Joint inversion of normal-mode and finite-frequency S-wave data using an irregular tomographic grid, Geophysical Journal International

Christophe Zaroli, Sophie Lambotte, Jean-Jacques Lévêque

Geophysical Journal International

Geophys. J. Int. (2015) **203,** 1665–1681
GJI Seismology

doi: 10.1093/gji/ggv388

Joint inversion of normal-mode and finite-frequency *S*-wave data using an irregular tomographic grid

Christophe Zaroli, Sophie Lambotte and Jean-Jacques Lévêque

Institut de Physique du Globe de Strasbourg, UMR 7516*, Université de Strasbourg, EOST/CNRS, Strasbourg, France. E-mail: c.zaroli@unistra.fr*

Accepted 2015 September 11. Received 2015 September 10; in original form 2015 June 3

SUMMARY

Global-scale tomographic models should aim at satisfying the full seismic spectrum. For this purpose, and to better constrain isotropic 3-D variations of shear velocities in the mantle, we tackle a joint inversion of spheroidal normal-mode structure coefficients and multiple-frequency *S*-wave delay times. In all previous studies for which normal modes were jointly inverted for, with body and/or surface waves, the mantle was laterally parametrized with uniform basis functions, such as spherical harmonics, equal-area blocks and evenly spaced spherical splines. In particular, spherical harmonics naturally appear when considering the Earth's free oscillations. However, progress towards higher resolution joint tomography requires a movement away from such uniform parametrization to overcome its computational inefficiency to adapt to local variations in resolution. The main goal of this study is to include normal modes into a joint inversion based upon a non-uniform parametrization that is adapted to the spatially varying smallest resolving length of the data. Thus, we perform the first joint inversion of normal-mode and body-wave data using an irregular tomographic grid, optimized according to ray density. We show how to compute the projection of 3-D sensitivity kernels for both data sets onto our parametrization made up of spherical layers spanned with irregular Delaunay triangulations. This approach, computationally efficient, allows us to map into the joint model multiscale structural informations from data including periods in the 10–51 s range for body waves and 332–2134 s for normal modes. Tomographic results are focused on the 400–2110 km depth range, where our data coverage is the most relevant. We discuss the potential of a better resolution where the grid is fine, compared to spherical harmonics up to degree 40, as the number of model parameters is similar. Our joint model seems to contain coherent structural components beyond degree 40, such as those related to the Farallon subduction. Assessing their robustness is postponed to a future work. A wider application of this tomographic workflow, holding promise to better understand mantle dynamics at various spatial scales, should primarily consist in adding surface-wave data and extending our sets of normal-mode and body-wave data.

Key words: Inverse theory; Body waves; Surface waves and free oscillations; Seismic tomography.

1 INTRODUCTION

It is crucial to build global-scale tomographic models of the Earth's mantle satisfying the full seismic spectrum, since both high and low frequencies matter to improve the overall resolution. For this purpose, to better constrain 3-D velocity variations in the mantle, seismologists should attempt to exploit all the structural information contained in body-wave, surface-wave and normal-mode data.

The type of model parametrization used in a tomographic experiment inherently limits the size of variations in the resulting solution. The physics behind the data should ultimately guide us towards a suitable model parametrization. For instance, spherical harmonics naturally appear when considering free oscillations of the Earth. To the best of our knowledge, in all previous tomographic studies for which normal modes were jointly inverted for, with body and/or surface waves, the spatial variability in the data's resolving power has been ignored by laterally parametrizing the Earth's mantle with uniform basis functions, such as spherical harmonics (e.g. Ritsema *et al.* 1999, 2011), equal-area blocks (Masters *et al.* 2000a) or evenly spaced spherical splines (Moulik & Ekström 2014).

The worldwide network of digital seismometers has continuously been expanding in recent years, including permanent and temporary stations on land or islands, ocean-bottom seismometers and the newly developed submarines MERMAIDS (e.g. Simons

et al. 2006). Significative tomographic improvements could come from better exploiting the present-day receivers coverage (e.g. Zaroli *et al.* 2014). However, since the spatial distribution of earthquakes remains similar through time, global tomography based on body and/or surface waves has to cope with strongly inhomogeneous earthquakes–receivers distributions. Uneven data coverage has led to the use of non-uniform parametrizations for body- and/or surface-wave tomography (e.g. Michelini 1995; Spakman & Bijwaard 2001; Montelli *et al.* 2004b; Nolet 2008; Simmons *et al.* 2009; Rawlinson *et al.* 2010; Zaroli *et al.* 2013). For example, such parametrizations can be made up of irregularly spaced blocks, tuned to the assumed local resolving power of the data. While such irregular parametrizations attempt to maximize the extraction of structural information from the data, the computational price to pay may be significant. Sophisticated computational algorithms are required for building, storing and searching through an irregular mesh (e.g. Sambridge & Rawlinson 2005). Is it worth using an irregular parametrization? One may suspect that uniform parametrizations do not extract all the available structural information contained in body- and/or surface-wave data because the minimum scale length (e.g. highest spherical-harmonic degree, size of regular blocks) is always chosen as a compromise between the data constraints and the computational convenience.

In this study, to better constrain isotropic 3-D variations of shear velocities in the mantle, we tackle a joint inversion of spheroidal normal-mode structure coefficients and multiple-frequency *S*-wave delay times. We shall present an irregular parametrization approach, aimed at fully exploiting the structural information in both normal-mode and body-wave data, whose sensitivity to mantle structure strongly differs in terms of resolution lengths. That is, we believe that progress towards higher resolution joint tomography requires a movement away from uniform parametrization, such as spherical harmonics, to overcome its computational inefficiency to adapt to local variations in resolution. Therefore, the cornerstone of this study will be to include normal-mode data into a joint inversion based upon a non-uniform parametrization that is locally adapted to the spatially varying resolving length of the data. A computationally efficient way for capturing the large range of scale lengths contained in both normal-mode and body-wave data is to use an irregular tomographic grid, optimized according to ray density and made up of spherical layers spanned with irregular Delaunay triangulations. We shall show that using such an irregular parametrization allows us to map into the joint model multiscale structural information from data including periods in the 10–51 s range for body waves and 332–2134 s for normal modes. In particular, we shall discuss the potential of a better resolution where the grid is fine, compared to spherical harmonics up to degree 40, as the number of model parameters is similar. In the first parts of this paper (Sections 2–5), all the technical ingredients for setting up and solving our joint inverse problem are presented. The last part (Section 6) consists in analysing the obtained tomographic model, the first one to be derived from normal-mode and finite-frequency body-wave data, while using an irregular model parametrization.

2 DATA SETS AND FORWARD PROBLEMS

We aim to build a tomographic model of the whole mantle, $m(\mathbf{r})$, representing isotropic 3-D variations of the shear-velocity parameter, $\beta(\mathbf{r})$, with respect to a radial reference model, $\beta_0(r)$, so that:

$$m(\mathbf{r}) = \beta(\mathbf{r})/\beta_0(r) - 1. \quad (1)$$

Table 1. Number of multiple-frequency *S* and *SS* cross-correlation delay times used in this study.

Period	10 s	15 s	22 s	34 s	51 s
S	15 739	31 264	43 263	43 263	35 457
SS	2763	14 518	36 142	36 142	27 807

The spatial location, $\mathbf{r}$, of a given point inside the Earth will be expressed with either spherical (radius r, colatitude θ, longitude ϕ) or Cartesian coordinates. Shear-velocity anomalies will alternatively be denoted as $\delta \ln \beta(\mathbf{r})$, and IASP91 will be used as the reference 1-D model (Kennett & Engdahl 1991). In the following, we shall present our data sets and their associated forward problems.

2.1 Multiple-frequency *S*-wave delay times

To better constrain the 3-D structure of the Earth's mantle, innovative theoretical developments on seismic wave propagation have recently received increasing attention in tomography. In the last decade, taking into account the finite-frequency behaviour of body waves has proved to pay back for better constraining tomographic models at various spatial scales (e.g. Hung *et al.* 2004; Montelli *et al.* 2004a, 2006; Sigloch *et al.* 2008; Tian *et al.* 2011; Zaroli *et al.* 2013). This interest for finite-frequency tomography has been supported by continued evidence for body-wave traveltime dispersion related to different forms of scattering (e.g. Hung *et al.* 2004; Sigloch & Nolet 2006; Yang *et al.* 2006; Zaroli *et al.* 2010; Schuberth *et al.* 2012, 2015). In this study, we aim at exploiting a set of frequency-dependent *S*-wave delay times, using the finite-frequency approach of Dahlen *et al.* (2000) to compute the corresponding 3-D sensitivity kernels. A delay time, δt, is the time-lag maximizing the cross-correlation of an observed waveform with its corresponding ray-theoretical synthetic waveform, both filtered around a central period, T, so that δt depends upon T. The forward problem is linear (e.g. Nolet 2008):

$$d_i^{\mathcal{B}} = \oint_{\mathbf{r}\in V_i} K_i^{\mathcal{B}}(\mathbf{r})\, m(\mathbf{r})\, \mathrm{d}^3\mathbf{r}, \quad (2)$$

where the datum $d_i^{\mathcal{B}}$ represents the ith delay time. Note that $\mathcal{B}$ stands for 'body wave'. The kernel $K_i^{\mathcal{B}}(\mathbf{r})$ also depends upon T. In practice, we compute the kernel using the analytical formulas derived by Zaroli *et al.* (2013) for a Gaussian source power spectrum, over a volume V_i where its amplitude is significant. Since at each period an observed waveform is influenced by a different weighted average of the Earth, through its corresponding kernel, taking into account multiple-frequency delay times should increase the number of independent informations in the inverse problem and allow us to better constrain short-scale seismic features in the mantle (e.g. Zaroli *et al.* 2010; Mercerat *et al.* 2014; Maceira *et al.* 2015). As summarized in Table 1, our body-wave data set consists in 287 078 globally distributed *S* and *SS* cross-correlation delay times measured at 10, 15, 22, 34 and 51 s periods, with individual uncertainty estimates. Traveltime measurements are corrected from effects related to the crust (CRUST2.0; Bassin *et al.* 2000), attenuation (PREM; Dziewonski & Anderson 1981), Earth's ellipticity and topography, as detailed in Zaroli *et al.* (2010, 2013).

2.2 Spheroidal normal-mode structure coefficients

Free oscillations of the Earth, excited by earthquakes with large magnitude, $M_w \geq 7.0$, can provide significative constraints on the very long wavelength mantle structure, through meticulous

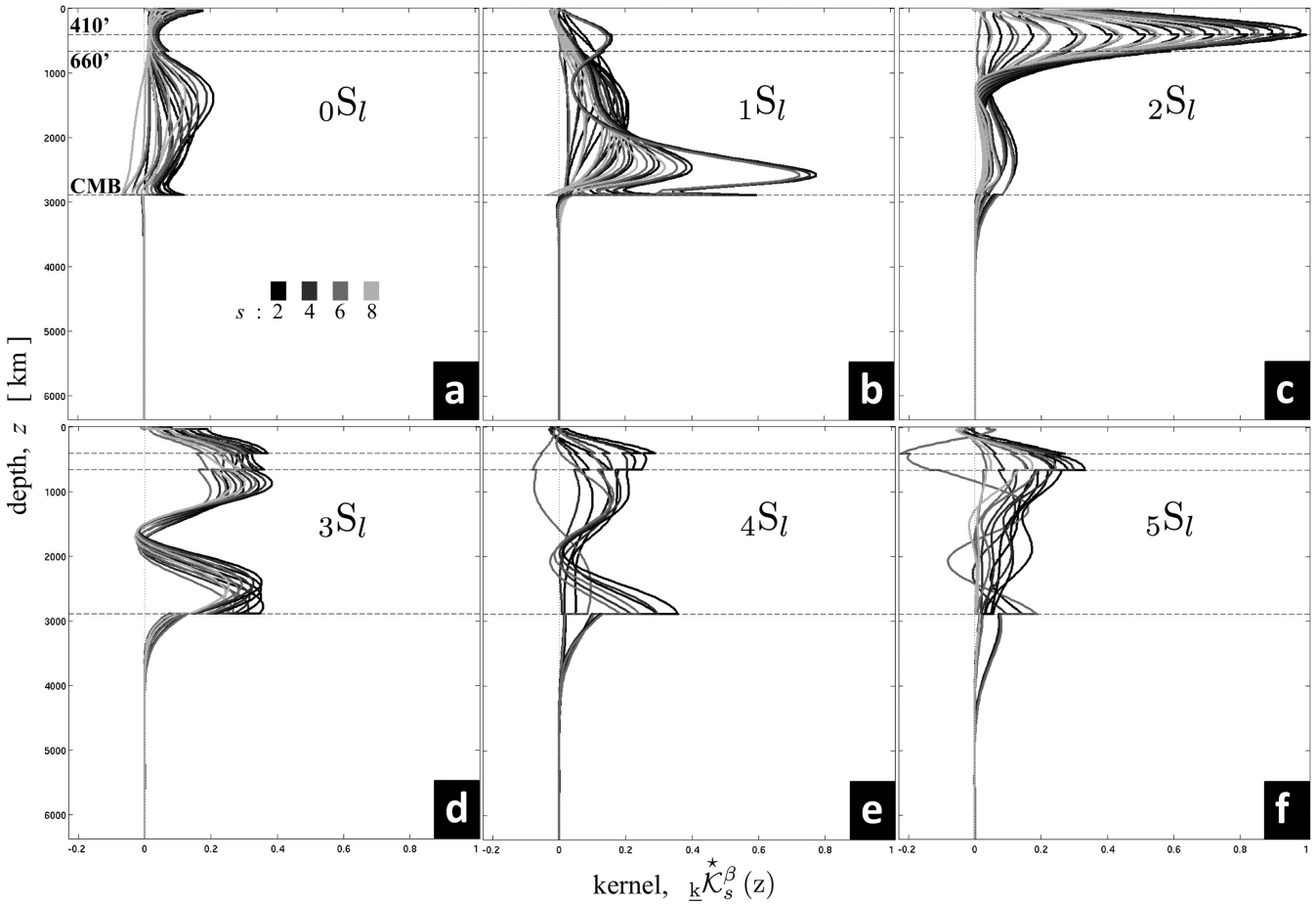

Figure 1. Normalized radial sensitivity kernels corresponding to all the spheroidal normal-mode data used in this study.

