekström, G., Nettles, M. & Dziewoński, A., 2012. The global cmt project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.*, **200**, 1–9.
- Fichtner, A., Zunino, A. & Gebraad, L., 2019. Hamiltonian Monte Carlo solution of tomographic inverse problems, *Geophys. J. Int.*, **216**(2), 1344–1363.
- Giardini, D., Li, X.-D. & Woodhouse, J.H., 1987. Three-dimensional structure of the Earth from splitting in free-oscillation spectra, *Nature*, **325**(6103), 405.
- Gu, Y.J., Dziewonski, A.M., Su, W. & Ekström, G., 2001. Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities, *J. geophys. Res.*, **106**(B6), 11 169–11 199.
- He, X. & Tromp, J., 1996. Normal-mode constraints on the structure of the Earth, *J. geophys. Res.*, **101**(B9), 20 053–20 082.
- Ishii, M. & Tromp, J., 1999. Normal-mode and free-air gravity constraints on lateral variations in velocity and density of Earth's mantle, *Science*, **285**(5431), 1231–1236.
- Ishii, M. & Tromp, J., 2004. Constraining large-scale mantle heterogeneity using mantle and inner-core sensitive normal modes, *Phys. Earth planet. Inter.*, **146**(1–2), 113–124.
- Karato, S.-I., 1993. Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, **20**(15), 1623–1626.
- Kellogg, L.H., Hager, B.H. & Van der Hilst, R.D., 1999. Compositional stratification in the deep mantle, *Science*, **283**(5409), 1881–1884.
- Koelemeijer, P., Ritsema, J., Deuss, A. & Van Heijst, H.-J., 2016. SP12RTS: a degree-12 model of shear- and compressional-wave velocity for Earth's mantle, *Geophys. J. Int.*, **204**(2), 1024–1039.
- Koelemeijer, P., Deuss, A. & Ritsema, J., 2017. Density structure of Earth's lowermost mantle from Stoneley mode splitting observations, *Nat. Commun.*, **8**, 15241.
- Kuo, C. & Romanowicz, B., 2002. On the resolution of density anomalies in the Earth's mantle using spectral fitting of normal-mode data, *Geophys. J. Int.*, **150**(1), 162–179.
- Lau, H.C., Mitrovica, J.X., Davis, J.L., Tromp, J., Yang, H.-Y. & Al-Attar, D., 2017. Tidal tomography constrains Earth's deep-mantle buoyancy, *Nature*, **551**(7680), 321–326.
- Li, X.-D., Giardini, D. & Woodhouse, J.H., 1991. Large-scale three-dimensional even-degree structure of the Earth from splitting of long-period normal modes, *J. geophys. Res.*, **96**(B1), 551–577.
- Masters, G., Laske, G. & Gilbert, F., 2000. Matrix autoregressive analysis of free-oscillation coupling and splitting, *Geophys. J. Int.*, **143**(2), 478–489.
- Megnin, C. & Romanowicz, B., 1995. Estimation of inner core mode splitting functions combining a genetic algorithm with a direct iterative inversion scheme, *EOS, Trans. Am. geophys. Un.*, **76**, 355.
- Mooney, W.D., Laske, G. & Masters, T.G., 1998. CRUST 5.1: A global crustal model at 5° × 5°, *J. geophys. Res.*, **103**(B1), 727–747.
- Pachhai, S., Tkalcic, H. & Masters, G., 2016. Estimation of splitting functions from Earth's normal mode spectra using the neighbourhood algorithm, *Geophys. J. Int.*, **204**(1), 111–126.
- Pratt, R.G., Shin, C. & Hick, G., 1998. Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion, *J. geophys. Int.*, **133**(2), 341–362.
- Resovsky, J. & Ritzwoller, M., 1998. New and refined constraints on three-dimensional Earth structure from normal modes below 3 mHz, *J. geophys. Res.*, **103**(B1), 783–810.
- Resovsky, J. & Ritzwoller, M., 1999. A degree 8 mantle shear velocity model from normal mode observations below 3 mHz, *J. geophys. Res.*, **104**(B1), 993–1014.
- Resovsky, J. & Trampert, J., 2003. Using probabilistic seismic tomography to test mantle velocity–density relationships, *Earth planet. Sci. Lett.*, **215**(1–2), 121–134.
- Rickers, F., Fichtner, A. & Trampert, J., 2013. The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: evidence from full-waveform inversion, *Earth planet. Sci. Lett.*, **367**, 39–51.
- Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., 1999. Complex shear wave velocity structure imaged beneath Africa and Iceland, *Science*, **286**(5446), 1925–1928.
- Ritsema, J., Deuss, A., Van Heijst, H.J. & Woodhouse, J.H., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.*, **184**(3), 1223–1236.
- Romanowicz, B., 2001. Can we resolve 3D density heterogeneity in the lower mantle? *Geophys. Res. Lett.*, **28**(6), 1107–1110.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Vol. **89**, SIAM.
- Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.*, **20**(2), 219–232.
- Trampert, J., Deschamps, F., Resovsky, J. & Yuen, D., 2004. Probabilistic tomography maps chemical heterogeneities throughout the lower mantle, *Science*, **306**(5697), 853–856.
- Wessel, P., Smith, W.H., Scharroo, R., Luis, J. & Wobbe, F., 2013. Generic mapping tools: improved version released, *EOS, Trans. Am. Geophys. Un.*, **94**(45), 409–410.



- Woodhouse, J., 1980. The coupling and attenuation of nearly resonant multiplets in the Earth's free oscillation spectrum, *Geophys. J. Int.*, **61**(2), 261–283.
- Woodhouse, J. & Dahlen, F., 1978. The effect of a general aspherical perturbation on the free oscillations of the Earth, *Geophys. J. Int.*, **53**(2), 335–354.
- Yang, H.-Y. & Tromp, J., 2015. Synthetic free-oscillation spectra: an appraisal of various mode-coupling methods, *Geophys. J. Int.*, **203**(2), 1179–1192.

## SUPPORTING INFORMATION

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

**Figure S1.** Model slices showing even degrees 2, 4, 6, 8 for (a) starting model S20RTS; (b) resulting model for the one-step inversion; (c) one-step inversion result starting from PREM (as shown in the main paper).

**Figure S2.** Model slices showing even degrees 2, 4, 6, 8 for (a) starting model S20RTS; (b) resulting model for the two-step

inversion; (c) two-step inversion result starting from PREM (as shown in the main paper).

**Figure S3.** Model slices showing even degrees 2, 4, 6, 8 for (a) random starting model; (b) resulting model for the one-step inversion; (c) one-step inversion result starting from PREM (as shown in the main paper).

**Figure S4.** Model slices showing even degrees 2, 4, 6, 8 for (a) random starting model; (b) resulting model for the two-step inversion; (c) two-step inversion result starting from PREM (as shown in the main paper).

**Figure S5.** L-curves of starting models PREM (red), S20RTS (black) and random model (grey) and both inversion methods for model size versus average spectral misfit.

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