splitting measurements of normal-mode spectra (e.g. Resovsky & Ritzwoller 1998; Masters *et al.* 2000b,c; Widmer-Schniring 2002; Deuss *et al.* 2011, 2013). In this study, for simplicity, we shall only treat spheroidal modes, further denoted as:

$$\underline{k} = \{_n\mathrm{S}_l\}\,, \tag{3}$$

where n is the overtone number and l the angular order. Each mode can be associated to a splitting function, $F_{\underline{k}}(\theta,\phi)$, used to visualize the geographical distribution of the radial average of the Earth's 3-D structure as 'seen' by the mode (Woodhouse & Giardini 1985):

$$F_{\underline{k}}(\theta,\phi) = \sum_{s\geq 0}\sum_{t=-s}^{s} {}_{\underline{k}}c_{st}\, Y_{st}(\theta,\phi)\,, \tag{4}$$

where $Y_{st}(\theta,\phi)$ are the complex fully normalized spherical harmonics (Edmonds 1996), with degree s and order t. We shall work in the 'self-coupling' approximation, meaning that modes are considered as isolated from their neighbours and are only sensitive to even-degree structure. The normal-mode structure coefficients, ${}_{\underline{k}}c_{st}$, will represent our tomographic normal-mode data, with $s \in \{2, 4, 6, 8\}$. They are linearly related to the 3-D relative perturbations to a radial reference model in shear (β) and compressional (α) velocities, density (ρ), and to the variations in discontinuity topography ($\delta d/a$), with a the Earth radius. The forward problem is linear (e.g. Dahlen & Tromp 1998):

$${}_{\underline{k}}c_{st} = \int_b^a \xi(r)\,\mathrm{d}r + \sum_d {}_{\underline{k}}\mathcal{K}^d_s\,[\delta d/a]_{st}\,, \tag{5}$$

where we have set:

$$\xi(r) = {}_{\underline{k}}\mathcal{K}^{\beta}_{s}(r)\, m^{\beta}_{st}(r) + {}_{\underline{k}}\mathcal{K}^{\alpha}_{s}(r)\, m^{\alpha}_{st}(r) + {}_{\underline{k}}\mathcal{K}^{\rho}_{s}(r)\, m^{\rho}_{st}(r)\,, \tag{6}$$

with ${}_{\underline{k}}\mathcal{K}^{\beta|\alpha|\rho}_{s}(r)$ the radial sensitivity kernels, in terms of either β, α or ρ, respectively (e.g. Woodhouse & Dahlen 1978). Since we aim at using modes whose radial sensitivity is almost confined to the mantle (see Fig. 1), the integration in eq. (5) starts from the CMB (core–mantle boundary) radius, b. The summation over the index d in eq. (5) includes all seismic discontinuities in the reference 1-D model, typically the free surface, 410', 660' and CMB. The spherical harmonic decomposition of the mantle structure is:

$$m^{\beta|\alpha|\rho}_{st}(r) = \iint_{(\theta,\phi)\in\Omega} m^{\beta|\alpha|\rho}(r,\theta,\phi)\, Y^{*}_{st}(\theta,\phi)\,\mathrm{d}\Omega\,, \tag{7}$$

where Ω denotes the unit sphere, * the complex conjugate and $\mathrm{d}\Omega$ the surface element: $\mathrm{d}\Omega = \sin\theta\mathrm{d}\theta\mathrm{d}\phi$. To obtain a 3-D tomographic model of $\delta\ln\beta$, we opt for reducing the number of physical parameters. The effects of boundary topography are neglected (except for crustal corrections) and the relative variations in density and compressional velocity are related to those in shear-velocity through the scaling factors ν_ρ and ν_α (e.g. Resovsky & Ritzwoller 1999; Ishii & Tromp 2001):

$$\nu_\rho = \delta\ln\rho/\delta\ln\beta \simeq 0.2, \quad \nu_\alpha = \delta\ln\alpha/\delta\ln\beta \simeq 0.55. \tag{8}$$

Therefore, for the purpose of this study, eq. (5) can be rewritten as:

$${}_{\underline{k}}\mathrm{c}_{st} = \int_b^a {}_{\underline{k}}\mathcal{K}^{\star\,\beta}_{s}(r)\, m^{\beta}_{st}(r)\,\mathrm{d}r, \tag{9}$$

Table 2. Spheroidal normal modes $\{{}_nS_l\}$ used in this study; n is the overtone number and l the angular order.

$n = 0$	$\{{}_0S_3\ {}_0S_4\ {}_0S_5\ {}_0S_6\ {}_0S_7\ {}_0S_8\ {}_0S_9\}$
$n = 1$	$\{{}_1S_2\ {}_1S_3\ {}_1S_4\ {}_1S_5\ {}_1S_6\ {}_1S_7\ {}_1S_8\ {}_1S_9\ {}_1S_{10}\ {}_1S_{14}\}$
$n = 2$	$\{{}_2S_4\ {}_2S_5\ {}_2S_6\ {}_2S_7\ {}_2S_8\ {}_2S_9\ {}_2S_{10}\ {}_2S_{11}\ {}_2S_{12}\ {}_2S_{13}\}$
$n = 3$	$\{{}_3S_6\ {}_3S_7\ {}_3S_8\ {}_3S_9\}$
$n = 4$	$\{{}_4S_2\ {}_4S_3\ {}_4S_4\ {}_4S_5\}$
$n = 5$	$\{{}_5S_3\ {}_5S_4\ {}_5S_5\ {}_5S_6\}$

where we have set:

$$\underset{k}{\overset{\star}{\mathcal{K}}}{}_s^{\beta}(r) = \underset{k}{\mathcal{K}}{}_s^{\beta}(r) + \underset{k}{\mathcal{K}}{}_s^{\alpha}(r)v_{\alpha} + \underset{k}{\mathcal{K}}{}_s^{\rho}(r)v_{\rho}. \tag{10}$$

As seen from eq. (9), spherical harmonics naturally appear when considering free oscillations of the Earth. Our set of normal-mode data consists in 1326 real or imaginary components of structure coefficients. They represent a subset of the splitting function measurements obtained by Deuss *et al.* (2011, 2013), and correspond to 39 spheroidal modes (see Table 2 and Fig. 1); individual uncertainties were estimated using cross-validation. The minimum and maximum periods of the modes used in this work are 332 s and 2134 s, corresponding to ${}_5S_6$ and ${}_0S_3$, respectively (Deuss *et al.* 2013). Splitting functions were measured from large earthquakes spectra, starting from PREM and using non-linear iterative least-square inversion technique (e.g. Tarantola & Valette 1982; Li *et al.* 2001). Thus, we apply corrections to the original structure coefficients aimed at recasting the normal-mode data with respect to IASP91, rather than PREM, as the reference model. We also apply standard crustal and topographic corrections to the measured splitting function coefficients, based on the same crustal model as used to correct our *S*-wave measurements.

3 MODEL PARAMETRIZATION

In this section, we shall quantitatively show that using an irregular tomographic grid, optimized according to ray density, is a more efficient way to exploit the structural information in our body-wave data compared to spherical harmonics up to degree 40, while the number of model parameters is similar.

3.1 Spherical harmonics

Spherical harmonics, $Y_{\ell m}(\theta, \phi)$, form a set of uniform basis functions on the sphere, and thus can be used to laterally parametrize the Earth's interior. The vertical parametrization may consist in various kind of radial functions, $q_z(r)$, such as splines or layers. The continuous tomographic model is given by:

$$m(r, \theta, \phi) = \sum_{z=1}^{z_{\max}} \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} m_{z\ell m} q_z(r)\, Y_{\ell m}(\theta, \phi), \tag{11}$$

where $m_{z\ell m}$ are the unknown complex parameters, that one aims at estimating through the inversion process. The total number of model parameters, $M_{\rm sh}$, is:

$$M_{\rm sh} = z_{\max} (\ell_{\max} + 1)^2. \tag{12}$$

When using spherical harmonics to uniformly parametrize shear-velocity variations over a sphere with radius r, the associated characteristic wavelength $\lambda_{\rm sh}$ can be approximated as:

$$\lambda_{\rm sh} \simeq 2\pi r / \ell_{\max}. \tag{13}$$

The corresponding lateral resolving length, $L_{\rm sh}$, that can theoretically be reached is about half the wavelength (e.g. Nolet 2008):

$$L_{\rm sh} \simeq \lambda_{\rm sh}/2. \tag{14}$$

One should keep in mind that, in practice, the actual value of $L_{\rm sh}$ may be larger, for instance due to data errors, model regularization, non-diagonal resolution matrix (e.g. Lévêque *et al.* 1993; Trampert 1998; Aster *et al.* 2012).

Since spherical harmonics naturally appear when considering the Earth's free oscillations, they have often been employed in tomographic joint inversions of normal modes with body and/or surface waves. For instance, S40RTS is a recent tomographic model of isotropic 3-D shear-velocity variations in the whole mantle, resulting from joint inversion of Rayleigh wave phase velocity, teleseismic shear wave traveltime and normal-mode splitting function measurements (Ritsema *et al.* 2011). It is parametrized laterally with spherical harmonics up to degree $\ell_{\max} = 40$ and vertically with 21 spline functions; its total number of unknowns is 35 301. In such a degree-40 model, the lateral resolving-length that can potentially be achieved, at best, further referred to as $L_{\rm sh40}$, linearly varies from 500 km at the surface to 250 km at the CMB. We shall see that our body-wave data are sensitive to velocity anomalies of lateral extent smaller than $L_{\rm sh40}$ (see Section 3.3). However, decreasing the resolving-length $L_{\rm sh}$ further than $L_{\rm sh40}$, to better exploit our body-wave data, may quickly become computationally inefficient in terms of model size. For instance, if one wishes to reach the resolving-length $L_{\rm sh} = L_{\rm sh40}/2$, one needs to double the highest spherical-harmonic degree, from 40 to 80, and the price to pay is an increase by four of the corresponding number of model parameters.

3.2 Irregular grid

In the following, we briefly present the irregular model parametrization used in this tomographic study (see Fig. 2). The model $\mathbf{m}$ is described by a finite number M of parameters m_j, such that:

$$\mathbf{m} = (m_j)_{1 \le j \le M}. \tag{15}$$

The whole mantle is divided into a set of spherical layers. They are laterally spanned with Delaunay triangulations (e.g. Barber *et al.* 1996), whose spatial distributions are optimized according to ray density—a proxy for the varying *S*-wave resolution length. The parametrization is then made up of spherical triangular prisms, as shown in Fig. 2(a). The pth spherical triangular prism, $\mathcal{T}_p$, is defined by the spatial locations of the three vertices of its top spherical triangle, $\mathcal{F}_p$. The three vertices of $\mathcal{F}_p$, which correspond to three nodes of the tomographic grid, represent three parameters of $\mathbf{m}$. As illustrated in Fig. 2(b), it will prove to be convenient to also index them with respect to prism $\mathcal{T}_p$, that is $m_q^{\mathcal{T}_p}$, with $q = \{1, 2, 3\}$. Each parameter $m_q^{\mathcal{T}_p}$ can be mapped back to its globally indexed, unique, parameter m_j. This index back-mapping is denoted as:

$$(\mathcal{T}_p, q) \rightarrow j. \tag{16}$$

Note that each index j corresponds to several pairs $(\mathcal{T}_p, q)$. Our aim is to compute an optimum spatial distribution of the nodes (or parameters m_j) to maximize the extraction of structural information from our *S*-wave data set, as shown in Fig. 2(c). We follow the non-linear optimization approach by Nolet & Montelli (2005). It mainly consists in generating, within each layer z, a set of nodes $\mathfrak{S}_z$ that

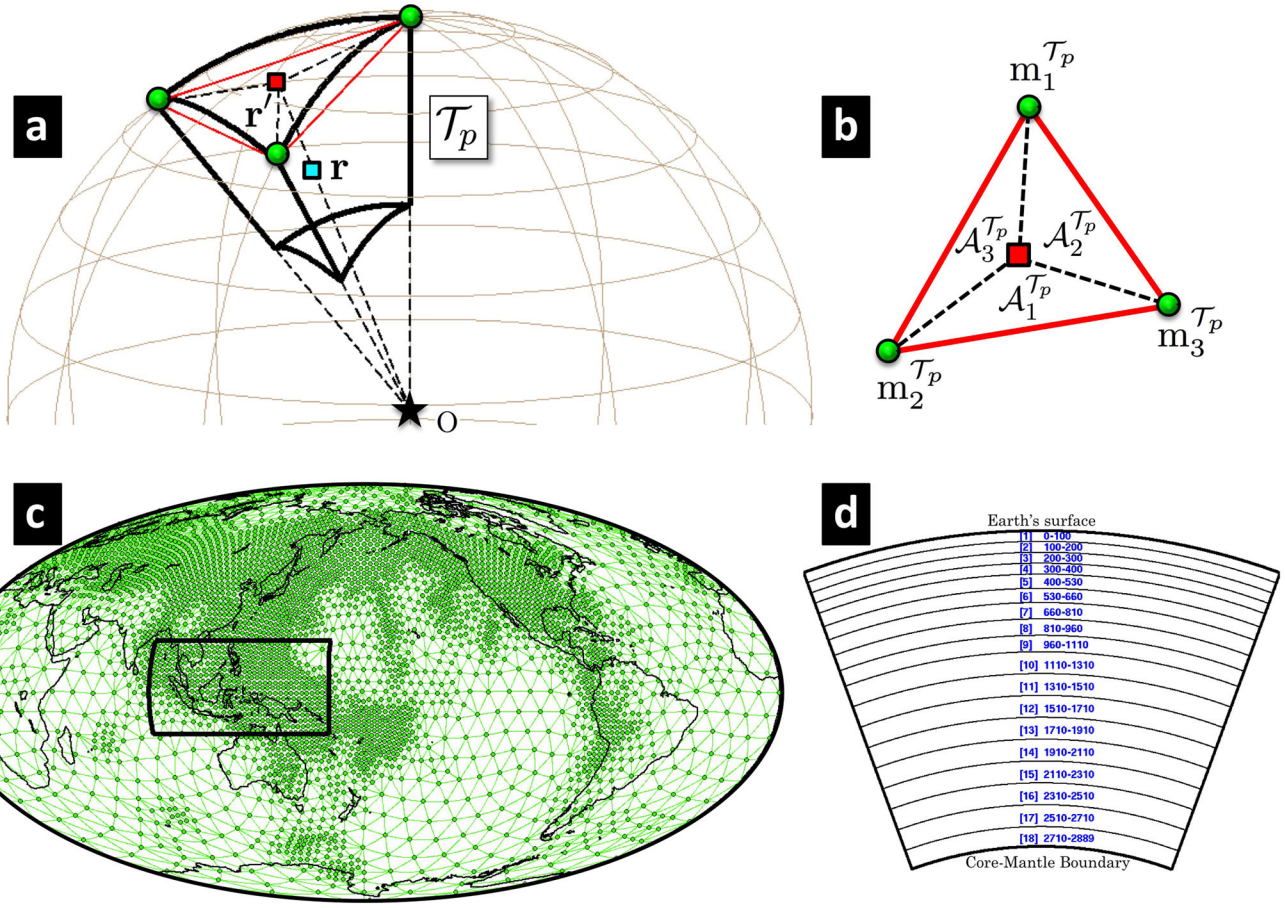

Figure 2. Schematic illustration of the irregular tomographic grid, optimized according to ray density. The parametrization is made up of spherical layers spanned with irregular Delaunay triangulations. (a,b) Spherical triangular prism $\mathcal{T}_p$ enclosing a point $\mathbf{r}$, with $\mathcal{F}_p$ the (black) spherical triangle located at its top. The three vertices (green nodes) of $\mathcal{F}_p$ represent three model parameters, indexed as $\{m_1^{\mathcal{T}_p}, m_2^{\mathcal{T}_p}, m_3^{\mathcal{T}_p}\}$. The point $\mathbf{r}'$ denotes the radial projection of $\mathbf{r}$ onto the (red) planar triangle $\overline{\mathcal{F}}_p$, whose vertices are identical to those of $\mathcal{F}_p$. Barycentric coordinates, $b_q^{\mathcal{T}_p}(\mathbf{r})$, represent normalized subareas of $\overline{\mathcal{F}}_p$, as defined in Section 3.2, where $\{\mathcal{A}_1^{\mathcal{T}_p}, \mathcal{A}_2^{\mathcal{T}_p}, \mathcal{A}_3^{\mathcal{T}_p}\}$ are the three considered subareas. The point O denotes the Earth's centre. (c) Optimized grid of nodes obtained in the ninth layer (960–1110 km depth) and its associated Delaunay mesh; note how it is driven by the irregular spatial coverage of *S*-waves. (d) Radial parametrization: 18 spherical layers whose thickness gently increases with depth, from 100 km thick at the surface to 200 km thick in the lowermost mantle.

is adapted to a given distribution $\mathcal{L}(\mathbf{r})$ of resolving lengths in the mantle, that is, which is minimizing the penalty function:

$$E_z = \sum_{j\in\mathfrak{S}_z}\sum_{k\in\mathbb{N}_j}\left(\frac{\mathcal{D}_{jk}}{\mathcal{L}_{jk}}-1\right)^2, \tag{17}$$

where $\mathbb{N}_j$ is the set of natural-neighbour nodes with respect to j, $\mathcal{D}_{jk}$ the actual distance between j and k, and $\mathcal{L}_{jk}$ the 'average' resolving length between j and k. For global tomographic purposes, Vasco *et al.* (2003) show that the ray density may provide a first-order estimate of $\mathcal{L}(\mathbf{r})$. In general, the heterogeneous earthquakes–receivers distribution leads to closely spaced nodes in Northern Hemisphere, and coarser nodes spacing in Southern Hemisphere. For further details on the nodes layout, the reader is referred to Zaroli (2010) and Zaroli *et al.* (2013). As illustrated in Fig. 2(d), our radial model parametrization consists in 18 spherical layers whose thickness increases with depth, from 100 km thick at the surface to 200 km thick in the lowermost mantle. Note that thinner layers are used in the upper-mantle to anticipate for future joint inversions including surface-wave data.

The model size, that is, the total number of nodes spanning the mantle, is $M = 38\,125$; it is similar to the 35 301 parameters for the degree-40 model S40RTS (see Section 3.1). For such a tomographic grid, the lateral resolving-length which can potentially be achieved, at best, further referred to as L_{grid}, is of the order of the nodes spacing (e.g. Nolet & Montelli 2005). Throughout the mantle, the grid is characterized by a minimum and maximum lateral distance between two adjacent nodes of about 200 and 1000 km, respectively. Thus, in mantle regions with high *S*-wave data coverage, that is, where the grid is fine, the smallest resolving-length L_{grid} that could theoretically be reached is about 200 km. In practice, its actual value may be larger (see Section 3.1).

Finally, the model interpolation consists in relating the value of the continuous model, $m(\mathbf{r})$, at a given spatial location, $\mathbf{r}$, in function of the model parameters. Since our tomographic equations are integrals (see Section 4), a linear interpolation is appropriate, which can be accomplished using the concept of barycentric coordinates (e.g. Montelli *et al.* 2004b; Wu *et al.* 2005; Zaroli 2010):

$$m(\mathbf{r}) = \sum_{q=1}^{3} b_q^{\mathcal{T}_p}(\mathbf{r})\, m_q^{\mathcal{T}_p}, \quad \forall \mathbf{r}\in\mathcal{T}_p. \tag{18}$$

The barycentric coordinates $b_q^{\mathcal{T}_p}(\mathbf{r})$, with $q = \{1, 2, 3\}$, represent normalized weights associated to the three nodes of the prism $\mathcal{T}_p$:

$$\sum_{q=1}^{3} b_q^{\mathcal{T}_p}(\mathbf{r}) = 1, \quad \forall \mathbf{r} \in \mathcal{T}_p. \tag{19}$$

Their use guarantees, at first order, the lateral continuity of the interpolated field. In practice, we compute them as:

$$b_q^{\mathcal{T}_p}(\mathbf{r}) = b_q^{\mathcal{T}_p}(\mathbf{r}') = \frac{\mathcal{A}_q^{\mathcal{T}_p}}{\sum_{q=1}^{3} \mathcal{A}_q^{\mathcal{T}_p}}, \quad \forall \mathbf{r} \in \mathcal{T}_p. \tag{20}$$

The spatial location $\mathbf{r}'$ and the three sub-areas $\mathcal{A}_q^{\mathcal{T}_p}$ are illustrated and defined in Figs 2(a) and (b) and associated caption. As a final remark, note that dealing with a strongly irregular parametrization turns the 'point location problem' into a crucial task in terms of computational cost (e.g. Sambridge & Rawlinson 2005). The reader is referred to Wu *et al.* (2005) and Zaroli (2010) for technical details on how signed barycentric coordinates can efficiently be used for locating a query point in such spherical Delaunay triangulations.

3.3 Uniform versus irregular: *S*-waves beyond degree 40

First, let $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$ denotes our purely body-wave tomographic model obtained using our irregular grid; its derivation will be further discussed in Sections 5 and 6. Let $\mathbf{m}^{\mathcal{B}}_{\text{sh40}}$ represents the model $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$ from which all the spherical-harmonic components with degree superior to 40 have been filtered out. We aim at comparing these two models, $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$ and $\mathbf{m}^{\mathcal{B}}_{\text{sh40}}$, in terms of their respective fit to our body-wave data, $\mathbf{d}^{\mathcal{B}}$. In tomographic experiments, the reduced chi-square functional $\chi^2_{\text{red}}(\mathbf{m}, \mathbf{d})$ is commonly used as a direct measure of misfit of data $\mathbf{d}$ by model $\mathbf{m}$ (e.g. Nolet 2008); its definition will be recalled in Section 5.2. The deterioration of the fit of our body-wave data, when filtering out spherical-harmonic degrees greater than 40 in our body-wave model, can thus be defined and computed as:

$$\chi^2_{\text{red}}\left(\mathbf{m}^{\mathcal{B}}_{\text{sh40}}, \mathbf{d}^{\mathcal{B}}\right) / \chi^2_{\text{red}}\left(\mathbf{m}^{\mathcal{B}}_{\text{grid}}, \mathbf{d}^{\mathcal{B}}\right) - 1 \simeq 14\%. \tag{21}$$

It means that the components of degree superior to 40, in our model $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$, significantly contribute to improve the body-wave data fit. Note that truncating our model $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$ at degree 80 would still result in a 5 per cent data fit deterioration.

Moreover, we aim at comparing the shear-velocity variations of the two models $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$ and $\mathbf{m}^{\mathcal{B}}_{\text{sh40}}$. As an example, Figs 3(a) and (b) displays these models within the transition zone, at 400–530 km depth. At such depth, the minimum values of the lateral resolving-lengths L_{grid} and L_{sh40} are about 200 km and 450 km, respectively (see Sections 3.1 and 3.2). Several trenches of large subduction zones, and their associated slabs, are pointed out in Figs 3(a) and (b). One sees that the prominent high-velocity features, located along the plates boundaries and interpreted as the signature of deep subducting slabs, are better sharpened and imaged in the model $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$. For instance, the short-scale slab feature of New-Hebrides cannot be resolved in the degree-40 model $\mathbf{m}^{\mathcal{B}}_{\text{sh40}}$. Therefore, short-wavelength components in the model solution, corresponding to degrees superior to 40, seem to be relevant in terms of structural interpretation. Figs 3(c) and (d) show the model difference, $\{\mathbf{m}^{\mathcal{B}}_{\text{grid}} - \mathbf{m}^{\mathcal{B}}_{\text{sh40}}\}$, and the corresponding tomographic grid. As expected, seismic features of degree greater than 40 are only retrieved in regions with high ray density, that is, where the grid is fine.

Finally, for the purpose of global finite-frequency tomography, one should aim at exploiting the 'true' kernels, $K_i^{\mathcal{B}}(\mathbf{r})$, to exploit the structural dispersion observed in our frequency-dependent delay times (Zaroli *et al.* 2010). Theoretical kernels are always projected on the model parametrization designed for the tomographic experiment (see Section 4). A fine parametrization is thus needed to capture the short-scale spatial variations of finite-frequency kernels (e.g. Chevrot *et al.* 2012). The *i*th kernel projected onto our irregular tomographic grid, $K^{\mathcal{B}}_{i,\text{grid}}(\mathbf{r}_j)$, at the *j*th node location, $\mathbf{r}_j$, is:

$$K^{\mathcal{B}}_{i,\text{grid}}(\mathbf{r}_j) = \frac{\sum_{(\mathcal{T}_p,q)\to j}\left(\oint_{\mathbf{r}\in(\mathcal{T}_p\cap V_i)} K_i^{\mathcal{B}}(\mathbf{r})\, b_q^{\mathcal{T}_p}(\mathbf{r})\, \mathrm{d}^3\mathbf{r}\right)}{\sum_{(\mathcal{T}_p,q)\to j}\left(\oint_{\mathbf{r}\in(\mathcal{T}_p\cap V_i)} b_q^{\mathcal{T}_p}(\mathbf{r})\, \mathrm{d}^3\mathbf{r}\right)}, \tag{22}$$

where $\sum_{(\mathcal{T}_p,q)\to j}$ means a sum over all the pairs $(\mathcal{T}_p, q)$ satisfying to eq. (16). The continuous kernel $K^{\mathcal{B}}_{i,\text{grid}}(\mathbf{r})$ can be computed from its values at the nodes locations by using the same linear interpolation rule as in eq. (18). Let $K^{\mathcal{B}}_{i,\text{sh40}}(\mathbf{r})$ be the kernel from which all the spherical-harmonic components with degree superior to 40 have been filtered out; it corresponds to the true kernel projected onto a degree-40 spherical-harmonics parametrization. The situation is illustrated in Fig. 4 for a direct *S* wave at 10 and 51 s periods, within the 960–1110 km depth range. At such depth, the minimum values of the lateral resolving lengths L_{grid} and L_{sh40} are about 200 and 400 km, respectively. A visual comparison of the three kernels $K_i^{\mathcal{B}}$, $K^{\mathcal{B}}_{i,\text{grid}}$ and $K^{\mathcal{B}}_{i,\text{sh40}}$ is displayed in Fig. 4. It clearly shows that the finite-frequency *S*-wave kernels, which are effectively used in global tomographic studies in the 10–51 s period range, are significantly less degraded by using our irregular grid rather than spherical-harmonics limited to highest degree 40.

4 SENSITIVITY MATRICES WITH IRREGULAR GRID

As seen in Section 3, progress towards higher resolution joint tomography can be made by using an irregular parametrization, to take advantage of its flexibility to adapt to local variations in the data's resolving-power, while keeping computationally manageable the number of model parameters. In this section, we shall show how to efficiently compute the body-wave and normal-mode sensitivity matrices when using our irregular tomographic grid.

4.1 Body waves

The forward problem is linear, so that it can be written as:

$$d_i^{\mathcal{B}} = \sum_{j=1}^{M} G^{\mathcal{B}}_{i,j} m_j, \quad \text{with } G^{\mathcal{B}}_{i,j} = \frac{\partial d_i^{\mathcal{B}}}{\partial m_j}. \tag{23}$$

In the case of our irregular model parametrization, the elements of the body-wave sensitivity matrix, $G^{\mathcal{B}}_{i,j}$, can be expressed as:

$$G^{\mathcal{B}}_{i,j} = \sum_{(\mathcal{T}_p,q)\to j}\left(\oint_{\mathbf{r}\in(\mathcal{T}_p\cap V_i)} K_i^{\mathcal{B}}(\mathbf{r})\, b_q^{\mathcal{T}_p}(\mathbf{r})\, \mathrm{d}^3\mathbf{r}\right). \tag{24}$$

We compute the volumetric integral in eq. (24) using a simple Riemann sum, that is, by splitting the volumes $(\mathcal{T}_p \cap V_i)$ into regular cubic cells with edges of 20 km, around a regularly spaced grid of points, which are much smaller than the prisms $\mathcal{T}_p$.

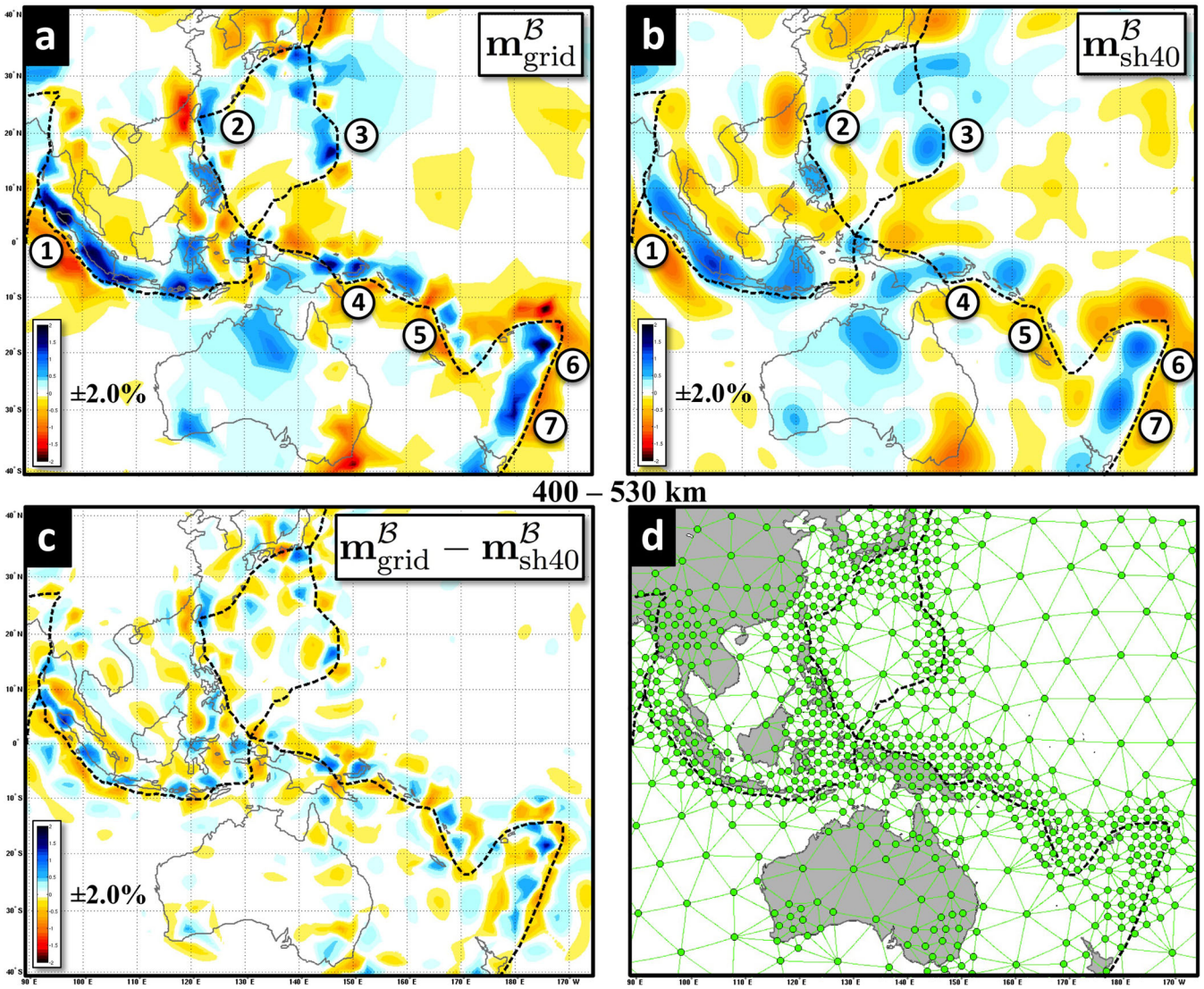

Figure 3. On the relevance of considering *S*-waves beyond degree 40. We focus on a region within the transition zone, at 400–530 km depth, where several trenches of large subduction zones and their associated slabs (high velocities) are pointed out with circled numbers: (1) Java; (2) Ryukya; (3) Mariana; (4) New-Britain/Solomon; (5) New-Hebrides; (6) Tonga; (7) Kermadec. Black dashed lines denote tectonic plate boundaries. (a) Body-wave model $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$, obtained using our irregular tomographic grid. (b) Body-wave model $\mathbf{m}^{\mathcal{B}}_{\text{sh40}}$, representing the model $\mathbf{m}^{\mathcal{B}}_{\text{grid}}$ from which all the spherical-harmonic components with degree superior to 40 have been filtered out. For instance, note that the short-scale slab feature (5) cannot be resolved with the degree 40 limitation. (c) Model difference, $\{\mathbf{m}^{\mathcal{B}}_{\text{grid}} - \mathbf{m}^{\mathcal{B}}_{\text{sh40}}\}$. (d) Corresponding tomographic grid, optimized according to ray density.

4.2 Normal modes

The cornerstone of this work consists in efficiently computing the normal-mode sensitivity matrix when using our irregular grid. For this purpose, we reformulate eq. (9) as:

$$d_i^{\mathcal{N}} = \oint_{(r,\theta,\phi)\in\oplus} K_i^{\mathcal{N}}(r,\theta,\phi)\, m(r,\theta,\phi)\, \mathrm{d}V, \tag{25}$$

where $\mathcal{N}$ stands for 'normal-mode', $\oplus$ denotes the whole mantle in terms of spherical coordinates, $d_i^{\mathcal{N}}$ represents the ith normal-mode datum, $[\underline{\mathrm{k}}\mathrm{c}_{st}]_i$ and $\mathrm{d}V$ is the volume element: $\mathrm{d}V = r^2 \mathrm{d}r\mathrm{d}\Omega$. The Fréchet functional $K_i^{\mathcal{N}}(r,\theta,\phi)$ is the 3-D sensitivity kernel corresponding to the ith normal-mode datum; it can be written as:

$$K_i^{\mathcal{N}}(r,\theta,\phi) = \underline{\mathrm{k}}\overset{\star}{\mathcal{K}}{}_s^{\beta}(r) r^{-2} Y_{st}^*(\theta,\phi). \tag{26}$$

The forward problem is linear, and thus can be expressed as:

$$d_i^{\mathcal{N}} = \sum_{j=1}^{M} G_{i,j}^{\mathcal{N}} m_j, \quad \text{with } G_{i,j}^{\mathcal{N}} = \frac{\partial d_i^{\mathcal{N}}}{\partial m_j}. \tag{27}$$

As a preliminary remark, note that with a parametrization based upon spherical harmonics, the normal-mode sensitivity matrix is:

$$G_{i,z\ell m}^{\mathcal{N}} = \begin{cases} \int_b^a \underline{\mathrm{k}}\, \overset{\star}{\mathcal{K}}{}_s^{\beta}(r) q_z(r) \mathrm{d}r & \text{if } (\ell = s, m = t) \\ 0 & \text{otherwise,} \end{cases} \tag{28}$$

where each triplet (z, ℓ, m) refers to one model parameter m_{zlm}. Such a sensitivity matrix, with elements $G_{i,z\ell m}^{\mathcal{N}}$, is highly sparse and straightforward to compute, which are two clear computational

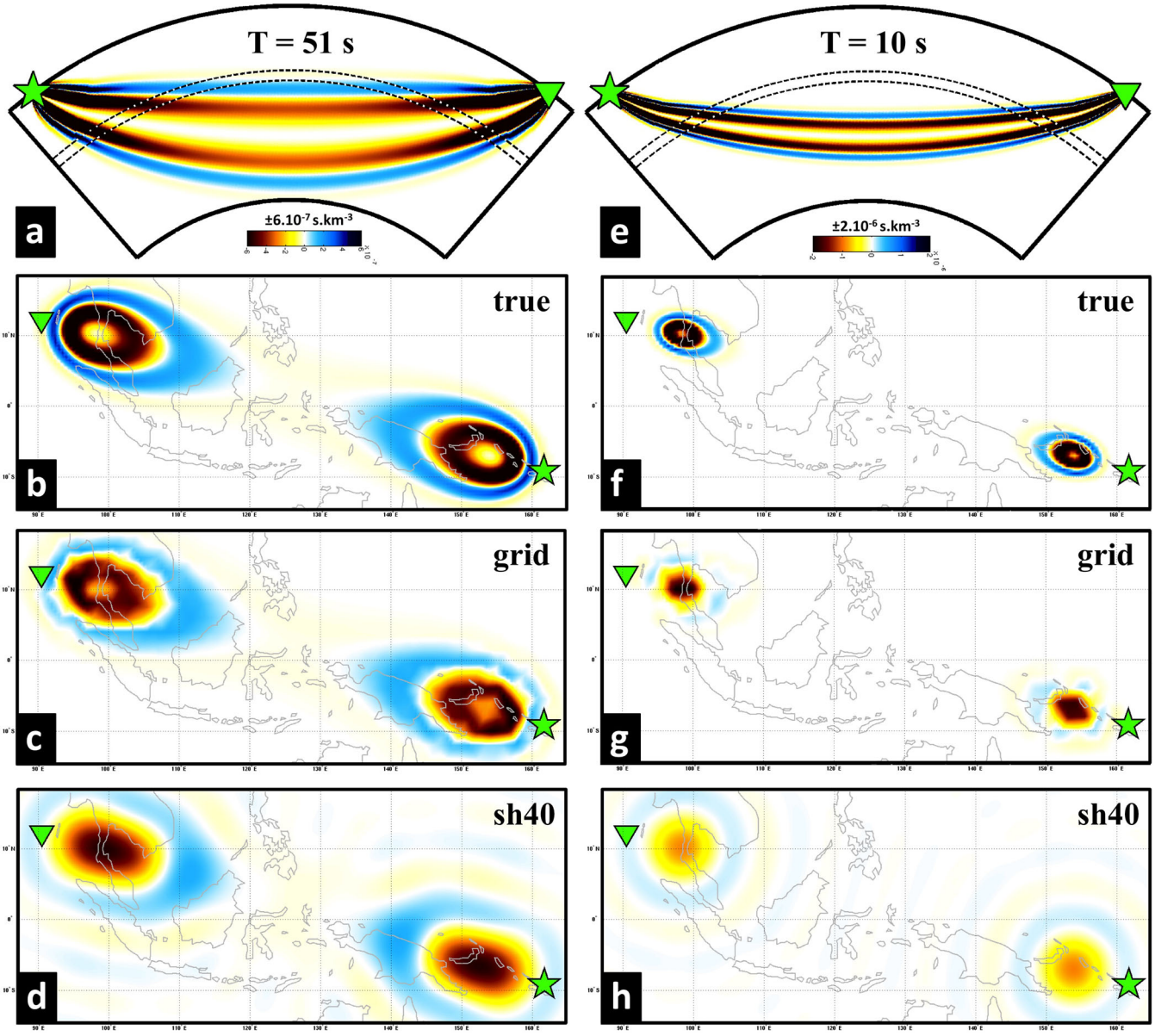

Figure 4. Illustration of the need for a fine model parametrization to capture the short-scale spatial variations of finite-frequency *S*-wave kernels in the 10–51 s period range. (a) Cross-section view of the true kernel, $K^{\mathcal{B}}$, for an *S* phase at 51 s central period, recorded at 74° of epicentral distance. The green star and triangle denote the earthquake and receiver locations, respectively. The ninth layer (960–1110 km depth) of the radial parametrization is marked with dashed lines. (b) Lateral view of the true kernel in the ninth layer. (c) Lateral view of the degraded kernel, $K^{\mathcal{B}}_{\text{grid}}$, after projection of the true kernel onto the irregular grid; the corresponding grid is highlighted with a black frame in Fig. 2(c). (d) Lateral view of the degraded kernel, $K^{\mathcal{B}}_{\text{sh40}}$, after projection of the true kernel onto the spherical harmonics limited to highest degree 40. (e–h) Same as (a–d) but for an *S* phase at 10 s central period, respectively.

advantages. However, when using our irregular grid, the elements $G^{\mathcal{N}}_{i,j}$ of the normal-mode sensitivity matrix are now given by:

$$G^{\mathcal{N}}_{i,j} = \sum_{(\mathcal{T}_p,q)\to j} \left(\oint_{(r,\theta,\phi)\in\mathcal{T}_p} K^{\mathcal{N}}_i(r,\theta,\phi)\, b^{\mathcal{T}_p}_q(r,\theta,\phi)\, \mathrm{d}V \right). \qquad (29)$$

Note that the sensitivity matrix $G^{\mathcal{N}}_{i,j}$ is not sparse anymore and would be highly time consuming if numerically computed with a Riemann sum similar to the one used for the body-wave sensitivity matrix (see Section 4.1). This heavy computational cost can drastically be reduced, by exploiting the fact that the lateral variations of the 3-D normal-mode kernels occur at much longer wavelengths than those characterizing our tomographic grid. That is, their lateral variations are fully constrained by the spherical-harmonics terms, $Y^*_{st}(\theta,\phi)$, which weakly vary inside each prism $\mathcal{T}_p$ when considering structure degree s up to 8 (see Section 2.2). Indeed, the lateral resolving-length associated to degree 8 linearly varies from 2500 km at the surface to 1350 km at the CMB (see Section 3.1), which is much larger than the nodes spacing of our grid (see Section 3.2). One can then reasonably assume that the value of $Y^*_{st}(\theta,\phi)$ remains constant inside each prism $\mathcal{T}_p$, so that the elements $G^{\mathcal{N}}_{i,j}$ of the sensitivity matrix can be fairly approximated as:

$$G^{\mathcal{N}}_{i,j} \simeq \frac{1}{3} \sum_{(\mathcal{T}_p,q)\to j} \left(Y^*_{st}(\theta_p,\phi_p)\mathcal{E}_p \int_{r\in\mathcal{T}_p} \underline{\mathrm{k}}\, \overset{\star}{K}{}^{\beta}_{s}(r)\mathrm{d}r \right), \qquad (30)$$

where (θ_p, ϕ_p) denote the colatitude and longitude of the barycentre point associated to the spherical triangle, $\mathcal{F}_p$, located atop the prism $\mathcal{T}_p$, and where the 'spherical excess' from $\mathcal{F}_p$ is defined as:

$$\mathcal{E}_p = \widehat{\alpha}_p + \widehat{\beta}_p + \widehat{\gamma}_p - \pi, \tag{31}$$

where $(\widehat{\alpha}_p, \widehat{\beta}_p, \widehat{\gamma}_p)$ are the angles at the three vertices of $\mathcal{F}_p$. To formally derive eq. (30), we make use of the Albert Girard's theorem, which gives the surface area $\mathcal{S}$ of a spherical triangle $\mathcal{F}$, with spherical excess $\mathcal{E}$ and lying upon a sphere with radius r, such that:

$$\mathcal{S} = \mathcal{E} r^2. \tag{32}$$

We also make use of the quadrature formula:

$$\oint_{\mathbf{r}\in\mathcal{T}_p} b_q^{\mathcal{T}_p}(\mathbf{r})\,\mathrm{d}^3\mathbf{r} = \mathcal{V}_p/3, \tag{33}$$

where $\mathcal{V}_p$ represents the volume of the prism $\mathcal{T}_p$, that is:

$$\mathcal{V}_p = \mathcal{E}_p \int_{r\in\mathcal{T}_p} r^2 \mathrm{d}r. \tag{34}$$

We compute the integral in eq. (30) using a Riemann sum with a radial step of 10 km. Since the normal-mode data and sensitivity matrix are complex, their real and imaginary parts have to be separately considered to set up the corresponding tomographic equations. We now have at hands all the needed tools to perform such a joint inversion tailored to our irregular grid.

5 JOINT INVERSION

The joint inversion of our normal-mode and body-wave data consists in solving a linear inverse problem of the usual form:

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \tag{35}$$

where $\mathbf{d}$ and $\mathbf{m}$ are the vectors of data (size N) and unknown model parameters (size M), and $\mathbf{G}$ is the sensitivity matrix (size $N \times M$).

5.1 LSQR with irregular grid

Assuming that the prior covariance matrices of the data, $\mathbf{C}_\mathrm{d}$, and model parameters, $\mathbf{C}_m$, follow Gaussian probability functions, the optimum estimate of $\mathbf{m}$ can be obtained by minimizing the following function (e.g. Tarantola & Nercessian 1984; Tarantola 1987):

$$f(\mathbf{m}) = (\mathbf{d}-\mathbf{G}\mathbf{m})^\mathrm{T}\mathbf{C}_\mathrm{d}^{-1}(\mathbf{d}-\mathbf{G}\mathbf{m}) + \mathbf{m}^\mathrm{T}\mathbf{C}_\mathrm{m}^{-1}\mathbf{m}, \tag{36}$$

where $(\cdot)^\mathrm{T}$ and $(\cdot)^{-1}$ are the transpose and inverse operators, respectively. For simplicity reasons, we shall use diagonal data and model covariance matrices of the form:

$$\begin{cases} \mathbf{C}_\mathrm{d} = \underset{i\in[\![1;N]\!]}{\mathrm{diag}}\left(\sigma_{\mathrm{d},i}^2\right) \\ \mathbf{C}_\mathrm{m} = \sigma_\mathrm{m}^2\mathbf{I}_M, \end{cases} \tag{37}$$

where $\sigma_{\mathrm{d},i}$ represent the individually estimated data uncertainties, σ_m denotes the prior model variance and $\mathbf{I}_M$ is an identity matrix (size $M \times M$). Minimizing the function $f(\mathbf{m})$ leads to solving a system of normal equations (e.g. Nolet 2008):

$$\begin{pmatrix}\mathbf{d}'\\ \mathbf{0}\end{pmatrix} = \begin{pmatrix}\mathbf{G}'\\ \Theta\mathbf{I}_M\end{pmatrix}\mathbf{m}, \quad \text{with:} \begin{cases}\mathbf{d}' = \mathbf{C}_\mathrm{d}^{-1/2}\mathbf{d}\\ \mathbf{G}' = \mathbf{C}_\mathrm{d}^{-1/2}\mathbf{G},\end{cases} \tag{38}$$

where the system has been scaled with data uncertainties, to end up with an univariant data vector, and the regularization parameter, Θ, hereafter referred to as the damping parameter, is expressed as:

$$\Theta = 1/\sigma_\mathrm{m}. \tag{39}$$

Inverse problems are usually ill-posed and require a degree of regularization to deal with data errors and stabilize the tomographic solution. For each value of damping Θ, we compute the corresponding model solution $\mathbf{m}_\Theta$ with the LSQR algorithm (e.g. Paige & Saunders 1982; Grunberg 2006). It is an iterative row action method, converging to the least-squares solution of eq. (38):

$$\mathbf{m}_\Theta = \arg\min\left(\|\mathbf{d}'-\mathbf{G}'\mathbf{m}\|^2 + \Theta^2\|\mathbf{m}\|^2\right), \tag{40}$$

where $\|\cdot\|$ denotes the Euclidean norm. The value of Θ influences the solution by damping the model norm, and thus allows us to regularize our inverse problem. Such a simple regularization is sufficient, in our experience, to obtain smooth solutions (Zaroli *et al.* 2013, 2014). The LSQR algorithm implicitly requires that the norm of the discretized model $\mathbf{m}$, involved in the cost function to be minimized in eq. (40), be equal to the norm of the continuous model $\mathrm{m}(\mathbf{r})$. Because of the irregular pattern of our model parametrization, the volume $\mathbb{V}_j$ associated to each model parameter m_j is also irregular, which implies that the LSQR pre-requisite is not met:

$$\|\mathbf{m}\|^2 = \sum_{j=1}^{M}\mathrm{m}_j^2 \neq \frac{1}{\mathcal{V}_\circledcirc}\oint_\circledcirc \mathrm{m}^2(\mathbf{r})\,\mathrm{d}^3\mathbf{r}, \tag{41}$$

where $\circledcirc$ represents the whole mantle, in terms of Cartesian coordinates, and the volume $\mathcal{V}_\circledcirc = \oint_\circledcirc \mathrm{d}^3\mathbf{r}$. Thus, one needs to modify the system of normal equations to be solved by LSQR as follows:

$$\begin{pmatrix}\mathbf{d}'\\ \mathbf{0}\end{pmatrix} = \begin{pmatrix}\mathbf{G}''\\ \Theta\mathbf{I}_M\end{pmatrix}\mathbf{m}', \quad \text{with:} \begin{cases}\mathbf{G}'' = \mathbf{G}'\mathbf{D}_\mathrm{v}\\ \mathbf{m}' = \mathbf{D}_\mathrm{v}^{-1}\mathbf{m},\end{cases} \tag{42}$$

where the diagonal matrix $\mathbf{D}_\mathrm{v}$ is:

$$\mathbf{D}_\mathrm{v} = \underset{j\in[\![1;M]\!]}{\mathrm{diag}}\left(\sqrt{\mathcal{V}_\circledcirc/\mathbb{V}_j}\right). \tag{43}$$

Solving for the new system in eq. (42) with LSQR leads to:

$$\begin{cases}\mathbf{m}'_\Theta = \arg\min(\|\mathbf{d}'-\mathbf{G}''\mathbf{m}'\|^2 + \Theta^2\|\mathbf{m}'\|^2)\\ \mathbf{m}_\Theta = \mathbf{D}_\mathrm{v}\mathbf{m}'_\Theta,\end{cases} \tag{44}$$

which means that we first compute the LSQR solution $\mathbf{m}'_\Theta$ and then multiply it by the matrix $\mathbf{D}_\mathrm{v}$ to end up with the physical model solution $\mathbf{m}_\Theta$. Each term $\mathbb{V}_j$ has to be calculated such that it implies:

$$\|\mathbf{m}'\|^2 = \frac{1}{\mathcal{V}_\circledcirc}\sum_{j=1}^{M}\mathbb{V}_j\mathrm{m}_j^2 = \frac{1}{\mathcal{V}_\circledcirc}\oint_\circledcirc \mathrm{m}^2(\mathbf{r})\,\mathrm{d}^3\mathbf{r}. \tag{45}$$

If the parametrization consisted in non-overlapping blocks, $\mathbb{V}_j$ would simply be the volume of the jth block (e.g. Spakman & Nolet 1988; Spakman & Bijwaard 2001). In this study, the parameters are represented by a set of nodes, and its Delaunay mesh, so that we have to formally derive the proper expression for $\mathbb{V}_j$, and find:

$$\mathbb{V}_j = \frac{1}{3}\sum_{\mathcal{T}_p\in\mathbb{T}_j}\mathcal{V}_p, \tag{46}$$

where $\mathbb{T}_j$ represents the 1-ring set of natural-neighbour prisms with respect to node j, that is, all the prisms $\mathcal{T}_p$ such that one of their three nodes $\{m_1^{\mathcal{T}_p}, m_2^{\mathcal{T}_p}, m_3^{\mathcal{T}_p}\}$ corresponds to m_j. To derive eq. (46), we make use of eq. (33) and of the quadrature formulas:

$$\oint_{\mathbf{r}\in\mathcal{T}_p} b_\mu^{\mathcal{T}_p}(\mathbf{r})\,b_\nu^{\mathcal{T}_p}(\mathbf{r})\,\mathrm{d}^3\mathbf{r} = \begin{cases}\mathcal{V}_p/6 & \text{if} \quad \mu=\nu\\ \mathcal{V}_p/12 & \text{if} \quad \mu\neq\nu.\end{cases} \tag{47}$$

We also make use of the fact that, for a given prism $\mathcal{T}_p$, we have:

$$m_1^{\mathcal{T}_p} \simeq m_2^{\mathcal{T}_p} \simeq m_3^{\mathcal{T}_p}. \tag{48}$$

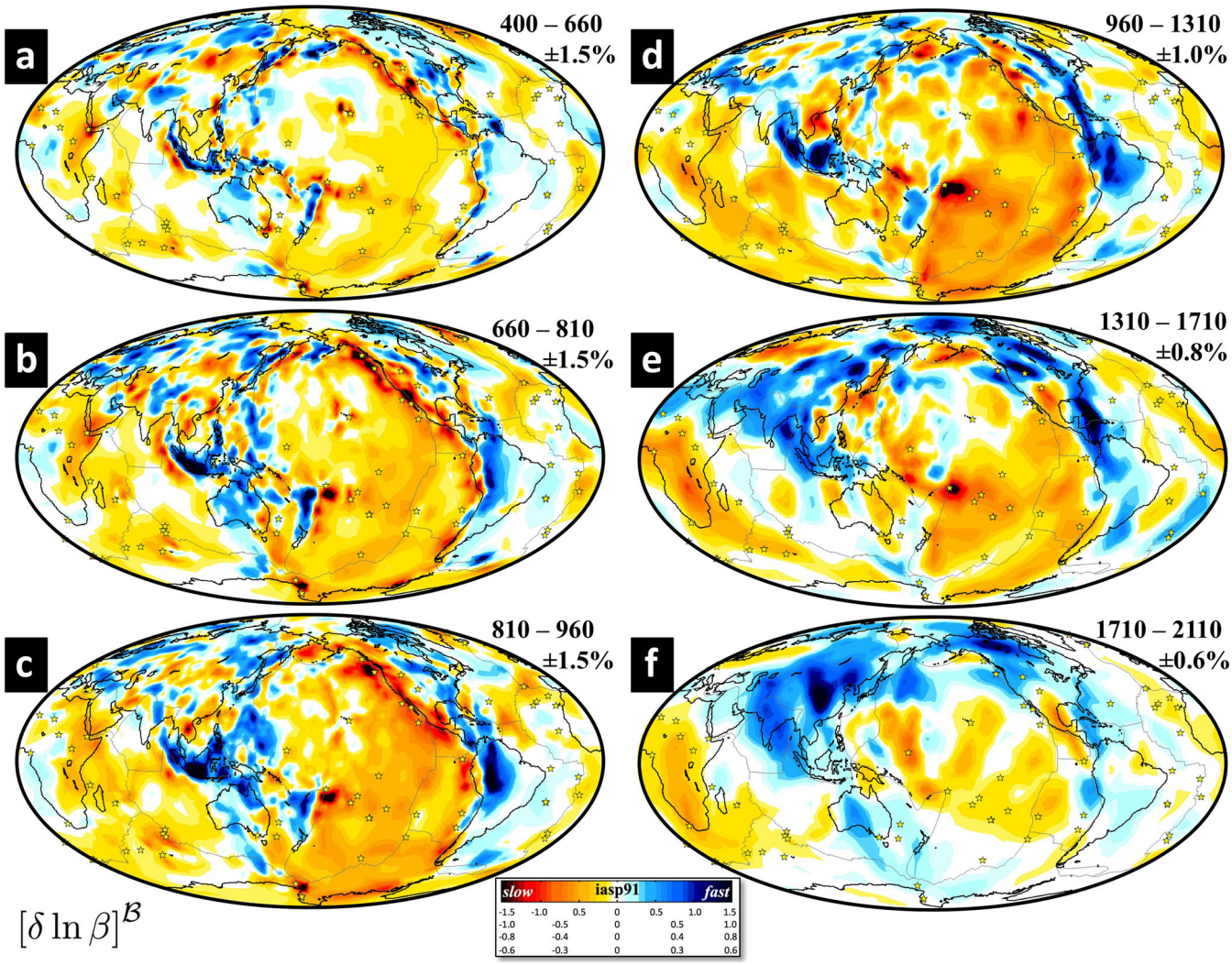

Figure 5. Tomographic model obtained from the inversion of body-wave data, using the irregular grid. Shear-velocity variations are displayed with respect to IASP91, in the 400–2110 km depth range. Note that the colour scale is changing in panels (d)–(f), and that each panel among (a, d, e, f) represents a vertical averaging of the model over two adjacent layers, *cf.* Fig. 2(d). Grey solid line: tectonic plates; yellow stars: hotspots.

As a final remark, note that in eqs (40) and (44) the data misfit term has remained the same, since it is trivial that $\mathbf{G}'\mathbf{m} = \mathbf{G}''\mathbf{m}'$.

5.2 Damping and data weighting

A convenient graphical tool for estimating the damping parameter, Θ, is to perform an L-curve analysis (e.g. Hansen & O'leary 1993; Aster *et al.* 2012; Zaroli *et al.* 2013). It consists in analysing the behaviour of the curve, parametrized by Θ, of trade-off between the continuous model norm, $\|\mathbf{m}'_\Theta\|^2$, and the data misfit, $\chi^2_{\rm red}$. The reduced chi-square functional, $\chi^2_{\rm red}$, is commonly defined in large-scale tomographic experiments as (e.g. Nolet 2008):

$$\chi^2_{\rm red}(\mathbf{m}_\Theta, \mathbf{d}) = \frac{1}{N}\|\mathbf{d}' - \mathbf{G}'\mathbf{m}_\Theta\|^2, \tag{49}$$

where N is the total number of data. Though the factor $1/N$ may slightly differ in the literature (e.g. Trampert & Woodhouse 2003), it is not crucial here as we only argue about ratios of data misfits. If the statistics of data uncertainties were perfectly known, the optimal solution would correspond to $\chi^2_{\rm red} \simeq 1$ near the bend of the L-curve. Data uncertainties are a mix of observational and modelling errors, and in practice are often just best guesses. Therefore, one usually faces the dilemma to choose a solution around the L-curve's corner as a best compromise, with a certain degree of subjectivity, between minimizing the data misfit and the model norm.

First, let us consider separately the inversions of either the body-wave or the normal-mode data. Based upon their respective L-curve shapes, we can estimate relevant values for their damping parameters (i.e. their prior model variances), hereafter referred to as $\Theta^{\mathcal{B}}$ and $\Theta^{\mathcal{N}}$. They are likely to differ, since our high- and low-frequency data sense the mantle structure at very different wavelengths. As shown in Figs 5 and 6, the corresponding body-wave and normal-mode tomographic models seem to be compatible with other studies (see Section 6.2). That is, when inverting for one kind of data set only, the subjectivity inherent to the damping choice can be hampered by *a priori* geophysical considerations of what physically plausible solutions can, or cannot, be.

However, it may be a bigger challenge to apprehend what to expect, or not, from jointly inverting intrinsically different data sets. As one usually has to cope with poorly constrained data uncertainties, it is often not an easy task, in a joint inversion, to estimate an adequate value for the damping, and it may be necessary to

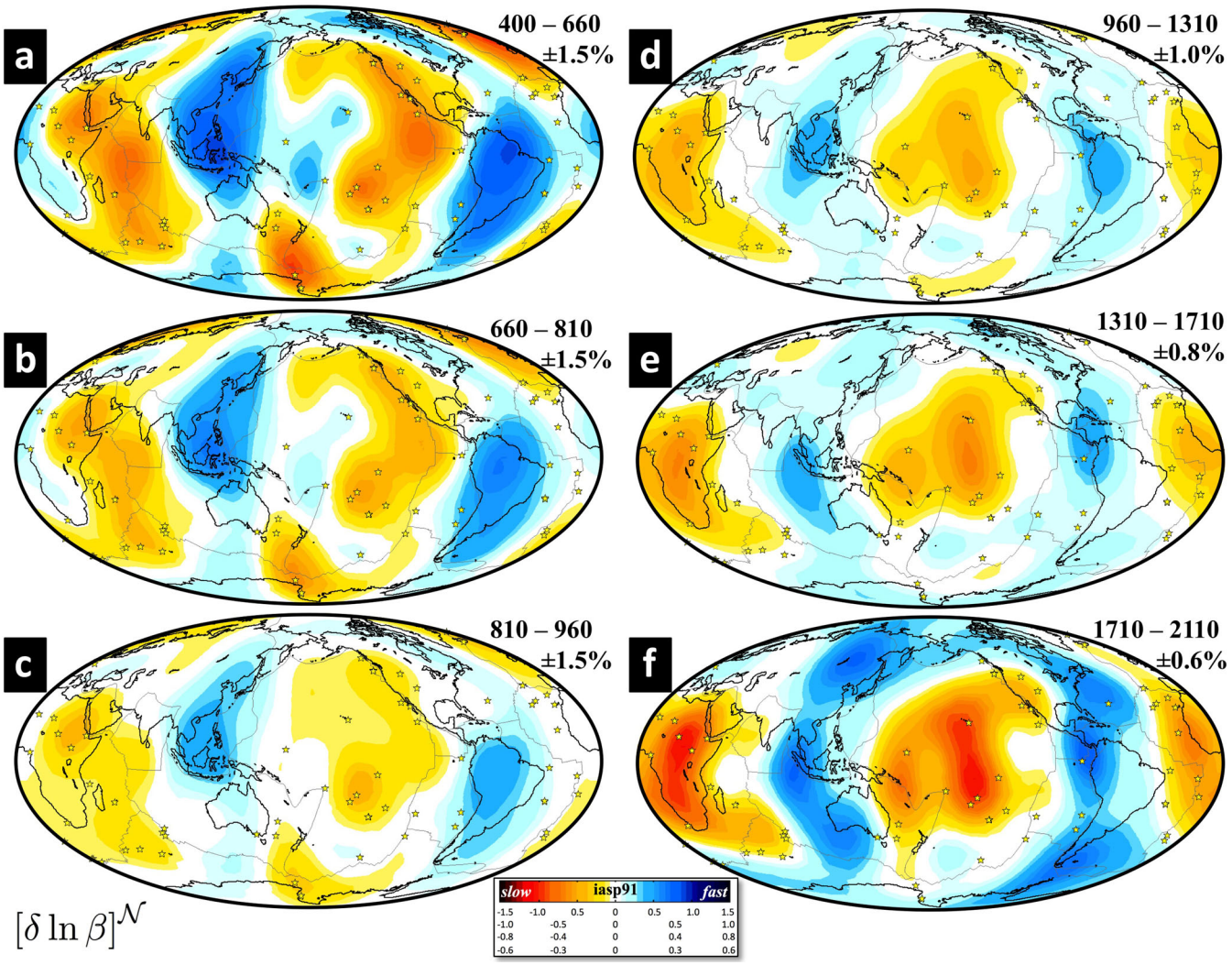

Figure 6. Same legend as in Fig. 5, except that this model results from the inversion of normal-mode data, using the irregular grid.

apply a relative weighting on the univariant data sets. In this study, our choice is to keep unchanged, in the joint inversion, the previously estimated damping parameters, $\Theta^{\mathcal{B}}$ and $\Theta^{\mathcal{N}}$. Since only a single damping value can be used for each LSQR inversion, let us consider, as a mathematical trick to meet our purpose of keeping unchanged both damping parameters, the weighted joint system:

$$\begin{pmatrix} \mathbf{D}_{\mathrm{w}}\mathbf{d}' \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{\mathrm{w}}\mathbf{G}'' \\ \Theta^{\mathcal{J}}\mathbf{I}_M \end{pmatrix} \mathbf{m}', \quad \text{with: } \mathbf{D}_{\mathrm{w}} = \underset{i\in[\![1;N]\!]}{\mathrm{diag}}\,(1/\mathrm{w}_i), \tag{50}$$

where $\mathcal{J}$ stands for 'joint', and the data weights are defined as:

$$\underset{i\in[\![1;N]\!]}{\mathrm{w}_i} = \begin{cases} 1 & \text{if } i \Leftrightarrow \text{body-wave} \\ \mathcal{W} & \text{if } i \Leftrightarrow \text{normal-mode.} \end{cases} \tag{51}$$

The diagonal matrix $\mathbf{D}_{\mathrm{w}}$ only depends upon the value of $\mathcal{W} \in \mathbb{R}_+^*$, which controls the relative weighting of both univariant data sets. Since the normal-mode equations are divided by $\mathcal{W}$, in eq. (50), the corresponding damping value, $\widetilde{\Theta}^{\mathcal{N}}$, is such that:

$$\widetilde{\Theta}^{\mathcal{N}} = \Theta^{\mathcal{N}}/\mathcal{W}. \tag{52}$$

One should then use in eq. (50) the values of $(\Theta^{\mathcal{J}}, \mathcal{W})$ defined as:

$$\begin{cases} \Theta^{\mathcal{J}} = \Theta^{\mathcal{B}} \\ \mathcal{W} = \Theta^{\mathcal{N}}/\Theta^{\mathcal{B}}. \end{cases} \tag{53}$$

Indeed, eqs (52–53) guarantee the aimed property:

$$\Theta^{\mathcal{J}} = \Theta^{\mathcal{B}} = \widetilde{\Theta}^{\mathcal{N}}. \tag{54}$$

As a concluding remark, note that the relative data weighting, $\mathcal{W}$, is here only based on the ratio of the two optimal damping values, $\Theta^{\mathcal{N}}$ and $\Theta^{\mathcal{B}}$, derived from separate inversions. We shall see that our approach to regularize the joint inversion leads to relevant and consistent tomographic results.

6 TOMOGRAPHIC RESULTS

We aim at analysing the first tomographic model to be derived from a joint inversion of normal-mode and body-wave data, while using an irregular parametrization. Three models will be considered: $\mathbf{m}^{\mathcal{B}}$ (inversion of body waves), $\mathbf{m}^{\mathcal{N}}$ (inversion of normal modes) and $\mathbf{m}^{\mathcal{J}}$ (joint inversion). All of them have been obtained using our irregular grid, and they correspond to the damping and/or data weighting parameters discussed in Section 5.2. Displayed tomographic images will be focused on the transition zone and mid lower-mantle, within

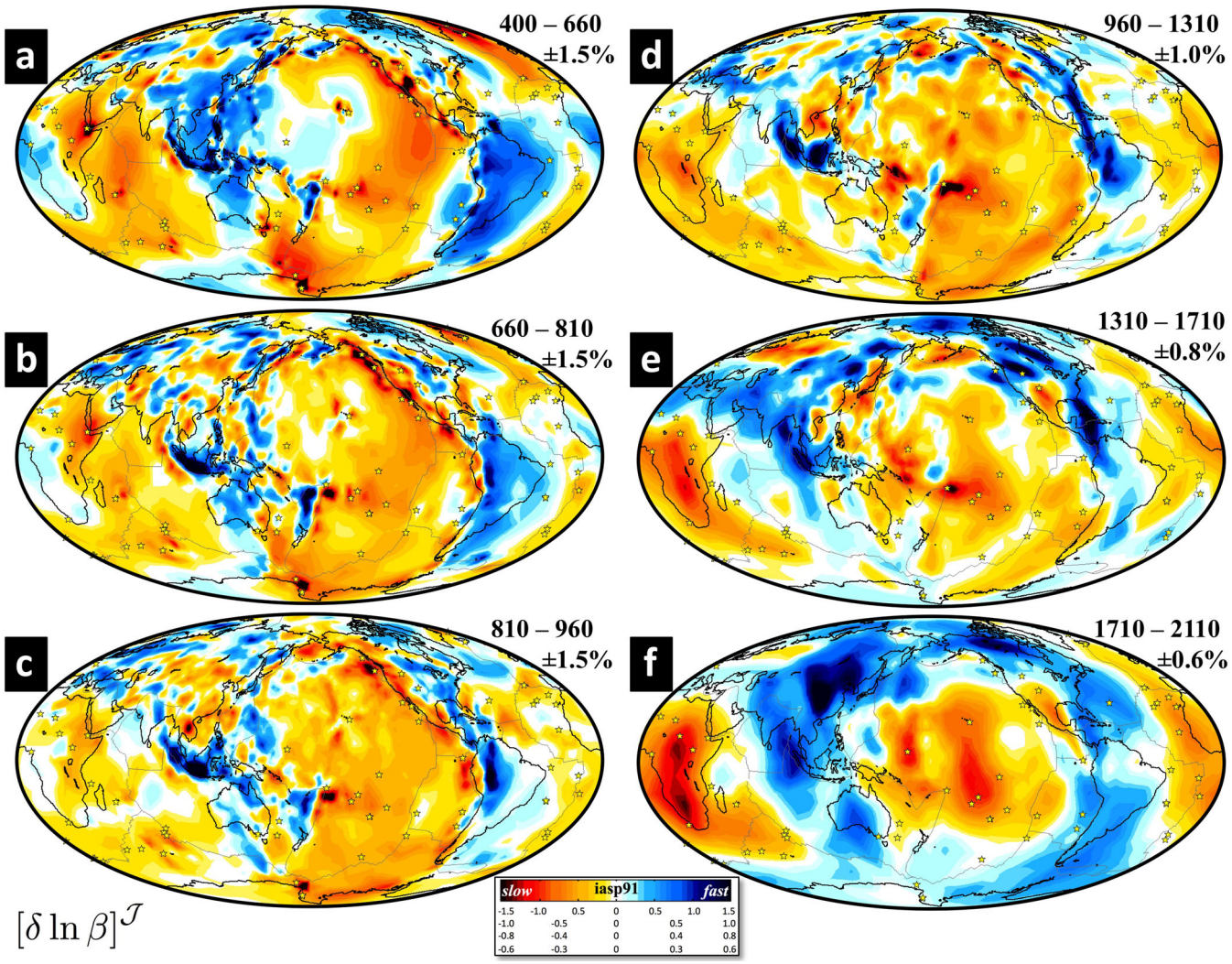

Figure 7. Same legend as in Fig. 5, except that this model results from the joint inversion of body-wave and normal-mode data, using the irregular grid.

the 400–2110 km depth range, where our *S*-wave data coverage is the most relevant (Zaroli *et al.* 2010, 2013).

6.1 Joint model analysis

First, we aim at quantifying the fit deterioration of the body-wave data by the joint model. This can be expressed and computed as:

$$\chi^2_{\rm red}(\mathbf{m}^{\mathcal{J}}, \mathbf{d}^{\mathcal{B}})/\chi^2_{\rm red}(\mathbf{m}^{\mathcal{B}}, \mathbf{d}^{\mathcal{B}}) - 1 \simeq 0.5\% \,. \tag{55}$$

The joint model, $\mathbf{m}^{\mathcal{J}}$, fits the body-wave data, $\mathbf{d}^{\mathcal{B}}$, almost equally well as the body-wave model, $\mathbf{m}^{\mathcal{B}}$, since there is only 0.5 per cent of fit deterioration. Similarly, the fit deterioration of the normal-mode data by the joint model is given by:

$$\chi^2_{\rm red}(\mathbf{m}^{\mathcal{J}}, \mathbf{d}^{\mathcal{N}})/\chi^2_{\rm red}(\mathbf{m}^{\mathcal{N}}, \mathbf{d}^{\mathcal{N}}) - 1 \simeq 3.5\%. \tag{56}$$

The joint model fits very well the normal-mode data, $\mathbf{d}^{\mathcal{N}}$, relatively to the normal-mode model, $\mathbf{m}^{\mathcal{N}}$, as the fit deterioration is no more than 3.5 per cent. The body-wave and normal-mode data seem to be overall nicely compatible, since the joint inversion does not require to drastically deteriorate the fit of one data set to the detriment of the other. Moreover, as expected, the normal-mode data are better fitted by the joint model than by the body-wave model:

$$\chi^2_{\rm red}(\mathbf{m}^{\mathcal{J}}, \mathbf{d}^{\mathcal{N}})/\chi^2_{\rm red}(\mathbf{m}^{\mathcal{B}}, \mathbf{d}^{\mathcal{N}}) \simeq 1/8. \tag{57}$$

Therefore, the joint inversion leads to a better model, from a data fit point of view, than the body-wave inversion, since the body-wave data are almost equally well fitted by both $\mathbf{m}^{\mathcal{J}}$ and $\mathbf{m}^{\mathcal{B}}$, while the normal-mode data are better fitted by $\mathbf{m}^{\mathcal{J}}$ than by $\mathbf{m}^{\mathcal{B}}$.

Furthermore, we aim at better apprehending, in the model space, the effects of incorporating in the joint inversion the normal modes in addition to the body waves. Figs 5–8 show, at global scale, the body-wave model, the normal-mode model, the joint model, and the irregular grid superimposed to the model difference $\{\mathbf{m}^{\mathcal{J}} - \mathbf{m}^{\mathcal{B}}\}$, respectively. In several regions where the grid is fine, meaning high *S*-wave coverage, one sees that incorporating the normal-mode data do imply significative long-wavelength changes in the joint model, with respect to the body-wave model. For instance, Fig. 8 shows that normal modes require a shift towards slightly lower shear-velocities in Eastern-Asia/Western-Pacific, at 810–960 km depth, whereas the grid is fine. As expected, a great majority of the model differences observed in Fig. 8 occur in regions where the grid is coarse, meaning poor *S*-wave data coverage, such as in Central/Eastern South-America within the transition-zone and uppermost lower-mantle. Other examples are numerous in the Southern Hemisphere or the oceans, where the grid is often coarse. In particular, Fig. 9 focuses on the Central Pacific area, in the 1310–1710 km depth range. Though our *S*-waves cannot resolve the surroundings of Hawaii at such depth, where the grid is coarse, a large-scale low-velocity feature

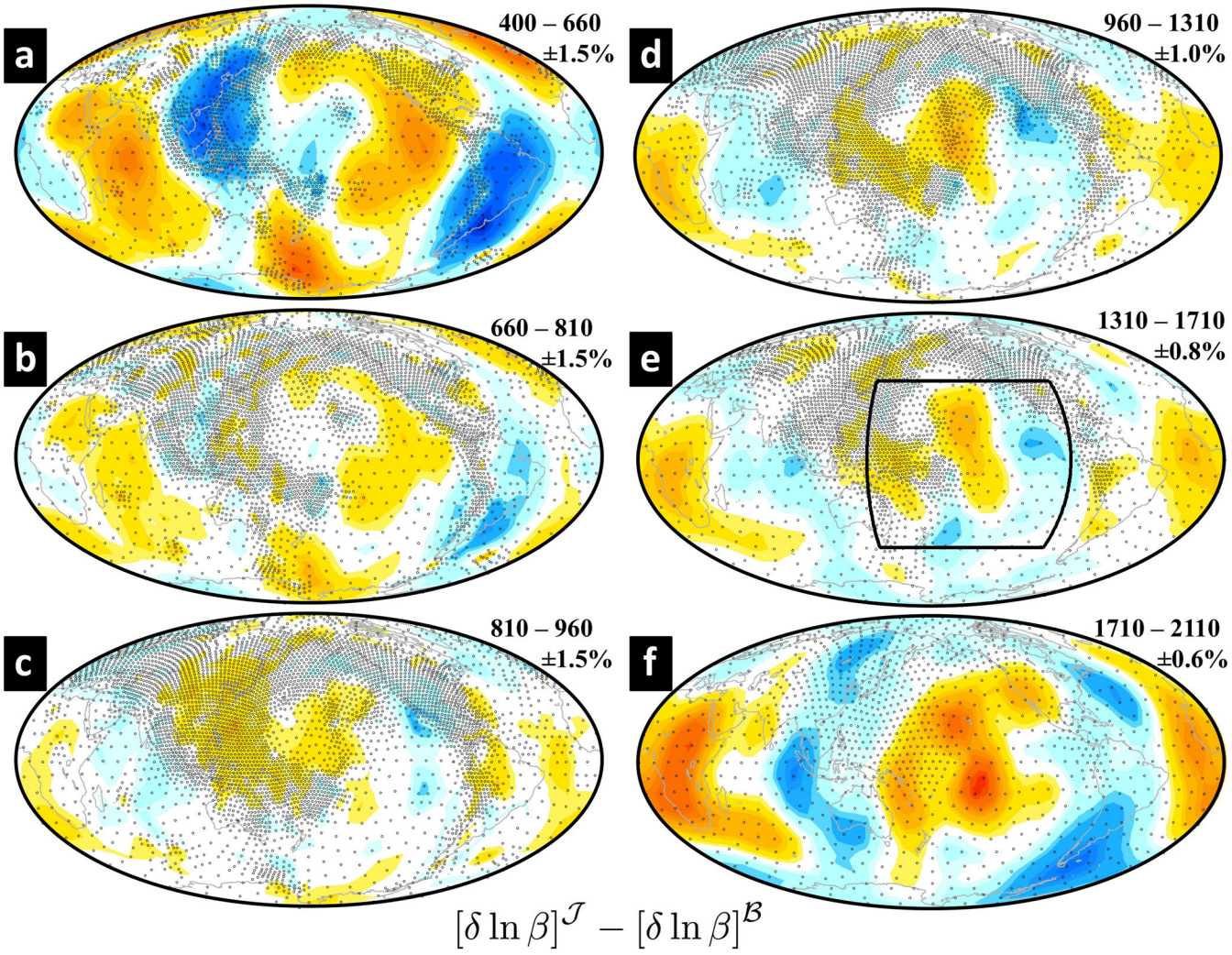

Figure 8. Analysis, in the model space, of the effects of incorporating in the joint inversion the normal modes in addition to the body waves. The model difference, between our joint and body-wave models, is plotted at different depths. Superimposed with black circles is the grid of nodes—which is tuned to the ray density. The nodes corresponding to the layers (5, 9, 11, 13) are shown in panels (a), (d), (e) and (f), respectively. Note that the long-wavelength sensitivity of normal modes helps to remedy the local lack of body-wave information, that is, where the grid is coarse. Same colour scales as in Figs 5–7.

roughly extending from Tahiti to Hawaii seems to be constrained by normal modes. Fig. 9 shows that significative changes related to normal modes occur in the joint model beneath Hawaii. A clear benefit is then to use the global-scale, long-wavelength information of normal modes in mantle regions where body-wave information is lacking.

6.2 Long-wavelength seismic features

For completeness, we briefly point out some long-wavelength features in our joint model (see Fig. 7). There is a significant contrast in the pattern of shear-velocity structure across the 660 km depth discontinuity (e.g. Gu *et al.* 2001; Ritsema *et al.* 2004, 2011). Low-velocity anomalies are well apparent, and often located nearby a known hotspot (Anderson & Schramm 2005). In particular, our model features in the entire 400–2110 km depth range a vertically continuous and strong low-velocity structure, located close to the Samoa hotspot, that could potentially be interpreted as a plume conduit because of its relatively small lateral extent (e.g. Montelli *et al.* 2006; Takeuchi 2009). Some low-velocity anomalies appear to be located under and parallel to several major slabs, within the transition-zone and uppermost lower-mantle. These sub-slab slow anomalies seem to be more pronounced in our model for large subduction-zone trenches: Java, Kuril, Peru-Chile, Tonga-Kermadec; no consensus seems to have emerged about the geodynamical processes behind (e.g. Fukao *et al.* 2001; Zhao 2004, 2012). Long-wavelength high-velocity structures are seen into circum-Pacific regions and under Asia, in the mid lower-mantle at about 960–2110 km depth, such as the seismic signatures of the Tethys remnants beneath Mediterranean/Southern Eurasia (e.g. Van der Hilst & Karason 1999) and of the ancient Farallon slab beneath North America (e.g. Grand *et al.* 1997).

6.3 The Farallon subduction beyond degree 40

We aim at pointing out some short-scale structure components in our joint model that would be out of reach if using a degree-40 lateral parametrization. As shown in Fig. 10, we shall focus on the Farallon region within the 660–1710 km depth range. In this region of interest, our tomographic grid is highly refined, as can be seen in Figs 8(b)–(e), so that the smallest lateral resolving-length

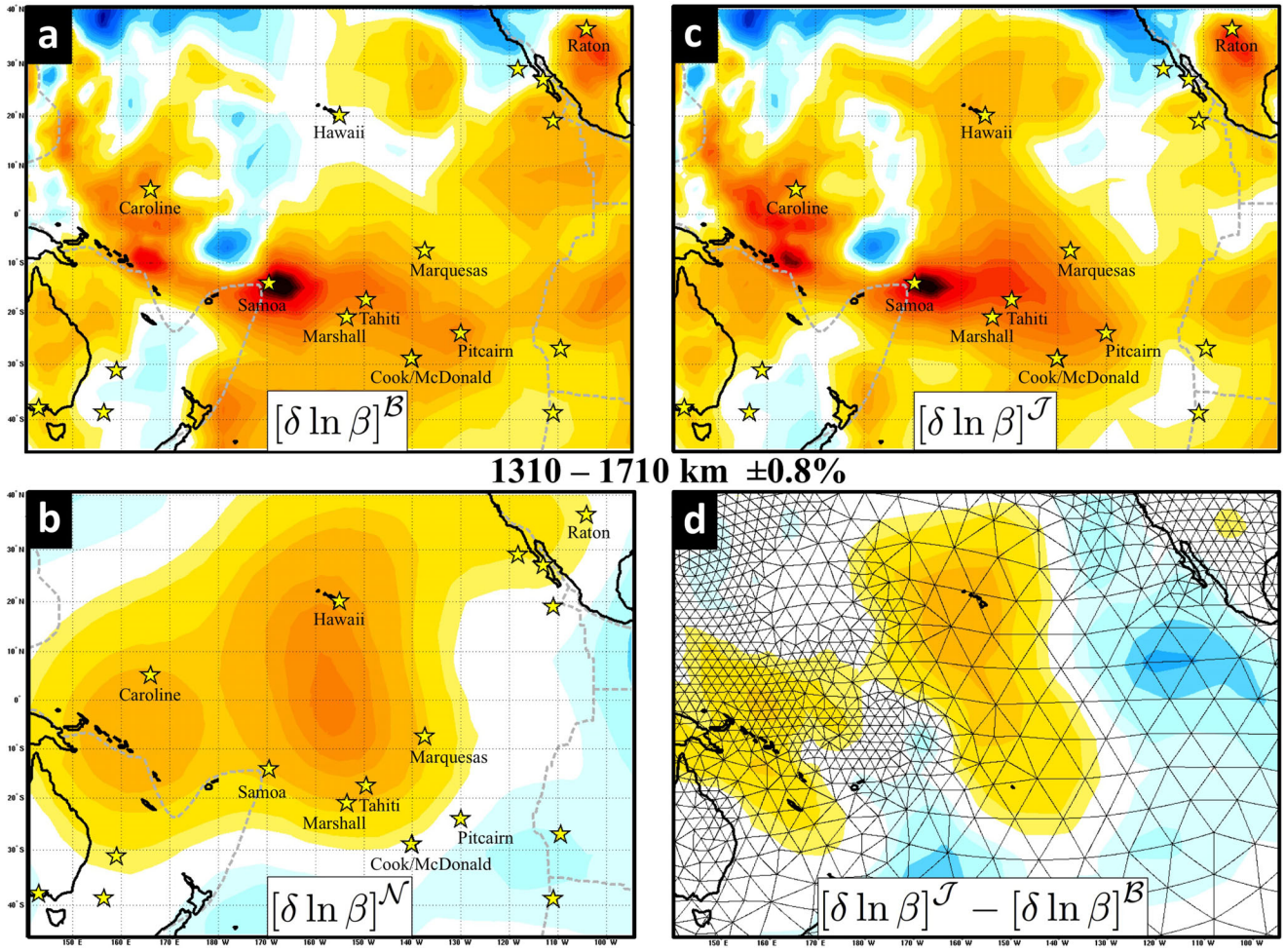

Figure 9. Illustration showing how the joint inversion attempts to satisfy both body-wave and normal-mode data. Zoom-in on the Central Pacific area, as outlined with a black frame in Fig. 8(e), within the 1310–1710 km depth range. (a) Body-wave model; (b) normal-mode model; (c) joint model; (d) difference between joint and body-wave models, superimposed to the irregular parametrization (corresponding to layer 11). Though our *S*-wave coverage cannot resolve the surroundings of Hawaii at such depth (coarse grid), one sees that a large-scale low-velocity feature roughly extending from Tahiti to Hawaii seems to be constrained by normal-mode data. As expected, significative changes between the joint model and the body-wave model mainly occur beneath Hawaii, where normal modes help to remedy the lack of body-wave information. Grey dashed line: tectonic plates; yellow stars: hotspots.

is potentially, at best, about 200 km for L_{grid}, while it varies with depth from 448 to 366 km for L_{sh40} (*cf.* Sections 3.1 and 3.2).

In addition to the long-wavelength ancient Farallon slab, mainly located beneath Eastern North-America, our model features a detached, thin and elongated slab fragment beneath Central North-America, referred to as F1 in Figs 10(c)–(e), within the 960–1510 km depth range. Another detached slab fragment, referred to as F2 in Fig. 10(e), appears at 1310–1510 km depth. Moreover, at 1510–1710 km depth, both features F1 and F2 seem to collapse together into one broad-scale detached slab, denoted as F2+F1 in Fig. 10(f). Such a complex Farallon subduction system, taking place in mid lower-mantle, has apparently not been so clearly and sharply identified in previously cited global-scale tomographic models. It could be coherent, at first glance, with some regional-scale *P*- and *S*-wave tomographic studies (e.g. Sigloch *et al.* 2008; Sigloch 2011; Tian *et al.* 2011).

Several short-scale features, including F1, are marked with dashed-line ellipses in Figs 10(a)–(c), within the 660–1110 km depth range on the west of the main Farallon slab. Inset frames on the left-hand side of Figs 10(a)–(c) show the corresponding joint model from which all the spherical-harmonic components with degree superior to 40 have been filtered out. Note that all the marked short-scale features, such as the elongated slab fragment F1 at 960–1110 km depth, cannot be retrieved with a spherical-harmonics parametrization limited to highest degree 40. Off course, this does not mean that those small-size features are robustly resolved in our joint model, but if they were, they could not be seen without using a lateral parametrization that goes well beyond degree 40.

As a final remark, further interpretations of the previously mentioned tomographic features would require a detailed model resolution analysis, to assess whether they are robust or not. That is out of the scope of this study. Indeed, we plan to further improve the presented model by adding surface-wave data and extending our sets of normal-mode and body-wave data. Doing so will necessarily affect the resolution in the whole mantle. Thus, in our view, it seems premature to make a model resolution analysis at this stage.

6.4 Perspectives

To further improve the presented joint model, other seismological data should be incorporated, as for instance: (1) surface waves, to better constrain the upper-mantle (e.g. Debayle & Ricard 2012); (2) other body waves, such as *ScS*, to better illuminate the

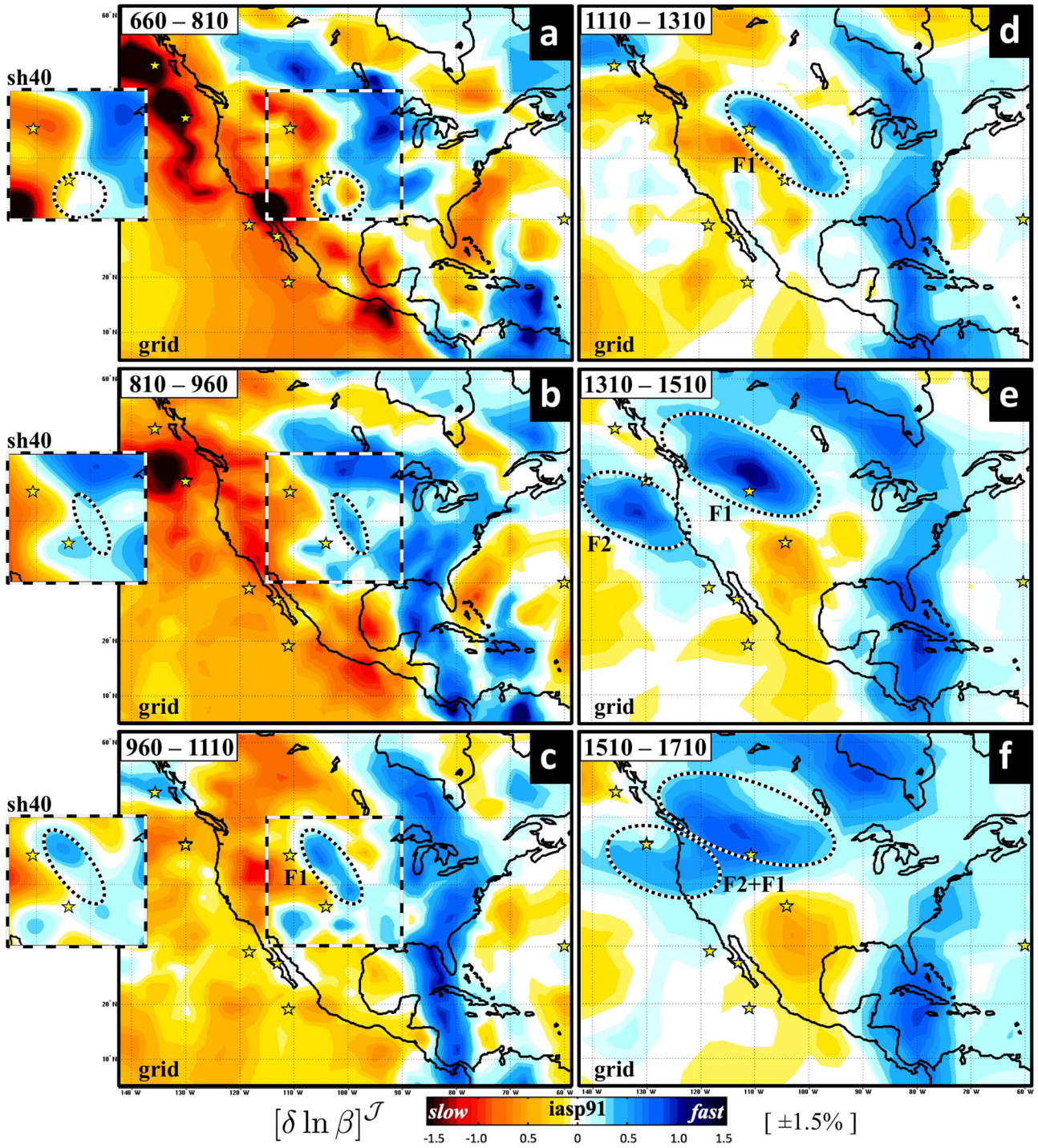

Figure 10. On the relevance of using an irregular parametrization for a joint inversion of normal-mode and body-wave data. Zoom-in on the Farallon subduction, as imaged in our joint model within the 660–1710 km depth range. Features F1, F2 and F2+F1 are described in Section 6.3. Several short-scale features are marked with dashed-line ellipses in panels (a)–(c). Inset frames, on the left-hand side, show the corresponding joint model from which all the spherical-harmonic components of degree greater than 40 have been filtered out. Note that the features marked with ellipses, such as the elongated slab fragment F1 at 960–1110 km depth, cannot be seen without using a lateral parametrization that goes well beyond degree 40.

lowermost mantle (Zaroli *et al.* 2010); (3) mode-coupling of the Earth's free oscillations to better constrain the long-wavelength uneven degree structure of the mantle (e.g. Resovsky & Ritzwoller 1999), as well as other single spheroidal modes. Increasing our data sets will imply to compute an updated tomographic grid that is adapted to the new data constraints. Moreover, as shown in Figs 4(c) and (g), to better exploit finite-frequency effects in body-wave (and/or surface-wave) data, the minimum spacing of nodes should ideally be further decreased, to deteriorate less the corresponding sensitivity kernels. Note that refining the irregular grid to potentially reach about 50–100 km of lateral resolving-length in particular mantle regions, such as subduction zones, would only imply a moderate increase of the total number of nodes, while it could be out of reach if using uniform basis functions such as the

spherical harmonics. Finally, it would be interesting to test the performance of alternative parametrizations, such as those based on wavelets (e.g. Simons *et al.* 2011; Chevrot *et al.* 2012; Charléty *et al.* 2013) or harmonic spherical splines (e.g. Amirbekyan *et al.* 2008), which are still in their infancy for global-scale tomographic purposes.

7 CONCLUSION

Global-scale tomographic models should aim at satisfying the full seismic spectrum, from the lowest to the highest frequencies. In this study, we have built a model of isotropic 3-D variations of shear velocities in the mantle, derived from data including periods in the 332–2134 s range for normal modes and 10–51 s for body waves. Though spherical harmonics naturally appear when considering the Earth's free oscillations, we have shown that progress towards higher resolution joint tomography requires a movement away from such uniform parametrization to overcome its computational inefficiency to adapt to local variations in resolution. The heart of this work has been to show how to include, for the first time, normal modes into a joint inversion based upon a non-uniform lateral parametrization, using an irregular tomographic grid optimized according to ray density. It has essentially consisted in efficiently computing the projection of 3-D normal-mode sensitivity kernels onto a parametrization made up of spherical layers spanned with irregular Delaunay triangulations.

Tomographic results have been focused on the 400–2110 km depth range, where our current data coverage is the most relevant. The obtained joint model does not significantly deteriorate the fit of one data set to the detriment of the other, which demonstrates our ability to map into the model multiscale structural informations from data at both ends of the seismic spectrum. For instance, the long-wavelength sensitivity of normal modes helps to remedy the local lack of body-wave information. We have discussed the potential of a better resolution where the grid is fine, compared to spherical harmonics up to degree 40, as the number of model parameters is similar. Our joint model seems to contain coherent structural components beyond degree 40, such as those related to the complex Farallon subduction system.

We conclude that using such an irregular grid, locally adapted to the spatially varying resolving length of the data, is a computationally efficient approach for building new global-scale tomographic models that will better satisfy the full seismic spectrum. A wider application of this joint inversion workflow should then primarily consist in adding surface-wave data and extending our sets of normal-mode and body-wave data, which holds promise for a better understanding of the Earth's interior at various spatial scales.

ACKNOWLEDGEMENTS

The authors thank Jeroen Ritsema, Jeannot Trampert and Kaiqing Yuan for helpful reviews that improved the original manuscript. CZ and SL received support from the SEISGLOB ANR 2011 Blanc SIMI 5-6-016-01.

REFERENCES

Amirbekyan, A., Michel, V. & Simons, F., 2008. Parameterizing surface wave tomographic models with harmonic spherical splines, *Geophys. J. Int.,* **174,** 617–628.

Anderson, D.L. & Schramm, K.A., 2005. Global hotspot maps, in *Plates, Plumes and Paradigm,* pp. 19–30, eds Foulger, G.R., Natland, J.H., Presnall, D.C. & Anderson, D.L., Geological Society of America.

Aster, R.C., Borchers, B. & Thurber, C., 2012. *Parameter Estimation and Inverse Problems,* Elsevier.

Barber, C.B., Dobkin, D.P. & Huhdanpaa, H.T., 1996. The quickhull algorithm for convex hulls, *ACM Trans. Math. Softw.,* **22,** 469–483.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS, Trans. Am. geophys. Un.,* **81**, F897.

Charléty, J., Voronin, S., Nolet, G., Loris, I., Simons, F., Sigloch, K. & Daubechies, I.C., 2013. Global seismic tomography with sparsity constraints: Comparison with smoothing and damping regularization, *J. geophys. Res.,* **118,** 1–13.

Chevrot, S., Martin, R. & Komatitsch, D., 2012. Optimized discrete wavelet transforms in the cubed sphere with the lifting scheme: implications for global finite-frequency tomography, *Geophys. J. Int.,* **191,** 1391–1402.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton University Press.

Dahlen, F.A., Hung, S.-H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.,* **141,** 157–174.

Debayle, E. & Ricard, Y., 2012. A global shear velocity model of the upper mantle from fundamental and higher rayleigh mode measurements, *J. geophys. Res.,* **117**(B10308), doi:10.1029/2012JB009288.

Deuss, A., Ritsema, J. & van Heijst, H.-J., 2011. Splitting function measurements for Earth's longest period normal modes using recent large earthquakes, *Geophys. Res. Lett.,* **38**(4), L04303, doi:10.1029/2010GL046115.

Deuss, A., Ritsema, J. & van Heijst, H.-J., 2013. A new catalogue of normal-mode splitting function measurements up to 10 mHz, *Geophys. J. Int.,* **193**(2), 920–937.

Dziewonski, A.M. & Anderson, D., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.,* **25,** 297–356.

Edmonds, A.R., 1996. *Angular Momentum in Quantum Mechanics,* Princeton University Press.

Fukao, Y., Widiyantoro, S. & Obayashi, M., 2001. Stagnant slabs in the upper and lower mantle transition region, *Rev. of Geophys.,* **39,** 291–323.

Grand, S.P., Van der Hilst, R.D. & Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the Earth, *GSA Today,* **7,** 1–7.

Grunberg, M., 2006. Conception d'une méthode de maillage 3D parallèle pour la construction d'un modèle de Terre réaliste par la tomographie sismique, *PhD thesis,* Strasbourg University, France.

Gu, Y.J., Dziewonski, A.M., Su, W.-J. & Ekström, G., 2001. Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities, *J. geophys. Res.,* **106,** 11 169–11 199.

Hansen, C. & O'leary, D., 1993. The use of the L-curve in the regularization of discrete ill-posed problems, *SIAM J. Sci. Comput.,* **14**(6), 1487–1503.

Hung, S.-H., Shen, Y. & Chiao, L.-Y., 2004. Imaging seismic velocity structure beneath the Iceland hotspot: a finite-frequency approach, *J. geophys. Res.,* **109,** B08305, doi:10.1029/2003JB002889.

Ishii, M. & Tromp, J., 2001. Even-degree lateral variations in the Earth's mantle constrained by free oscillations and the free-air gravity anomaly, *Geophys. J. Int.,* **145**(1), 77–96.

Kennett, B. & Engdahl, E., 1991. Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.,* **105,** 429–465.

Lévêque, J.-J., Rivera, L. & Wittlinger, G., 1993. On the use of the checkerboard test to assess the resolution of tomographic inversion, *Geophys. J. Int.,* **115,** 313–318.

Li, X.D., Giardini, D. & Woodhouse, J.H., 2001. Large scale three dimensional even degree structure of the Earth from splitting of long period normal modes, *J. geophys. Res.,* **96,** 551–577.

Maceira, M., Larmat, C., Porritt, R.W., Higdon, D.M., Rowe, C.A. & Allen, R.M., 2015. On the validation of seismic imaging methods: finite-frequency or ray theory?, *Geophys. Res. Lett.,* **42**(2), 323–330.

Masters, G., Laske, G., Bolton, H. & Dziewonski, A.M., 2000a. The relative behaviour of shear velocity, bulk sound speed and compressional velocity in the mantle: implications for chemical and thermal structure, in: *Earth's Deep Interior,* pp. 63–88, eds Karato, S., Forte, A., Liebermann, R.C., Masters, G. & Stixrude, L., AGU.

Masters, G., Laske, G. & Gilbert, F., 2000b. Autoregressive estimation of the splitting matrix of free-oscillation multiplets, *Geophys. J. Int.,* **141**(1), 25–42.

Masters, G., Laske, G. & Gilbert, F., 2000c. Matrix autoregressive analysis of free-oscillation coupling and splitting, *Geophys. J. Int.,* **143**(2), 478–489.

Mercerat, E. D., Nolet, G. & Zaroli, C., 2014. Cross-borehole tomography with correlation delay times, *Geophys.,* **79**(1), R1–R12.

Michelini, A., 1995. An adaptive-grid formalism for traveltime tomography, *Geophys. J. Int.,* **121,** 489–510.

Montelli, R., Nolet, G., Dahlen, F.A., Masters, G., Engdahl, E.R. & Hung, S.-H., 2004a. Finite-frequency tomography reveals a variety of plumes in the mantle, *Science,* **303,** 338–343.

Montelli, R., Nolet, G., Masters, G., Dahlen, F.A. & Hung, S.-H., 2004b. Global P and PP traveltime tomography: rays versus waves, *Geophys. J. Int.,* **158,** 636–654.

Montelli, R., Nolet, G., Dahlen, F.A. & Masters, G., 2006. A catalogue of deep mantle plumes: new results from finite-frequency tomography, *Geochem. Geophys. Geosyst.,* **7**(11), doi:10.1029/2006GC001248.

Moulik, P. & Ekström, G., 2014. An anisotropic shear velocity model of the Earth's mantle using normal modes, body waves, surface waves and long-period waveforms, *Geophys. J. Int.,* **199,** 1713–1738.

Nolet, G., 2008. *A Breviary of Seismic Tomography,* Cambridge Univ. Press.

Nolet, G. & Montelli, R., 2005. Optimum parameterization of tomographic models, *Geophys. J. Int.,* **161,** 365–372.

Paige, C.C. & Saunders, M., 1982. LSQR: an algorithm for sparse, linear equations and sparse least squares, *ACM Trans. Math. Softw.,* **8,** 43–71.

Rawlinson, N., Pozgay, S. & Fishwick, S., 2010. Seismic tomography: a window into deep Earth, *Phys. Earth planet. Inter.,* **178**(3–4), 101–135.

Resovsky, J. & Ritzwoller, M.H., 1998. New and refined constraints on the three-dimensional Earth structure from normal-modes below 3 mHz, *J. geophys. Res.,* **103**(B1), 783–810.

Resovsky, J. & Ritzwoller, M.H., 1999. A degree 8 mantle shear velocity model from normal mode observations below 3 mHz, *J. geophys. Res.,* **104**(B1), 993–1014.

Ritsema, J., van Heijst, H.-J. & Woodhouse, J.H., 1999. Complex shear wave velocity structure imaged beneath Africa and Iceland, *Science,* **286,** 1925–1928.

Ritsema, J., van Heijst, H.-J. & Woodhouse, J.H., 2004. Global transition zone tomography, *J. geophys. Res.,* **109,** doi:10.1029/2003JB002610.

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,** 1223–1236.

Sambridge, M. & Rawlinson, N., 2005. Seismic tomography with irregular meshes, in *Seismic Earth: Array Analysis of Broadband Seismograms,* vol. 157, pp. 49–65, eds Levander, A. & Nolet, G., AGU.

Schuberth, B.S.A., Zaroli, C. & Nolet, G., 2012. Synthetic seismograms for a synthetic Earth: long-period P- and S-wave traveltime variations can be explained by temperature alone, *Geophys. J. Int.,* **188**(3), 1393–1412.

Schuberth, B.S.A., Zaroli, C. & Nolet, G., 2015. Traveltime dispersion in an isotropic elastic mantle: strong lower mantle signal in differential-frequency residuals, *Geophys. J. Int.,* doi:10.1093/gji/ggv389.

Sigloch, K., 2011. Mantle provinces under North America from multi-frequency P-wave tomography, *Geochem. Geophys. Geosyst.,* **12,** Q02W08, doi:10.1029/2010GC003421.

Sigloch, K. & Nolet, G., 2006. Measuring finite-frequency body-wave amplitudes and traveltimes, *Geophys. J. Int.,* **167,** 271–287.

Sigloch, K., McQuarrie, M. & Nolet, G., 2008. Two-stage subduction history under North America inferred from multiple-frequency tomography, *Nature Geosci.,* **1,** 458–462.

Simmons, N.A., Myers, S.C. & Ramirez, A., 2009. Multi-resolution seismic tomography based on recursive tessellation hierarchy, in *Proceedings of the 2009 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies,* LA-UR-09-05276, Vol. 1, pp. 211–220.

Simons, F., Nolet, G., Babcock, J., Davis, R. & Orcutt, J., 2006. A future for drifting seismic networks, *EOS, Trans. Am. geophys. Un.,* **31,** 305–307.

Simons, F., Loris, I., Nolet, G., Daubechies, I.C., Voronin, S., Judd, J.S., Vetter, P.A., Charlety, J. & Vonesch, C., 2011. Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity, *Geophys. J. Int.,* **187,** 969–988.

Spakman, W. & Bijwaard, H., 2001. Optimization of cell parameterizations for tomographic inverse problems, *Pure appl. Geophys.,* **158,** 1401–1423.

Spakman, W. & Nolet, G., 1988. Imaging algotithms, accuracy and resolution in delay-time tomography, in *Mathematical Geophysics,* pp. 155–187, eds Vlaar, N.J., Nolet, G., Wortel, M.J.R. & Cloetingh, S.A.P.L., Reidel.

Takeuchi, N., 2009. A low-velocity conduit throughout the mantle in the robust component of a tomographic model, *Geophys. Res. Lett.,* **36,** L07306, doi:10.1029/2009GL037590.

Tarantola, A., 1987. Inversion of traveltimes and seismic waveforms, in *Seismic Tomography*, pp. 135–157, ed. Nolet, G., Reidel.

Tarantola, A. & Nercessian, A., 1984. Three-dimensional inversion without blocks, *Geophys. J. R. astr. Soc.,* **76,** 299–306.

Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.,* **20,** 213–232.

Tian, Y., Zhou, Y., Sigloch, K., Nolet, G. & Laske, G., 2011. Structure of North American mantle constrained by simultaneous inversion of multiple-frequency SH, SS, and Love waves, *J. geophys. Res.,* **116,** B02307, doi:10.1029/2010JB007704.

Trampert, J., 1998. Global seismic tomography: the inverse problem and beyond, *Inverse Probl.,* **14,** 371–385.

Trampert, J. & Woodhouse, J.H., 2003. Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s, *Geophys. J. Int.,* **154**(1), 154–165.

Van der Hilst, R.D. & Karason, H., 1999. Compositional heterogeneity in the bottom 1000 kilometers of Earth's mantle: toward a hybrid convection model, *Science,* **283**(5409), 1885–1888.

Vasco, D.W., Johnson, L.R. & Marques, O., 2003. Resolution, uncertainty, and whole Earth tomography, *J. geophys. Res.,* **108,** doi:10.1029/2001JB000412.

Widmer-Schniring, R., 2002. Application of regionalized multiplet stripping to retrieval of aspherical structure constraints, *Geophys. J. Int.,* **148,** 201–213.

Woodhouse, J.H. & Dahlen, F.A., 1978. The effect of a general aspherical perturbation on the free oscillations of the Earth, *Geophys. J. R. astr. Soc.,* **53,** 335–354.

Woodhouse, J.H. & Giardini, D., 1985. Inversion for the splitting function of isolated low order normal mode multiplets, *EOS, Trans. Am. geophys. Un.,* **66**, 300.

Wu, Y., He, Y. & Tian, H., 2005. A spherical point location algorithm based on barycentric coordinates, in *Computational Science and Its Applications – ICCSA 2005,* LNCS 3482, pp. 1099–1108, eds Gervasi, O., Gavrilova, M.L., Kumar, V., Laganà, A., Lee, H.P., Mun, Y., Taniar, D. & Tan, C.J.K., Springer.

Yang, T., Shen, Y., van der Lee, S., Solomon, S. & Hung, S.-H., 2006. Upper mantle beneath the Azores hotspot from finite-frequency seismic tomography, *Earth Planet. Sci. Lett.,* **250,** 11–26.

Zaroli, C., 2010. Global multiple-frequency S-wave tomography of the Earth's mantle, *PhD thesis,* Strasbourg University, France.

Zaroli, C., Debayle, E. & Sambridge, M., 2010. Frequency-dependent effects on global S-wave traveltimes: wavefront-healing, scattering and attenuation, *Geophys. J. Int.,* **182,** 1025–1042.

Zaroli, C., Sambridge, M., Lévêque, J.-J., Debayle, E. & Nolet, G., 2013. An objective rationale for the choice of regularisation parameter with application to global multiple-frequency S-wave tomography, *Solid Earth,* **4,** 357–371.

Zaroli, C., Lévêque, J.-J., Schuberth, B.S.A., Duputel, Z. & Nolet, G., 2014. Global S-wave tomography using receiver pairs: an alternative to get rid of earthquake mislocation, *Geophys. J. Int.,* **199,** 1043–1057.

Zhao, D., 2004. Global tomographic images of mantle plumes and subducting slabs: insight into deep Earth dynamics, *Phys. Earth Planet. Inter.,* **146,** 3–34.

Zhao, D., 2012. Tomography and dynamics of Western-Pacific subduction zones, *Monog. Environ. Earth Planets,* **1,** 1–70.